**TELKOMNIKA Indonesian Journal of Electrical Engineering**
Vol. 12, No. 9, September 2014, pp. 6725 ~ 6731
DOI: 10.11591/telkomnika.v12i9.4593


# Intelligent Train Operation Models Based on Ensemble Regression Trees

**Dewang Chen*, Xiangyu Zeng, Guiwen Jia**
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University,
Beijing, 100044, China.
*Corresponding author, e-mail: dwchen@bjtu.edu.cn, 11120319@bjtu.edu.cn, 10125057@bjtu.edu.cn

***Abstract***

*Traditional control algorithms in Automatic Train Operation (ATO) system have some drawbacks, such as high energy consumption and low riding comfort. Combined with data mining methods and driving experience, two Intelligent Train Operation (ITO) models for the subway train control are proposed. Firstly the training data set was sorted out and sieved out from the real train operation data set by drivers in Beijing subway line Yizhuang to establish the standard database. By using Classification and Regression Trees (CART) algorithm and Bagging ensemble learning method which base on CART algorithm, two ITO models are dug out to represent the output of controller with limited speed, running time and gradient. In the train control simulation platform, ITO models were compared with the traditional PID (Proportional Integral Derivative) control algorithm of ATO systems. The simulation results indicate the proposed ITO models are better than PID control in energy consumption, riding comfort and switching times of controller's output. Furthermore, the ITO model with bagging ensemble learning method is better especially in energy consumption and riding comfort.*

***Keywords****: ensemble Learning, regression trees, data mining, automatic train operation, intelligent train operation*



## 1. Introduction

Control technology plays an important role in maintaining safe, reliable, and cost-effective operation of trains. Automatic Train Operation (ATO) is responsible for all the traction and braking controls.Facing with such real-time dynamic operational requirements, intelligent control strategies came into play in the 1980s. PID (Proportional Integral Derivative) control has been widely used in industrial control system because of its simple structure and robust performance, also it can be used in ATO system [1], Genetic algorithm (GA) is proposed to optimize train movements using appropriate coast control that can be integrated within ATO systems [2] or constructed optimal train driving strategy [3]. Least square estimation and an adaptive network based fuzzy inference system (ANFIS) were presented to estimate the train station parking error in urban rail transit [4-5]. However, these control methods are limited to the traditional control theory, which takes control accuracy as the main goal in the process of tracking operation speed curve, in order to make the real speed curve as close to the optimal one as possible. Moreover the controller of ATO systems need frequently switching in the process of train operation, which is not conductive to the riding comfort and energy saving, the life of controller is also greatly reduced in the meantime.

Different from ATO systems, experienced drivers can operate the train to the specified location on time with a few time of handle changing smoothly. A large amount of the data are generated by human drivers in the process of manual train control. From the calculation of actual data, we can find that manual driving by experienced drivers is better than automatic driving in energy consumption and riding comfort. In order to get better control effect, we are trying to find the intelligent train operation (ITO) model, the driving model mining from the large amount of manual driving data by data mining techniques [6].

*Received October 6, 2013; Revised Apr 21, 2014; Accepted May 6, 2014*

## 2. Collection of Field Data

There are two basic driving mode in subway train, automatic driving mode and manual driving mode. In this paper, field data in two days (20 times) from Yizhuang Line Beijing Subway were collected. We choose one block from Xiaohongmen station to Xiaocun station as an example. For this block, we collect 11235 groups of samples in two days, and choose 8 useful attributes in each sample. The eight attributes are limited speed, gradient, train speed, remaining time, remaining distance, changing value of next limited speed, remaining distance of next limited speed and controller's output (from -1 to 1, positive is traction, negative is braking, zero is idle running). The first seven attributes are used as input variables, and the last attribute is used as the output variable.

It is necessary to get the data with good performance from the obtained massive field data. The reason is, some of the data is produced by high-level drivers but some is produced by middle-level or low-level drivers. Moreover, the drivers may be affected by psychological or physical conditions, for example, long time operation may lead to fatigue and pressure. That is to say we need to pick out the data with low energy consumption, high riding comfort, and low running time error.

According to the statistical results for the manual driving data set, the variation of time error is [-5.2 7.8](s), the variation of switching times of controller's output is [4 16], the variation of impingement rate is [0.087 0.146]($m/s^3$) and the variation of energy consumption is [197.32 216.63](J). By trial-and-error, we set the following four rules to select data which satisfying the all rules.

Rule 1 Time error is within $\pm 5s$ ;

Rule 2 Switching times of controller's output is within 10 times;

Rule 3 Impingement rate is within 0.12 $m/s^3$;

Rule 4 Energy consumption is within 210J.

Through the above rules, 7312 groups of samples are sorted out for data mining.

## 3. Regression Trees and its Improvement by Ensemble Learning

### 3.1. Classification and Regression Trees Algorithm

Due to the multi-variable and large amount of data for the regression problem in this paper, traditional regression methods are not applicable. So we use CART (Classification and Regression Trees) algorithm to solve it.

CART algorithm [7] was proposed by Breiman, Friedman, Olshen and Stone in 1984. The letters CART indicate that trees may be used not only to classify entities into a discrete number of groups, but also as an alternative approach to regression analysis in which the value of a response (dependent) variable is to be estimated, given the value of each variable in a set of explanatory (independent) variables.

But single CART algorithm can not achieve better results. To improve its performance , an ensemble learning algorithm is used with it.

### 3.2. Eesemble Learning

In 1997,T.G.Dietterich, authority in the field of machine learning, put ensemble learning in the first place of four research directions of machine learning [8].Resemble learning is one of the research hot spots in machine learning in recent years, and the main achievements are as follows: Bagging [9], Boosting [10], Random Forest [11] and so on. The main idea of resemble learning is training multiple weak learning systems and combining the results in a certain way, which can significantly improve the generalization ability of the learning systems. The brief introduction of bagging algorithm is shown as follow.

The bagging algorithm was proposed by Breiman in 1996 [9]. During the training phase of the algorithm, we repeatedly samples the original training samples with replication so that we get a new set of training samples.

Select training samples from the original set of training samples at random; train the samples using the given base learning algorithm, then we can get a model; put back the training samples. Repeat k times, so that we can get a set of k models. As for the regression problem, we can obtain a final forecasting model as follows:

$$F(x_i) = \frac{1}{K}\sum_{k=1}^{K} f_k(x_i) \quad (1)$$

The bagging algorithm is the most simple and intuitive ensemble learning method. Theoretically, when using the Bagging algorithm, about 36.8% of the samples will not appear in the new set of training samples averagely as we resample each time.

### 3.3. Algorithms Implementation

The bagging ensemble learning algorithm can be used for regression analysis. In this paper, CART algorithm is used as the weak learning machine of bagging algorithm, we call this algorithm as B-CART, and we can also use a single CART algorithm for analyzing. Here we use these two methods to mine the ITO models from the data. After that, we will analyze and compare the results.

B-CART is a method integrating the regression trees, and it can set the iteration times. Here, we set the iteration times ranging from 1 to 100. The mean absolute error is shown in Figure 1.

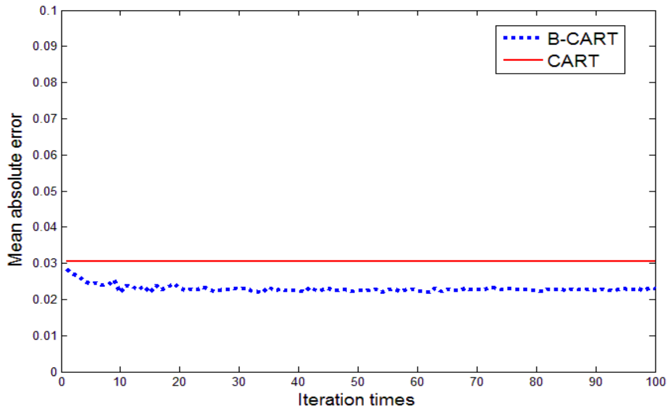


Figure 1. Mean Absolute Error Changes with Iteration Times

In Figure 1, compared with the results of CART algorithm, B-CART algorithm can always be superior to it. The results obtained are coincident to the theoretical analysis. With the iteration times increasing, the calculating speed of the algorithm will be greatly reduced. So we set the iteration times as 50 in our simulation.

## 4. ITO Model and its Simulating Platform

### 4.1. Operation Requirements for each Stage

Safety is important in the train operation, and the most basic requirement for it is that the running speed should not exceed the limited speed. So we divide the train operation into 4 stages, acceleration stage, idle running stage, deceleration stage and stopping stage. The requirements for each stage are different.

Acceleration stage and Idle running stage: If $V \geq 0.95 * V_{\max}$, then $a = 0$.

Deceleration stage: If $V \geq 0.9 * V_{\max}$, then $a = -0.5$. Where *V* is speed, $V_{max}$ is limited speed, $a$ is controller's output.

### 4.2. Simulation Platform

The ITO model simulation platform is established with Matlab Simulink. The platform includes five modules, input module, generator module, controller module, actuator module and display module. Figure 2 is the structure graph of ITO model.

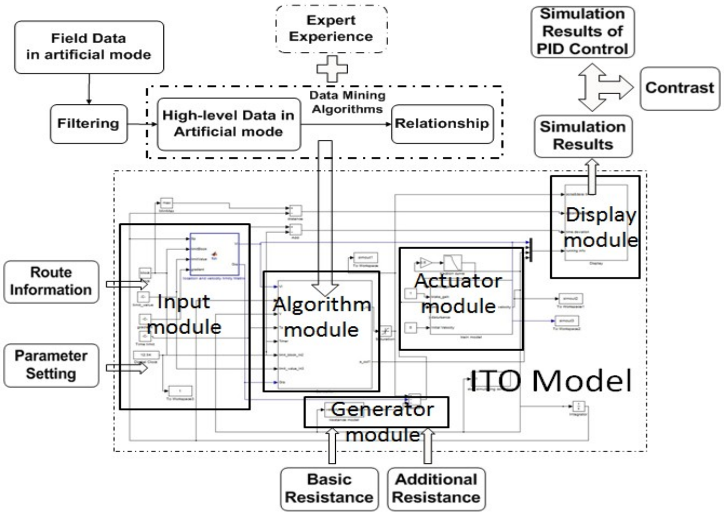


Figure 2. Structure Graph of ITO Model

## 5. Simulation and Evaluation

We also choose block 1 as an example. The distance of block 1 is 1341m, and the standard running time is 105s. The regression models mined by CART and B-CART are used in the ITO models, and the regression model mined by CART and B-CART are shown in Figure 3 and Figure 4.

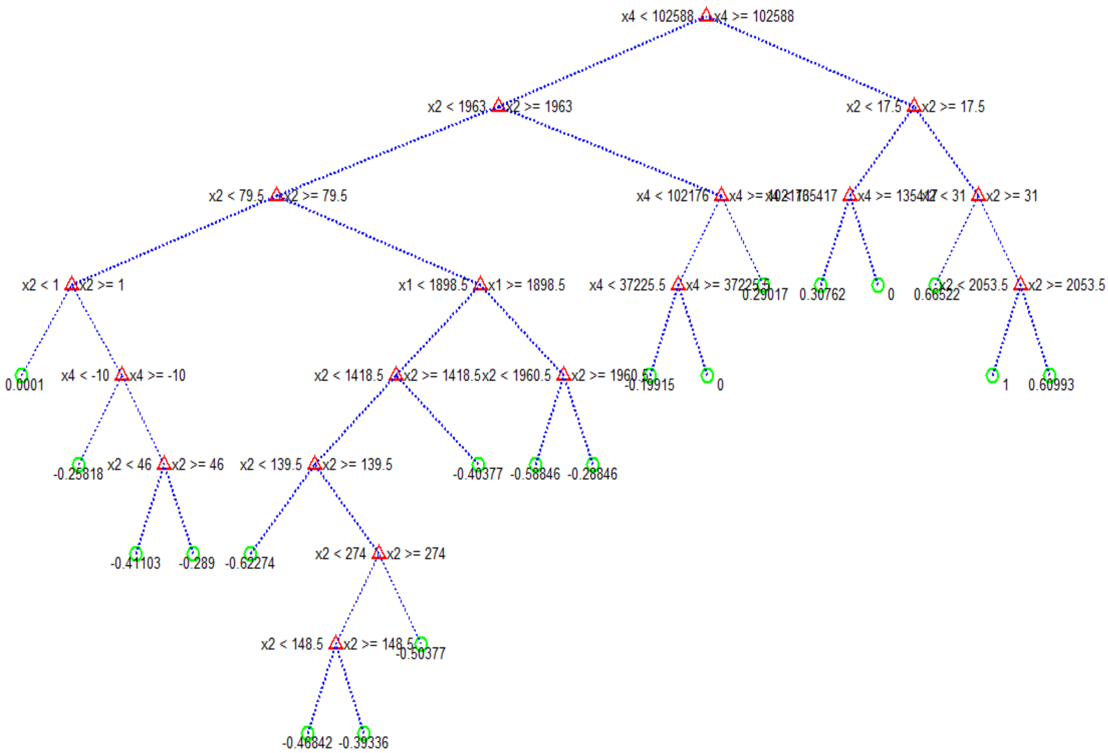


Figure 3. Regression Model of CART Algorithm

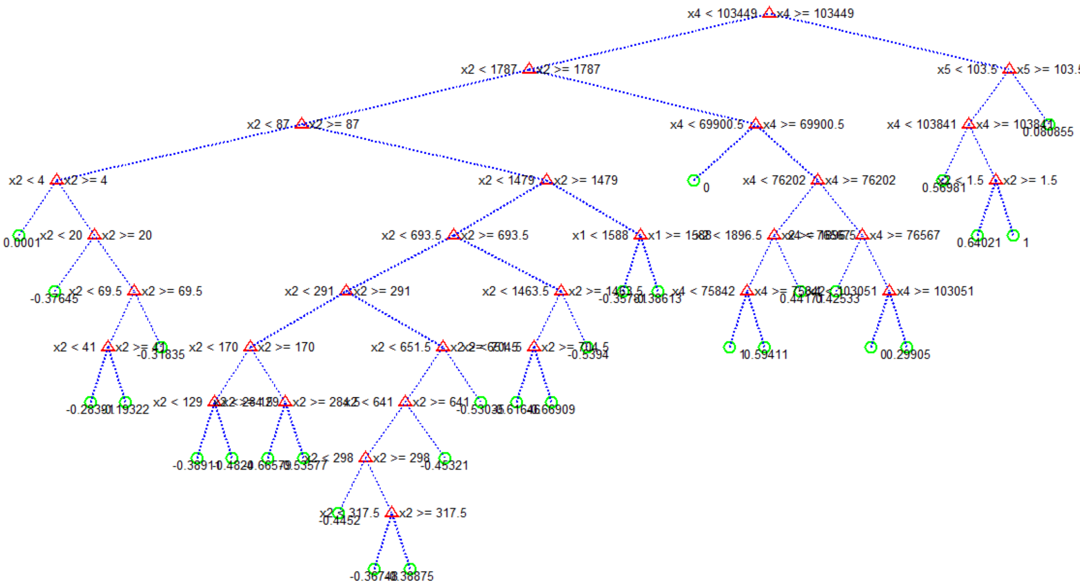

Figure 4. Regression Model of B-CART Algorithm

As can be seen in Figire 3 and Figure 4, the x1,x2,...,x7 represent limited speed, gradient, train speed, remaining time, remaining distance, changing value of next limited speed and remaining distance of next limited speed these seven input variables respectively. The regression model of B-CART algorithm is bigger than the regression model of CART algorithm, the numbers of leaf node in these two models are 19 and 30 respectively. From the root node to each leaf node corresponds to a rule, that is to say, there are 19 rules in regression model of CART algorithm and 30 rules in regression model of B-CART algorithm.

The comparison of speed and controller's output under CART algorithm, B-CART algorithm and PID control are shown in the Figure 5 and Figure 6.

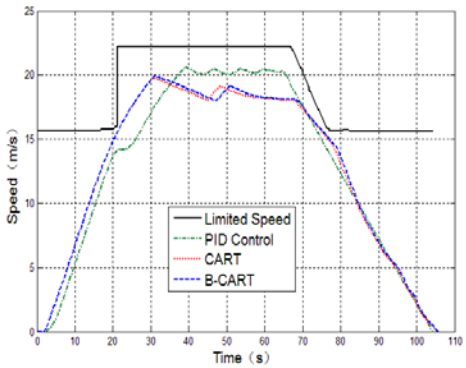


Figure 5. Comparison of Speed Curves

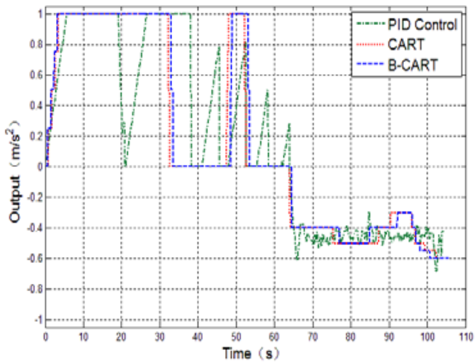


Figure 6. Comparison of Controller's Output

Figure 5 shows the operation of ITO model with CART or B-CART is smoother than PID control, and there is a long time of idle running in the middle stage, so the energy consumption will be less. In Figure 6, the switching times of controller's output in ITO model reduces obviously. The performance of ITO model with CART and B-CART are very similar with manual driving. We simulate 100 times using ITO model with CART or B-CART respectively, and calculate the average value. The function comparisons of PID control and ITO model with CART or B-CART are shown in Table 1.

Table 1. The Function Comparisons of PID Control, ITO model with CART or B-CART

| | PID Control | CART | B-CART |
|---|---|---|---|
| $e_t$(s) | 0.8 | **1.25** | 1.98 |
| $T_c$ | 14 | 6.34 | **5.56** |
| $I_r$ | 0.416 | 0.201 | **0.191** |
| $E$(J) | 236.7 | 206.48 | **201.52** |

Where $e_t$ is time error, $T_c$ is switching times of controller's output, $I_r$ is impingement rate, $E$ is energy.
Calculation of impingement rate:

$$I = \frac{1}{n}\sum_{i=1}^{n-1}\left|\frac{a_{i+1}-a_i}{\Delta t}\right| \tag{2}$$

The smaller the impingement rate, the better the riding comfort($C_r$), where $a_i$ is output of controller. Calculation of energy consumption:

$$E = \frac{\int Fv \cdot dt}{m} = \frac{\int m\left|a\right|v \cdot dt}{m} = \int \left|a\right| v \cdot dt \tag{3}$$

Where $a$ is the acceleration, $v$ is speed, this formula is a rough calculation, just for comparison. In Table 1, compared with PID control, the ITO model with CART or B-CART has some advantages: the $T_c$ and $E$ reduce obviously, and $C_r$ becomes better; although $e_t$ reduces slightly, it still meet the requirements.

In Figure 5 and Figure 6, the results of ITO model with CART and B-CART are very similar. From the four comparisons, they just have a little difference. The ITO model with B-CART performs better on $T_c$, $C_r$ and $E$. Above all, we use B-CART as the data mining algorithm of ITO model.

In order to verify the generality of ITO model, the other blocks of Yizhuang Line Beijing Subway are simulated too. Given that the space of paper is limited, we don't list the results here. The simulation results are similar with block 1, that is to say, the ITO model has achieved good performance in all blocks. Therefore, the generality of ITO model is very good.

## 6. Conclusion

In this paper, data in manual driving of excellent drivers are collected and filtered, and then the standard database is established. Through two data mining algorithms, two ITO models are dug out. We simulate all the blocks of Yizhuang Line Beijing Subway. From the results, ITO models have achieved good performance compared with PID control, especially in riding comfort and energy consumption. As for the two data mining algorithms, B-CART algorithm performs better, so we choose this algorithm as the data mining algorithm of ITO model.

There are some issues of this work need to be further researched, such as the robustness of ITO model, the adaptability of the ITO model for steep gradient and complex limited speed. Moreover, more advanced data mining algorithms are worth further studying for both simulation and real-world scenarios.

## Acknowledgements

This work is partially supported by the National High Technology Research and Development Program ("863" Program) of China under grant 2012AA112800, by New Scientific Star Program of Beijing under grant 2010B015, by the Fundamental Research Funds for the Central Universities under 2012JBM016, by the independent research project from the State Key Laboratory of Rail Traffic Control and Safety under grant RSC2011ZT001.

## References

[1] XX Chen, Y Zhang, H Huang. Train speed control algorithm based on PID controller and single-neuron PID controller. *Computer society*. 2010; 1: 107-110.
[2] CS Chang, SS Sim. *Optimising train movements through coast control using genetic algorithms*. Proc. IEEE Electric Power Applications. 1997; 144: 65–73.
[3] SH Han, YS Byen, JH Baek, TK An, SG Lee, HJ Park. *An optimal automatic train operation (ATO) control using genetic algorithms*. Proc. IEEE Region 10 Conf. TENCON. 1999; 1: 360–362.
[4] DW Chen, T Tang, CH Gao, RQ Mu. Research on the error estimation models and online learning algorithms for train station parking in urban rail transit. *Chin. Railway Sci.,* 2010; 31(6): 122-127.
[5] DW Chen, CH Gao. Soft computing methods applied to train station parking in urban rail transit. *Applied Soft Computing.* 2012; 12(2): 759-767.
[6] F Eibe, W Ian. Data Mining: Practical Machine Learning Tools and Techniques. *Morgan Kaufmann.* 2005.
[7] L Breiman, JH Friedman, RA Olshen, CJ Stone. Classification and Regression Trees. Belmont, CA: Wadsworth.1984.
[8] TG Dietterich. Machine Learning Research: Four Current Directions. AI Magazine. 1997; 18(4): 97-136.
[9] L Breiman. Bagging predictors. *Machine Learning.* 1996; 24(2): 123-140.
[10] RE Schapire, Y Freund, P Bartlet. Boosting the margin: a new explanation for the effectiveness of voting methods. *The Annals of Statistics*. 1998; 26(5): 1651-1686.
[11] L Breiman. Random forests. *Machine Learning.* 2001; 45(1): 5-32.