

# Train System Identification Based on the Analysis of Real-Time Running Data

Jiayi Yuan<sup>1+</sup>, Xiangxian Chen<sup>1</sup>, Zhujun Ling<sup>2</sup>, Wenbin Tang<sup>2</sup> and Guodong Teng<sup>2</sup>

<sup>1</sup> Department of Instrumentation Science and Engineering, Zhejiang University, Hangzhou, China

<sup>2</sup> Research and Development Center, Zhejiang United Science and Technology Co., Ltd, Hangzhou, China

**Abstract.** The train operation system is a time-varying nonlinear process. Most of the traditional train system identifications cannot describe the uncertain changes in actual operation. In this paper, we propose a time-varying nonlinear train model by analyzing the train operation process. AM-VFF-RLS method is applied to identify the parameters online. The simulation results verify that the model and identification method used in this paper can identify the characteristics of train operation system online and accurately.

**Keywords:** ATO; system identification; auxiliary model; recursive least square method; forgetting factor.

## 1. Introduction

As one of the most efficient passenger transport means, urban rail transit is now being widely used in different cities. The most important purpose of ATO (Automatic Train Operation) system is to adjust traction/braking force in different operation environment. Therefore, following the instructions, it can operate safely, reliably and efficiently [1].

The train operation process is a complicated process and it is a time-varying and nonlinear process. In traditional ATO research, most of the research methods are based on empirical modeling or offline identification. The results obtained are often difficult to adapt to the changing environment. Reference [2] presents an optimal ATO control using genetic algorithms, but there is no online optimization in the process of running. Reference [3] proposed a novel ATO algorithm based on iterative learning control theory. Since it requires a sufficient number of iterations, this method cannot meet the demand of real time. From the perspective of the actual engineering application, it is necessary to research an online train system identification method based on real-time running data.

In order to identify the train system online, the basic time-varying nonlinear model of train operation process is established according to Newton's laws of mechanics and the parameters are identified by AM-VFF-RLS (Auxiliary Model based Variable Forgetting Factor Recursive Least Squares) method.

## 2. Automatic Train Operation System

The most important purpose of ATO system is to operate the train in place of an experienced driver. The ATO system should automatically adjust the running speed of the train and park the train accurately and stably. The completion of the task of ATO relies on the coordination of the subsystems of ATC (Automatic Train Control) System. The ATP (Automatic Train Protection) System transfers travel permit, line parameters, speed limit information, train speed, train location and other information to ATO system. According to the real-time information received, ATO adjusts the running speed to adapt to the changing operation conditions. Fig. 1 shows the simplified process of adjusting speed by ATO.

---

<sup>+</sup> Corresponding author. Tel.: + 86 15957187314.  
E-mail address: sherryueda@126.com.

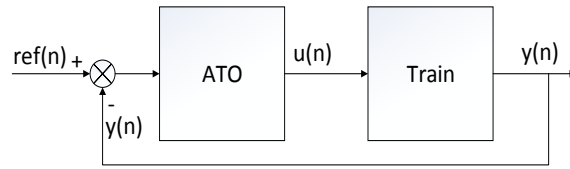


Fig. 1: Train model dialog.

According to the real-time information including reference speed (denoted as  $ref(n)$ ) and measured speed (denoted as  $y(n)$ ) received from ATP, the concrete calculation is carried out by ATO and the specified instructions (denoted as  $u(n)$ ) are sent to the train. So it can realize the tracking of the reference speed. The instruction of ATO in this paper is the traction/braking force and it is the input. And the output is the train running speed.

### 3. Mathematical Model

According to Fig. 1, we can establish the time-varying and nonlinear model of train as shown in Fig. 2.

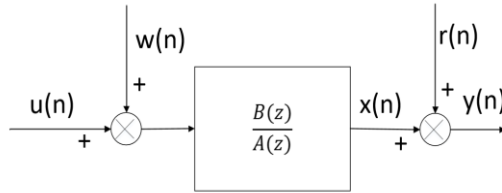


Fig. 2: Basic mathematical model of train.

ATO sends the traction/braking force  $u(n)$  to the train. The traction/braking force and the external resistance (denoted as  $w(n)$ ) are combined to form a resultant force on the train. According to Newton's law of mechanics, we can get the actual speed of train speed (denoted as  $x(n)$ ). The measured speed  $y(n)$  containing noise (denoted as  $r(n)$ ) is obtained by ATO.

The train movement is a complex and multi-degree of freedom process. In order to fully describe the dynamic characteristics of a train, the order of the system needs to be above 50 [4]. The slope of the train track is generally small, and the radius of the curve is far greater than the length of the train. In order to simplify and achieve the goal of on-line identification, the train can be regarded as a single particle in this paper. According to Newton's law of mechanics, the time-varying nonlinear model can be established as (1):

$$\begin{cases} dv/dt = cF \\ F = u - w \\ w = \alpha_0 + \alpha_1 v + \alpha_2 v^2 \end{cases} \quad (1)$$

where,  $t$  is time,  $v$  is the train speed,  $c$  is the acceleration coefficient,  $F$  is the resultant force of train,  $u$  is the traction/braking force,  $w$  is the resistance and  $w = \alpha_0 + \alpha_1 v + \alpha_2 v^2$  is an empirical formula.  $c$ ,  $\alpha_0$ ,  $\alpha_1$  and  $\alpha_2$  are time-varying coefficients. The change range of  $\alpha_0$ ,  $\alpha_1$  and  $\alpha_2$  may be quite wide.

Equation (1) can be written in a different form as:

$$x(n) = a_1(n)x(n-1)^2 + a_2(n)x(n-1) + b_1(n)u(n-1) + d_0(n) \quad (2)$$

where,  $x(n)$  is the train speed as well as output,  $u(n)$  is the traction/braking force as well as input,  $a_1(n)$ ,  $a_2(n)$ ,  $b_1(n)$  and  $d_0(n)$  are the time-varying parameters to be identified. And  $n$  is the sequence number.

### 4. Model Identification Method

Professor F. Ding proposed AM-RLS (Auxiliary Model based Recursive Least Squares) method [5]. AM-RLS method can be used to study systems with non-measurable internal variables and reduce the impact of noise. VFF-RLS (Variable Forgetting Factor Recursive Least Squares) method used in [6] has the

advantages of fast convergence and strong tracking ability. This paper combines the advantages of the two methods for parameter identification. According to the real-time running data, AM-VFF-RLS method is used to identify the time-varying parameters of train model.

#### 4.1. Recursive Least Square Method

RLS method is one of the basic methods of online identification. According to Fig. 2 and (2), the train model can be expressed as:

$$y(n) = \boldsymbol{\varphi}^T(n) \boldsymbol{\theta}(n) + r(n) \quad (3)$$

where,  $y(n)$  is the speed measurement,  $r(n)$  is the noise, and the parameter vector  $\boldsymbol{\theta}(n)$  and information vector  $\boldsymbol{\varphi}(n)$  are defined as follows:

$$\begin{cases} \boldsymbol{\theta}(n) = [a_1(n), a_2(n), b_1(n), d_0(n)]^T \\ \boldsymbol{\varphi}(n) = [y(n-1)^2, y(n-1), u(n-1), 1]^T \end{cases} \quad (4)$$

According to the minimization criterion function, we can get RLS method as follows:

$$\begin{cases} \varepsilon(n) = y(n) - \boldsymbol{\varphi}^T(n) \hat{\boldsymbol{\theta}}(n) \\ \hat{\boldsymbol{\theta}}(n) = \hat{\boldsymbol{\theta}}(n-1) + \mathbf{K}(n) \varepsilon(n) \\ \mathbf{K}(n) = \frac{\mathbf{P}(n-1) \boldsymbol{\varphi}(n)}{\boldsymbol{\varphi}^T(n) \mathbf{P}(n-1) \boldsymbol{\varphi}(n) + 1} \\ \mathbf{P}(n) = [\mathbf{I} - \mathbf{K}(n) \boldsymbol{\varphi}^T(n)] \mathbf{P}(n-1) \end{cases} \quad (5)$$

where,  $\varepsilon(n)$  is the prediction error,  $\mathbf{K}(n)$  is the Kalman gain vector,  $\mathbf{P}(n)$  is the error covariance matrix and  $\hat{\boldsymbol{\theta}}(n)$  is the estimated parameter vector.

#### 4.2. Variable Forgetting Factor

In order to prevent the phenomenon of data saturation and enhance the tracking ability, we can add a variable forgetting factor (denoted as  $\lambda(n)$ ) in the RLS method. If the value of  $\lambda(n)$  is smaller, then the tracking ability is stronger, but at the same time, the impact of noise is also greater. If the value of  $\lambda(n)$  is larger, then the tracking ability is weaker, but the impact of noise is also smaller and the convergence of the estimation error is smaller. In this paper, the updating equation of the forgetting factor is as (6):

$$\lambda(n) = 1 - \frac{\varepsilon(n)^2}{\sigma(\boldsymbol{\varphi}^T(n) \mathbf{P}(n-1) \boldsymbol{\varphi}(n) + 1)} \quad (6)$$

where,  $\sigma$  is a design parameter which affects the value of  $\lambda(n)$ ,  $\lambda(n)$  is the variable forgetting factor and the range of it is  $\lambda(n) \in (0, 1]$ .

#### 4.3. Auxiliary Model Identification

Since the collected train data contain colored noise, a direct usage of the RLS method cannot obtain the unbiased and consistent estimates of the train model. By applying auxiliary model identification idea, RLS method can be improved. An auxiliary model whose structure is consistent with the original model is established. The actual output of the system is replaced by the output of the auxiliary model (denoted as  $x_a(n)$ ) to identify the parameters. The system block diagram with auxiliary model is shown in Fig. 3.

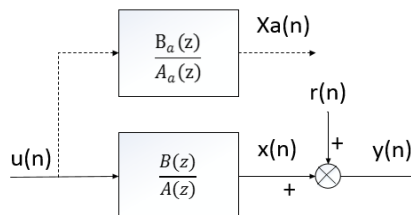


Fig. 3: System with auxiliary model.

According to [5], if we establish the auxiliary model by the estimated value of  $\theta(n)$ ,  $x_a(n)$  can be guaranteed to converge to  $x(n)$ . And considering the variable forgetting factor, we can get AM-VFF-RLS method and its calculation formula is as follows:

$$\begin{cases} \varepsilon(n) = y(n) - \hat{\phi}^T(n) \hat{\theta}(n) \\ \hat{\theta}(n) = \hat{\theta}(n-1) + \mathbf{K}(n) \varepsilon(n) \\ \lambda(n) = 1 - \frac{\varepsilon(n)^2}{\sigma(\hat{\phi}^T(n) \mathbf{P}(n-1) \hat{\phi}(n) + 1)} \\ \mathbf{K}(n) = \frac{\mathbf{P}(n-1) \hat{\phi}(n)}{\hat{\phi}^T(n) \mathbf{P}(n-1) \hat{\phi}(n) + \lambda(n)} \\ \mathbf{P}(n) = \frac{[\mathbf{I} - \mathbf{K}(n) \hat{\phi}^T(n)] \mathbf{P}(n-1)}{\lambda(n)} \\ \hat{\phi}(n) = [x_a(n-1)^2, x_a(n-1), u(n-1), 1]^T \\ x_a(n) = \hat{\phi}^T(n) \hat{\theta}(n) \end{cases} \quad (7)$$

## 5. Simulation Study

According to the proposed train model and the method of parameter identification, the real-time running data collected from Hangzhou Line 4 are used as an example for the simulation research.

AM-VFF-RLS is applied to identify the parameters of train model. The initial value of  $\lambda(n)$ ,  $\theta(n)$ ,  $\mathbf{P}(n)$  and  $\sigma$  in the simulation is set and the parameter identification results are shown in Fig. 6.

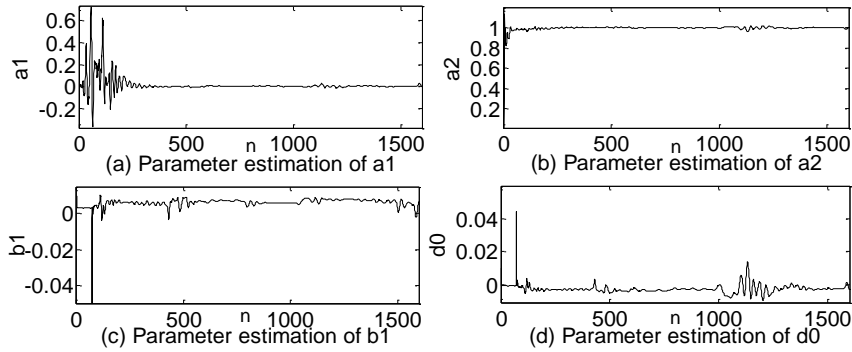


Fig. 4: Parameter identification results of AM-VFF-RLS.

We can see from the results that the parameters converge at about 200 sampling points. Each parameter has a different degree of change along with the change of environment.  $a_1(n)$  changes between -0.012 and 0.008 mostly;  $a_2(n)$  changes around 1.0 and the range is from 0.96 to 1.02;  $b_1(n)$  changes between -0.003 and 0.010 mostly;  $d_0(n)$  changes between -0.009 and 0.010 mostly and it changes fast around 1100 sampling points.

VFF-RLS and AM-RLS are also used to identify the same set of data. The estimated errors using AM-VFF-RLS, VFF-RLS and AM-RLS are compared. The compared results are shown in Fig. 7.

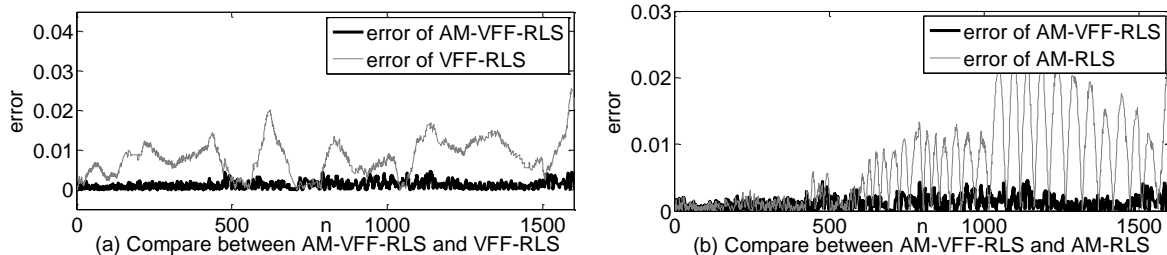


Fig. 5: Compared results.

As shown in Fig. 7, AM-VFF-RLS has a better performance than VFF-RLS and AM-RLS under the same conditions. The accuracy of AM-VFF-RLS is better than that of VFF-RLS and the error of AM-VFF-RLS does not change much with time. The error of AM-VFF-RLS is under 0.005 at most of the time. The error of AM-RLS is quite close to that of AM-VFF-RLS at beginning. But as time goes by, the error of AM-RLS will increase significantly.

According to the identification results of AM-VFF-RLS, the validity of the proposed model and the identification method in this paper is verified. This method has a better tracking ability and the estimated results are more close to the real value. When the environment changes, the method can still keep a high accuracy. The model and identification method used in this paper can fully recognize the time-varying nonlinear characteristics of train operation process.

## 6. Conclusion

The train operation process is a time-varying nonlinear process. In order to identify the train system on-line and accurately, we combine the RLS method and AM identification idea to identify train model on-line. In this paper, the basic time-varying nonlinear model of train is obtained by analyzing the force conditions of the train running process. We identify the time-varying parameters using AM-VFF-RLS method. We use the variable forgetting factor to enhance the tracking ability of the model. Auxiliary model is used to estimate the actual speed and weaken the impact of noise so that the model can be more accurate. The simulation results prove that the train model and identification method in this paper can identify the train system online and accurately. Besides, the results show that the identification method in this paper has a stronger tracking ability than AM-RLS method and is more accuracy than VFF-RLS method. The validity of the train system identification based on the analysis of real-time running data in this paper is verified.

## 7. Acknowledgements

The authors wish to acknowledge Jing Zheng and Yanni Zhang of Zhejiang University, who provided plenty of suggestions for this paper. This work is supported by the funding from the National Science and Technology Infrastructure Program of China under Grant 2015BAG19B03.

## 8. References

- [1] G. Young, M. Liu and L. Yu, "Nonlinear Predictive Control of Operation Process of High-speed Train", *Journal of The China Railway Society*, 2013, **35**(8):16-21.
- [2] S.H. Han, Y.S. Byen, J.H. Baek, T.K. An, S.G. Lee and H.J. Park, "An Optimal Automatic Train Operation (ATO) Control Using Genetic Algorithms", *Tencon 99 IEEE Region 10 Conference*, 1999, vol.1, pp.360-362,.
- [3] Y. Wang, Z. Hou and X. Li, "A Novel Automatic Train Operation Algorithm Based on Iterative Learning Control Theory", *Proc IEEE Conf. Service Operations, Logistics and Informatics*, 2008, vol. 2, pp.1766-1770.
- [4] C. Wang, "Parameter Identification and Its Application in Automatic Train Control", Ph.D. Thesis, Beijing Jiaotong University, 2014.
- [5] F. Ding, "System identification .Part D:Auxiliary model identification idea and methods", *Journal of Nanjing University of Information Sciences & Technology: Natural Science Edition*, 2011, vol. 4, pp.289-318.
- [6] Z. Li, Z. Young, H. Young and J. Liu, "Generalized Predictive Control with Dual Adaptation Method of High Speed Train", *China Railway Science*, 2015, vol. 36, pp.120-127.