ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-WEPGW064

MACHINE LEARNING APPLICATION IN BUNCH LONGITUDINAL PHASE MEASUREMENT *

X.Y.Xu †, Y.M.Zhou, Shanghai Institute of Applied Physics, CAS, Shanghai, China
also at University of Chinese Academy of Sciences, Beijing, China
Y.B.Leng#, Shanghai Advanced Research Institute, CAS, Shanghai, China

Abstract

High resolution bunch-by-bunch longitudinal phase measurement has been realized at Shanghai Synchrotron Radiation Facility (SSRF). In order to fully exploit the potency of the bunch phase monitor, the transient state during injection is being further studied. A longitudinal phase fitting method was used to study the synchrotron damping oscillation in injection events, where we can get the energy offsets between the injector and the storage ring, refilled bunch arrived time and the synchrotron damping time. However, manual multi-parameter fitting of nonlinear functions is awfully complex and slow. Machine learning algorithms, such as gradient descent and artificial neural network (ANN) is more suitable to do this fitting. Through these methods, we can quickly obtain more accurate fitting parameters and further realize online measurement of the refilled charge arrived time, energy offsets between the injector and storage ring, and the synchrotron damping time.

INTRODUCTION

The technology of bunch-by-bunch beam diagnosis plays an important role in understanding the motion state of the beam and the operation of the machine. Since 2012, studies on the bunch-by-bunch diagnostic technology have been carried out at the Beam Instrumentation (BI) group of SSRF (Shanghai Synchrotron Radiation Facility). by using a channel digital signal collecting BPM signal analyser and Zero - crossing detection methods of treatment, we have been able to realize the precise measurement of the beam longitudinal phase.

The Shanghai Synchrotron Radiation Facility (SSRF), which consists of a 150 MeV linear accelerator, a 180 m,3.5 GeV booster, and a 432 m, 3.5 GeV storage ring, is a high-energy third-generation light source [1]. To improve the efficiency and quality of the light, a top-up filling pattern was adopted. The instability of the injection beam can be obtained by measuring the longitudinal phase of the filling beam, which is advantageous for optimizing the injection scheme [2] .In addition, by fitting the longitudinal phase information of the refilled bunch at the time of injection, we can obtain the relevant parameters of the synchrotron damping oscillation caused by the mismatch between the injector and the storage ring, which is an important parameter for us to evaluate the operation state of the accelerator machine.

*Work supported by National Natural Science Foundation of China (No.11575282) and Ten Thousand Talent Program and Chinese Academy of Sciences Key Technology Talent Program
†xuxingyi@sinap.ac.cn
#lengyongbin@sinap.ac.cn

Theoretically speaking, the longitudinal phase of refilled bunch satisfies the following equation under the condition that the amplitude of oscillation is relatively small：

$$z_d = z_m \sin(\sqrt{\Omega^2 - \alpha_s^2}\, t + \varphi_0) \cdot e^{-\alpha_s t}. \quad (1)$$

In this equation, αs is the reciprocal of synchrotron damping time (τ), Ω is the frequency of synchrotron, z_m is the oscillation amplitude and φ_0 is the refilled charge arrived time. It can be found that this is a nonlinear equation with many parameters. How to quickly and accurately fit the measured longitudinal phase information of the refilled bunch according to the equation to obtain these key parameters is a crucial task.

The traditional manual adjustment parameter process is time consuming and laborious and the accuracy is difficult to quantify. On the other hand, the parameter fitting function library in some formed commercial software is prone to local convergence when it is used poorly, and it is not suitable for processing high-time repetition rate data in large quantities while ensuring accuracy. What's more, in order to achieve future online real-time processing, we hope to integrate a series of algorithms into the FPGA, so the non-open source of commercial software function library will limit our use.

As we know, machine learning is a powerful tool for data processing and analysis, which has gradually begun to get more and more in-depth applications in the field of accelerators. In this paper, we try to use the gradient descent algorithm and the artificial neural network (ANN) method to achieve a fast and accurate fitting of the longitudinal phase parameters of the refilled bunch.

FITTING BY GRADIENT DESCENT ALGORITHM

Gradient Descent is one of the most commonly used methods when solving model parameters of machine learning algorithms (unconstrained optimization problems). As the name implies, the calculation process of the gradient descent method is to solve the minimum value of the cost function along the direction of the gradient descent. The iteration formula is:

$$x_{k+1} = x_k - \alpha_k \nabla^k. \quad (2)$$

ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-WEPGW064

Gradient Descent is one of the most commonly used methods when solving model parameters of machine learning algorithms (unconstrained optimization problems).

As the name implies, the calculation process of the gradient descent method is to solve the minimum value of the cost function along the direction of the gradient descent.

Where $\bigtriangledown$ represents the gradient direction. α_k arepresents the search step size in the gradient direction. Gradient direction can be obtained by deriving the function. The determination of the step size is more complicated. If it is too large, it may diverge. If it is too small, the convergence speed is too slow. Therefore, we need to determine the size of the step by the specific conditions of the comprehensive fitting process.

In this problem, we select the sum of the squared differences of each data point (a group of about 2000) and the corresponding theoretical value obtained by the equation containing the fitting parameters as a cost function. Through the running of the multiple gradient descent algorithm, it is found that the stability of the machine itself during the injection process is degraded, which results in poor longitudinal phase data of the injected beam. In addition, because the capture of the injection process is prone to human error, a portion of the data in the database is not the longitudinal phase of the desired injection beam. Therefore, we choose to use a smaller learning step and add an algorithm interrupt decision that plays the role of data filtering in the gradient descent process, and only process the data as judged as good ones.

The results are shown in Fig. 1. We can see that the parameters (parameter 1 for z_m, parameter 2 for $\sqrt{\Omega^2 - \alpha_s^2}$, parameter 3 for φ_0, and parameter 4 for α_s, reference formula 1) are stable after many gradient iterations. The convergence of the cost function for evaluating the fitting accuracy in the iterative process is shown in Fig. 2. Using the gradient descent algorithm, the cost function can converge to a range of less than 0.00001 in a few seconds, and the parameter accuracy can be far within the BPM signal resolution.

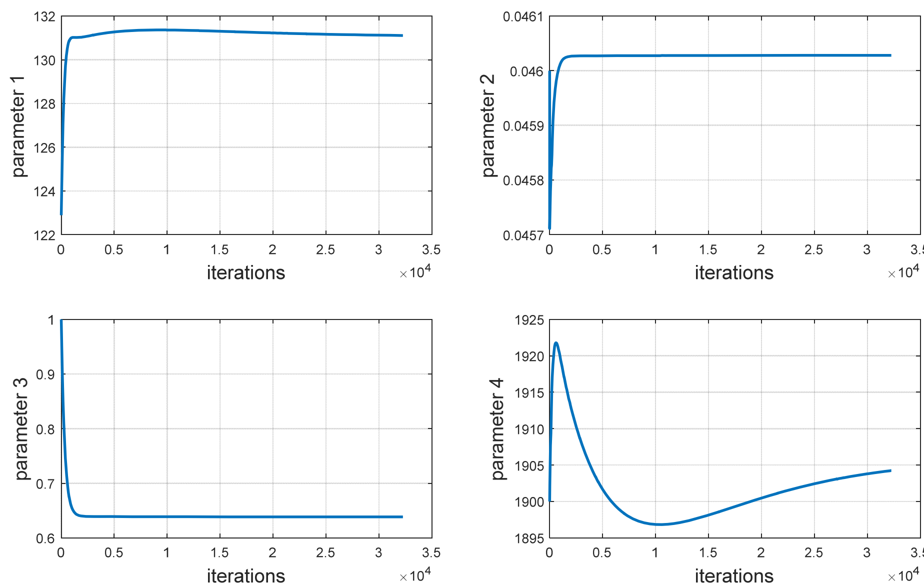

Figure 1: Convergence curve of parameters with the number of iterations.

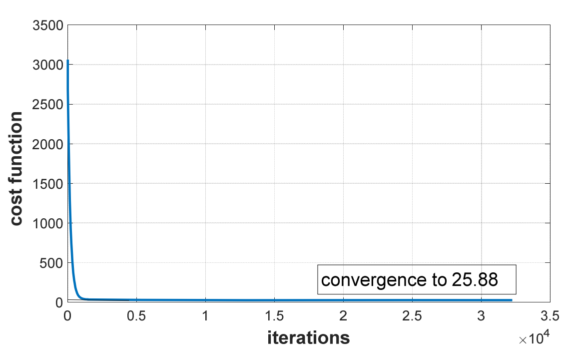

Figure 2: Cost function with the number of iterations.

FITTING BY ARTIFICIAL NEURAL NETWORK

Artificial Neural Network (ANN) is a research hotspot in the field of artificial intelligence since the 1980s. Artificial neural network is a nonlinear, adaptive information processing system composed of a large number of processing unit interconnections. It is based on the results of modern neuroscience research, trying to process information by simulating the way the brain neural network processes and memorizes information. It abstracts the human brain neuron network from the perspective of information processing, establishes a simple model, and forms different networks according to different connection methods. The artificial neural network has a strong ability to fit. Although the previous model training is more

ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-WEPGW064

complicated, after the training is completed, the model itself is very simple, and we can use the model for fast calculation [3]. In this paper, we hope to use artificial neural networks to achieve super fast and large-scale longitudinal phase parameter fitting.

We select the fitting parameters obtained by the gradient descent algorithm as the parameter true values, and use these fitting parameters together with the longitudinal phase obtained by the BPM measurement system as the training data to train the artificial neural network model (shown in Fig. 3.). After completing the model, we import the longitudinal phase data into the model, and the model can give the predicted fitting parameters.

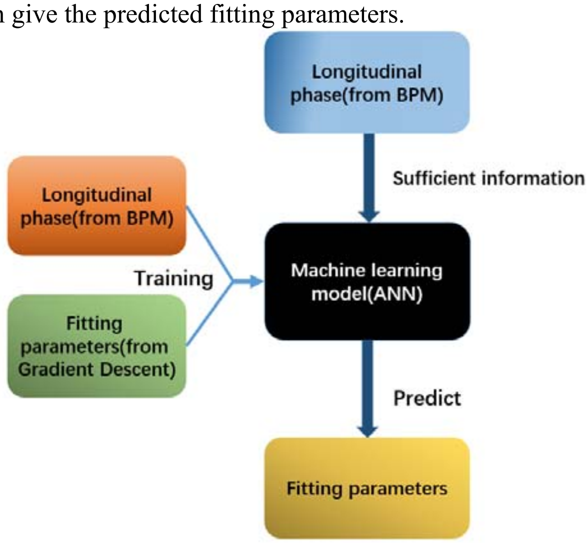

Figure 3: Machine learning schematic to predict fitting parameters.

This is a three-layer artificial neural network model (shown in Fig. 4.). It contains an input layer (1762 nodes), a hidden layer (200 nodes), and an output layer (2 nodes). The model uses a forward propagation approach in its use. Input layer input longitudinal beam phase data (1762 points) of each group, after the processing of the hidden layer, the two nodes of the output layer respectively output the oscillation amplitude and the synchrotron damping time, which are the two most concerned fitting parameters (because the frequency of synchrotron and the refilled charge arrived time are of little physical value and are not included in the prediction objectives of this model).

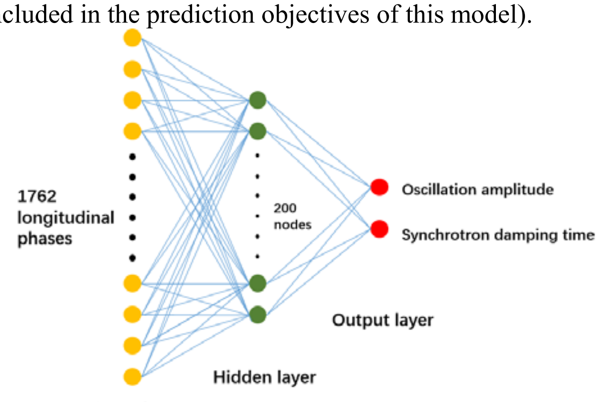

Figure 4: Schematic of the Neural network.

We divide the data into training set and test set according to 7:3. After training the model with training set data, the test set data is used to evaluate the accuracy of the model prediction. The true fitting parameter data and model prediction parameter data for the 252 sets of test sets are shown in Fig. 5. It can be seen that the model prediction parameters are highly consistent with the real fitting parameter data. After calculation, the prediction error of the model for the parameters is about 0.5%, which meets our design requirements.

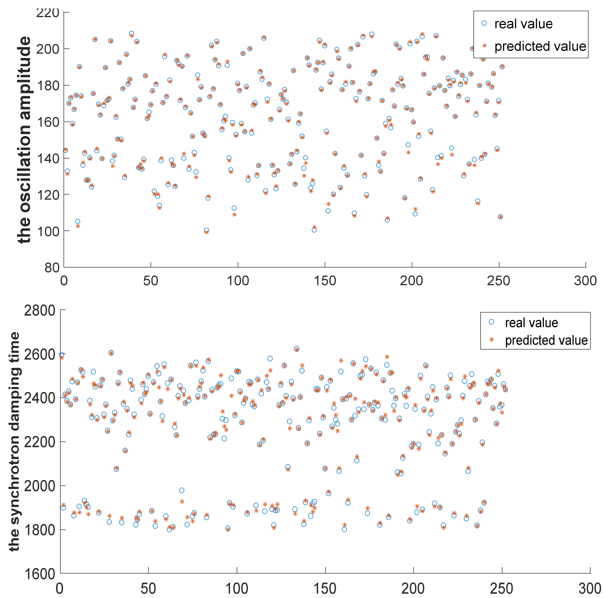

Figure 5: Results of parameters with the NN learning model.

CONCLUSION

In this paper, in order to achieve high-precision and high-speed measurement of the parameters related to the damping of the synchrotron, we explore the application of machine learning algorithms (gradient descent, artificial neural network) in multi-parameter fitting of nonlinear equations. To this end, we have established two models of gradient descent and artificial neural network. The gradient descent model can be easily improved in terms of fitting accuracy. Our model achieves a cost function that converges to 0.00001 in a few seconds, which far exceeds the resolution limit of the measurement system itself. In addition, we use the fitting parameters obtained by the gradient descent model as the true value to train the neural network model, and construct a three-layer neural network consisting of one input layer, one hidden layer and one output layer. After testing, the model can predict the accuracy of the parameters up to 0.5%.

In terms of application prospects, the gradient descent algorithm is suitable for high-precision measurement. In the case of willing time and computational cost, we can make the fitting parameters closer to the real value through more iterations. However, in many cases, we do not need to know every parameter in the equation. The neural network algorithm can perform high-speed prediction for the

ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-WEPGW064

required parameters. The neural network we build can complete the fitting process of thousands of data in one second. And by optimizing the training process, it achieves an accuracy of 0.5%. This meets our application needs. In addition, the neural network model is very simple after completing the offline training in the early stage. The prediction process occupies very little computing resources and can be integrated into the FPGA. It has a good application prospect in the accelerator engineering.

REFERENCES

[1] L. G. Liu *et al.*, "Commissioning of the SSRF Storage Ring", in *Proc. 11th European Particle Accelerator Conf. (EPAC'08)*, Genoa, Italy, Jun. 2008, paper WEPC042, pp. 2079-2081.

[2] Zhou, Y.M. *et al.*, "Bunch-by-bunch longitudinal phase monitor at SSRF.Nuclear Science and Techniques, 2018. **29**: 113

[3] B. Gao, J. Chen, Y. B. Leng, and Y. M. Zhou, "Machine Learning Applied to Predict Transverse Oscillation at SSRF", in *Proc. 7th International Beam Instrumentation Conference (IBIC'18)*, Shanghai, China, Sep. 2018, pp. 512-515. doi:10.18429/JACoW-IBIC2018-WEPC15