Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/19/20231006

Intelligent optical transceiver technology based on federated learning traffic prediction

Yirui Zhang

School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, 100080, China

zhangyirui@bupt.edu.cn

Abstract. With the development of optical network, modern optical network needs better performance. Because the traditional optical transceiver technology has a delay according to the flow switching transmission configuration, the delay optical network service still adopts the original configuration transmission, so a certain degree of frequency spectrum resources waste and high blocking rate will be caused. The above situation can be improved if the transmission configuration can be deployed in advance based on the predicted traffic. Federated learning is a scheme of distributed training model, which can train the traffic prediction model in distributed way under the premise of ensuring the privacy of client data, which is very suitable for the traffic prediction of optical network terminals. This paper proposes an intelligent optical transceiver technology based on federal learning traffic prediction, applies the federal learning on the traffic prediction of optical communication network terminal, distributed training traffic prediction model, and deploy the optical transceiver early transmission configuration such as modulation format and baud rate parameters, thus to weaken the delay of optical transceiver technology, reduce the network blocking rate and improve the transmission performance of optical network.

Keywords: federated learning, traffic prediction, optical transceiver technology.

1. Introduction

With the development of optical network, business types and business traffic have been greatly increased, modern optical network needs better performance, and the optimization of more intelligent optical transceiver technology can provide possibilities for more flexible optical network. Because the traditional optical transceiver technology has a delay according to the flow switching transmission configuration, the delay optical network service still adopts the original configuration transmission, so a certain degree of frequency spectrum resources waste and high blocking rate will be caused. The above situation can be improved if the transmission configuration can be deployed in advance based on the predicted traffic. Federated learning is a scheme of distributed training model, which can train the traffic prediction model in distributed way under the premise of ensuring the privacy of client data, which is very suitable for the traffic prediction of optical network terminals.

Some scholars have made prospective research on the above aspects. McMahan B presented the concept of federated learning, distributing training data across mobile devices and forming a shared model by aggregating local computing updates (aggregating locally-computed updates [1]. Shangjing Lin proposed the wireless communication traffic prediction model based on federated learning, the traffic prediction model of each single grid base station, the central cloud server uses the federated average algorithm to integrate the parameters of the grid traffic model with similar traffic distribution [2]. K.Yonenaga introduced The traditional bandwidth variable optical transceiver technology which can dynamically optimize the transmission link by changing the baud rate and modulation format[3]. Li Shuai Introduced traffic prediction into optical transceiver scheme, the simulation results show that this scheme can significantly reduce the network blocking rate and improve the performance of optical network [4].

This paper innovatively proposes the intelligent optical transceiver technology based on federal learning traffic prediction, applies the federal learning on the traffic prediction of optical communication network terminal, distributed training traffic prediction model, and deploy the optical transceiver early transmission configuration such as modulation format and baud rate parameters, thus to weaken the delay of optical transceiver technology, reduce the network blocking rate and improve the transmission performance of optical network.

2. Theoretical basis

2.1. Traditional optical transceiver technology

Under the condition of considering the transmission limit of the physical layer (including transmission distance and ONSR, the bandwidth variable optical transceiver opportunity establishes a static query table under the SDON framework, that is, the transmission capacity of different sizes correspond to the corresponding modulation format and wave rate parameters [3] [5]. The modulation formats mainly include: BPSK, QPSK, 16 QAM, etc., such as 7 G Baud, 8 G Baud, etc. When the optical network traffic and SNR requirements change, the optical transceiver opportunity will be configured according to the configuration action in the query table to meet the different transmission requirements of the service. However, due to the certain configuration switching delay of the traditional optical transceiver, the original configuration transmission is still used in the switching process, which cannot be quickly switched to the corresponding configuration, resulting in the decline of the transmission performance of the optical network. If it fails to quickly switch to the corresponding configuration when a high transmission bandwidth is required, the network blocking rate will increase, or if the low transmission bandwidth cannot be quickly switched to the corresponding configuration, which will lead to a waste of spectrum resources.

2.2. Federal learning

In federated learning, we aim to train models across multiple child nodes that can communicate only with a common central node without the need to exchange their data samples [6]. The federated learning process is divided in four steps: ① The central node sends a global model to each child node. ② Each child node is locally updated with the method of stochastic gradient descent (SGD). ③ Send the parameters of the local model of each child node to the central node. ④ The central node aver the model parameters according to the weights to generate a global model for the next round of training.

The process of federal learning is shown in figure 1.

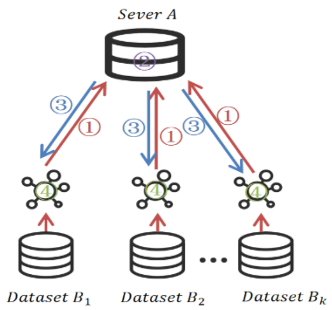

Figure 1. The process of federal learning.

DOI: 10.54254/2755-2721/19/20231006

In the above four-step cycle, the accuracy of the model will gradually improve with the number of training rounds.

2.3. Flow prediction method

In view of the traffic prediction problems, a large number of related studies have emerged at home and abroad. Traditional algorithms generally use statistical probability model or time series prediction model for flow prediction, such as AR, MA, ARMA, and ARIMA [7-8]. As the deep learning algorithm makes breakthroughs in various fields, the long CNN that can capture spatial correlation and the long and short memory LSTM networks that can capture temporal correlation are gradually applied to the field of communication traffic prediction. Because optical communication network terminals pay more attention to the relationship between traffic size and time, LSTM network is selected for prediction.

In the past, many researchers use RNN model to predict temporal flow data [9], but the RNN model cannot solve the problem of long-term dependence of temporal flow, it is difficult to learn the past long time flow, and the LSTM model as a special RNN model, on the basis of RNN model added a state, retained the to the current time t all historical information to realize the information within the model, which can solve the gradient disappear and explosion in long time series training, can have better prediction effect in long time series [10].

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c_t}, s_t = o_t \odot tanh(c_t), \tilde{c_t} = tanh(W_c x_t + U_c s_{t-1} + b_c) \tag{1}$$

$\odot$ is the product of matrix elements; c_{t-1} is the memory unit of the previous moment; $\tilde{c_t}$ is a candidate state obtained after mapping by nonlinear functions; and $tanh$ is a hyperbolic tangent function.

The LSTM model controls the transfer of information in the network by introducing Gating Mechanism, and the three gating units in the model are the input gate i_t, forgetting gate f_t, and output gateo_t. The values of the three gates in the LSTM model are between 0 and 1, representing passing through information in a certain proportion. The forgetting gate f_t controls the candidate state of the current moment $\tilde{c}_t$ retains partial information, the input gate i_t controls the retention of the candidate state at the current moment c_{t-1}; the output gate o_t selectively outputs the information to the hidden state h_t by controlling the current state c_t.

The calculation methods of the three doors are respectively:

$$i_t = σ(W_i x_t + U_i s_{t-1} + b_i), f_t = σ(W_f x_t + U_i s_{t-1} + b_f), o_t = σ(W_o x_t + U_{oi} s_{t-1} + b_o) \tag{2}$$

The structure of LSTM model is shown in figure 2.

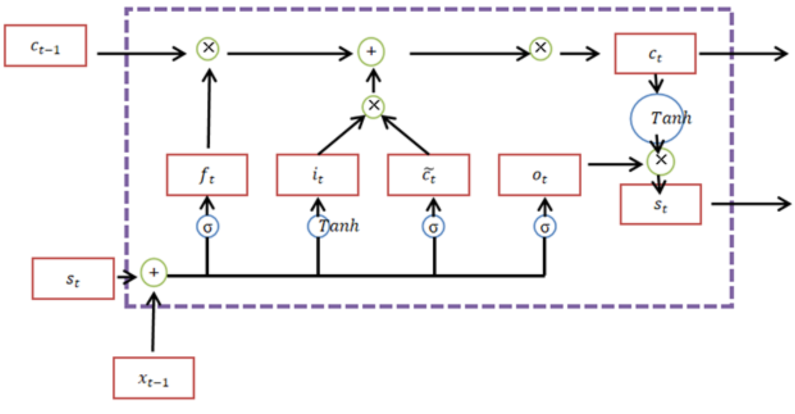

Figure 2. LSTM model.

3. Model of global traffic prediction

The global traffic prediction model is trained by applying the federated learning on the optical network terminal traffic prediction. The basic structure of the model adopts a model composed of one layer of LSTM and one fully connected layer, with the parameters set to input _ size = 2, hidden_size=4, output_size=1, and number _ layer = 1. First, each edge computing server collects traffic data and trains the local traffic prediction model. The training method is to find the model parameters that minimize the loss function. The optimized loss function in this paper is the mean square error of the predicted traffic and the true traffic;

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(observed_t - predicted_t)^2 \quad (3)$$

Second, the edge ccomputing server will be local traffic prediction model parameters uploaded to the central cloud server, by the central cloud server model parameters fusion, and the fusion of model parameters to the edge computing server, fusion parameters is the FedAvg, namely the local data proportion to weight the parameters of each node form the global model parameters;

$$w_{t+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} w_{t+1}^{k} \quad (4)$$

Finally, each edge computing server continues to conduct local model training on the basis of the converged model parameters until the end of the training to form a general traffic prediction model. The existing traffic prediction schemes generally adopt the centralized training framework, which requires to gather a large amount of traffic data on a central server for model training, with high communication cost and high model complexity. The model proposed in this paper is a distributed training framework. Each edge computing server only needs to collect traffic data at the regional level for model training. Therefore, the communication cost is small and the model complexity is low.

The intelligent optical transceiver technology architecture based on federal learning traffic prediction is shown in the following figure below. In the SDON framework, the trained traffic prediction module of the controller will monitor the optical transceiver traffic and predict the future traffic in real time, the control module of optical transceiver in the controller will query the optical transceiver configuration switch table and determine if the traffic exceeds a certain threshold, the optical transceiver, the modulation format, baud rate and other parameters to the appropriate position [11].

The structure of the intelligent optical transceiver technology is shown in figure 3.

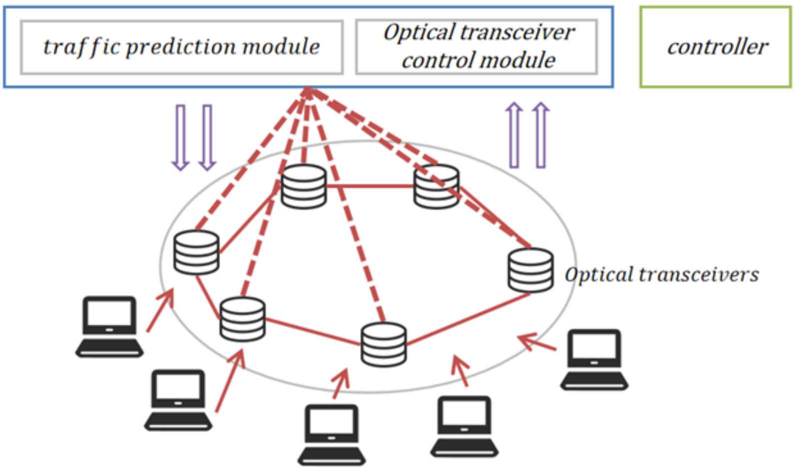

Figure 3. Intelligent optical transceiver technology.

4. Experimental results

We selected traffic data collected every five minutes for ISP communication equipment with centers in 11 European cities, from 07 to 11:17, and experimental data from two of these periods were selected as

DOI: 10.54254/2755-2721/19/20231006

local data for two sub nodes. The traffic prediction model is trained through federated learning, and the model is verified on the local data by summarizing the central node parameters in each round. The following figure shows the true and predicted values of the local traffic data in the ten rounds, respectively. The results of training in Node1 and Node2 is shown in figure 4,5,6,7.

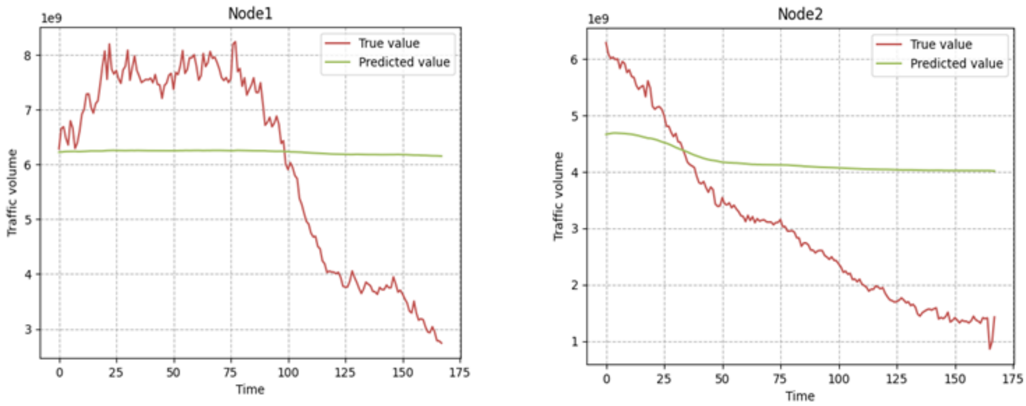

Figure 4. Results of the first round of training: (a) Node1 (b) Node2.

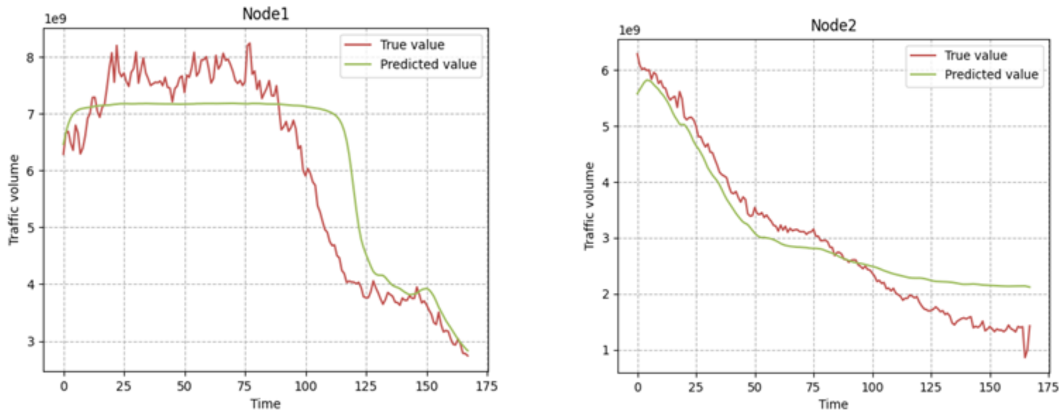

Figure 5. Results of the third round of training: (a) Node1 (b) Node2.

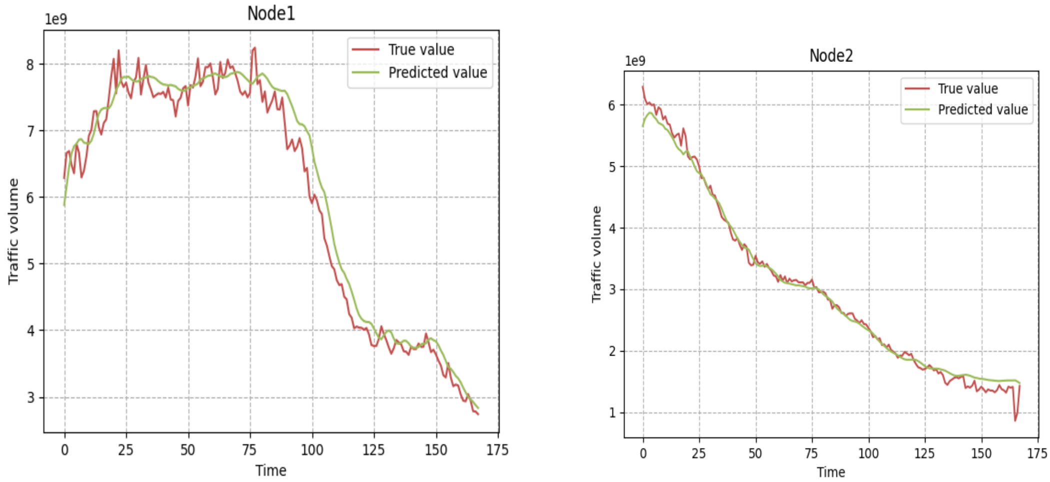

Figure 6. Results of the fifth round of training: (a) Node1 (b) Node2.

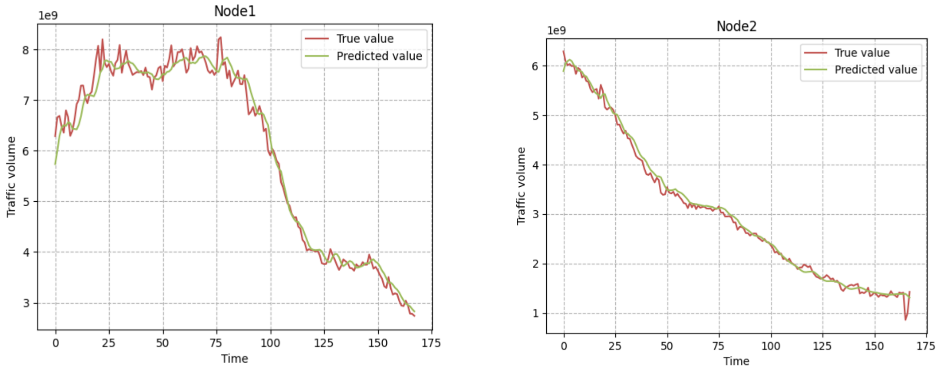

Figure 7. Results of the ninth round of training: (a) Node1 (b) Node2.

With the increase of the number of training rounds, the gap between the true value of traffic and the predicted value of traffic gradually decreases, indicating that the flow prediction model is more accurate. From this experiment, we can prove that federated learning can indeed distributed train the traffic prediction model on the premise of ensuring the privacy and difference of client data.

5. Conclusion

This paper innovatively proposes the intelligent optical transceiver technology. Through the transmission configuration deployment according to the predicted traffic, it can avoid the waste of spectrum resources and high blocking rate caused by the deployment delay of the traditional optical transceiver technology, to improve the transmission performance of the optical network. In this paper, we introduce federated learning into network traffic prediction, refine the federated learning scheme for optical terminal network traffic prediction, and prove its feasibility with experiments. This paper combines the above two points, proposes the intelligent optical transceiver technology based on federated learning traffic prediction, and optimizes the current traditional optical transceiver technology.

There are currently many improved versions of federated learning, and there has been quite a lot of research trying to improve the underlying FedAvg model. These studies can be divided into two major categories, improvements to local training and improvements to aggregation, such as personalized federated learning, federated learning based on contrast learning MOON algorithm, and so on. In the future, we can try which federated learning efficiency and effect are better in the application scenarios of intelligent optical transceiver technology. In the vision of 6G communication, network slicing technology will be widely used. If the network can be classified according to the characteristics of traffic, and then set different transmission configurations for different categories, the optical transceiver technology will be further optimized.

References

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-Efficient Learning of Deep Networks from Decentralized Data,” PMLR, vol. 54, pp. 1273–1282, Apr. 2017.

[2] S. Lin et al., “Wireless communication traffic prediction based on federated learning,” Computer application, vol. 0, Sep. 2022.

[3] K. Yonenaga et al., “Bit-rate-flexible all-optical OFDM transceiver using variable multi-carrier source and DQPSK/DPSK mixed multiplexing,” IEEE Xplore, Mar. 2009.

[4] S. Li, “Research on Optical Network Traffic Analysis and Intelligent Optical Transceiver Technology Based on Machine Learning - CNKI,”kns.cnki.net.

[5] Q. Yang, Y. Liu, T. Chen, and Y. Tong, “Federated Machine Learning,” ACM Transactions on Intelligent Systems and Technology, vol. 10, no. 2, pp. 1–19, Feb. 2019.

[6] J. K. Fischer et al., “Bandwidth-Variable Transceivers based on Four-Dimensional Modulation Formats,” Journal of Lightwave Technology, vol. 32, no. 16, pp. 2886–2895, Aug. 2014.

[7] R. H. Shumway and D. S. Stoffer, Time Series Analysis and Its Applications. Cham: Springer International Publishing, 2017.
[8] I. Rojas et al., “Soft-computing techniques and ARMA model for time series prediction,” Neurocomputing, vol. 71, no. 4–6, pp. 519–537, Jan. 2008.
[9] J. T. Connor, R. D. Martin, and L. E. Atlas, “Recurrent neural networks and robust time series prediction,” IEEE Transactions on Neural Networks, vol. 5, no. 2, pp. 240–254, Mar. 1994.
[10] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
[11] F. Tang, Z. Md. Fadlullah, B. Mao, and N. Kato, “An Intelligent Traffic Load Prediction-Based Adaptive Channel Assignment Algorithm in SDN-IoT: A Deep Learning Approach,” IEEE Internet of Things Journal, vol. 5, no. 6, pp. 5141–5154, Dec. 2018.