Routing Algorithm for Elastic Optical Network Based on Machine Learning assisted Traffic Prediction

Yutong Wang¹, Bokai Cheng¹, Cheng Hang¹, Jiaoliang Yu¹, Shixin Liu¹, Yufei Hu¹ and Jianhua Shen¹*

¹School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu, 210003, China
*Corresponding author’s e-mail: shenjh@njupt.edu.cn

Abstract. Elastic Optical Network (EON) is the promising technology for future network which may suffer congestion or fragment problems when huge amount of burst traffic arrives. This paper proposes an improved routing strategy for EON based on Artificial Intelligent (AI) assisted traffic prediction. Theoretical analysis and simulation results show that compared with traditional algorithms, the improved algorithm can effectively reduce the blocking rate of the network especially under heavy network load.

1. Introduction

With the rapid development of new applications such as cloud computing, high definition video, virtual reality, etc. both the service requirement and Internet throughput have grown explosively which causes a great challenge for optical bearer networks. However, the traditional Wavelength Division Multiplexing (WDM) network uses wavelength as the minimum resource allocation unit, which is difficult to adapt to the needs of users whose transmission rates vary greatly. Elastic Optical Network (EON) introduces Optical Orthogonal Frequency-division Multiplexing (O-OFDM) technology which uses fine-grained or ultra-large-grained Frequency Slot (FS) resources to adapt to dynamic customer requirements. It has been widely accepted as the promising bearer technology for future [1].

Although EON has an intrinsic ability to accommodate different service requirements flexibility, there may occur spectrum congestion problem due to the spectrum continuity and consistency constraints especially when huge massive burst traffic arrives frequently. How to effectively allocate and optimize available spectrum resources in EON has become the focus of attention for both academia and industry. Obviously, if the arrival time and bandwidth requirements of the arriving service can be predicted, and the network spectrum resources can be configured accordingly, the degree of congestion in EON will be reduced and resource utilization efficiency will be improved, respectively. Although great efforts have been carried out on network traffic prediction, most traditional predicting technologies may suffer unsatisfactory performance while many non-linear characteristics of network traffic such as randomness, time variability, and chaos are ignored. Therefore, it is difficult to precisely predict the network traffic for a long period of time, resulting in limited application scope [2]. In recent years, Artificial Intelligence based methods (Deep Learning or Reinforcement Learning) have been applied into network traffic prediction scenarios. Compared to traditional methods, AI assisted traffic prediction will introduce existing traffic records and their information for training neural networks to achieve the precise predicting future traffic [3].
In order to solve the Routing and Spectrum Allocation (RSA) problem in EON, many researchers have proposed related algorithms or solutions for resource assignment optimization. The main work of [4] focused on routing, which selects the optimal path based on the spectrum resource state and path length, and they applied the commonly used spectrum allocation strategy in the spectrum allocation stage. Jia et al proposed a dynamic RSA algorithm in [5], which improves the performance of routing algorithms by using back propagation neural network (BPNN) to predict the future information of each service with time-awareness from similar service sources. Ying Wang et al used Ant Colony Optimization Algorithm to solve the RSA problem in [6], and verified through simulation that the proposed algorithm has better performance. Generally, most of the proposed RSA algorithm based on machine learning does not consider the problem of heavy network load, i.e. if large amount of service arrives, the network performance may suffer.

This paper will use machine learning based on Recurrent Neural Network (RNN) to predict the future congestion of the network, and apply the prediction results to modify RSA strategy. Therefore, an improved routing distribution strategy based on network traffic prediction is proposed which uses RNN model to predict the congestion degree of link and adjust link select strategy based on the prediction results. In the case of high network load, the algorithm can effectively reduce the network blocking rate.

2. Principle of Improved Algorithm

2.1. General Symbol Definition

In order not to lose generality, this paper uses \( G=(V,E) \) to represent the physical network topology of EONs, where \( V=\{v_i|i=1,2,...,N\} \) represents the node set, and the total number of nodes in the network is \( N \). The link set is denoted as \( E=\{l_{mn}|m,n \in V\} \), and the fiber link between nodes \( v_m \) and \( v_n \) is represented as \( l_{mn} \). The total number of Frequency Slots (FS) on each fiber link is denoted as \( |X| \).

When time \( t=T \), the total number of services being carried by each link is \( |Y| \). Their set is expressed as \( R^w=\{R^w_1,R^w_2,...,R^w_{|Y|}\} \), the total number of incoming services is denoted as \( |Z| \), and their set \( R^r=\{R^r_1,R^r_2,...,R^r_{|Z|}\} \), the number of FS occupied or required by service \( R \) is \( S^r_R \), respectively.

2.2. Degree of Congestion on path and link

Both the path and link congestion degree is proposed to measure the level of path or link congestion between two node pairs in EON per unit time. Here, the link congestion is represented by \( \lambda \), and its expression on the link \( l_{v_av_b} \) can be denoted as follows:

\[
\lambda_{v_av_b} = \frac{\sum_j^{|Y|} s^r_{j} + \sum_j^{|Z|} s^r_{j'}}{|X|} \tag{1}
\]

The expression of the congestion degree of the path between nodes \( v_a \) and \( v_b \) can be denoted as follows:

\[
\mu_{v_av_b} = \left\{ \begin{array}{ll}
\max(\lambda_{v_av_a},...\lambda_{v_av_b}) & \text{When there is only one link between } v_a \text{ and } v_b (v_i \text{ and } v_j \text{ are nodes between } v_a \text{ and } v_b); \\
\min(\max(\lambda_{v_av_a},...\lambda_{v_av_b}),...\max(\lambda_{v_av_a},...\lambda_{v_av_b})) & \text{otherwise } (v_i,v_j \text{ and } v_k \text{ are nodes between } v_a \text{ and } v_b). 
\end{array} \right. \tag{2}
\]

The difference between \( \lambda \) and \( \mu \) can explained as follows: \( \lambda \) is the congestion degree of a single link, and \( \mu \) is equal to the maximum congestion degree of each link between two node pairs. When there is more than one path between node pairs, the congestion degree of the path will take the smallest congestion degree of the path. The purpose of this definition is that when \( \mu \leq 1 \), the path between two node pairs does not have congestion caused by the traffic volume exceeding the path carrying capacity. If congestion occurs at this time, it can be solved by spectrum defragmentation algorithm. When \( \mu > 1 \), it means that it is impossible to solve the blocking problem in the path by spectrum defragmentation alone. At this case, other path selection strategies should be used to avoid a large amount of service blocking.
Considering the example shown in Figure 1, there are two available paths between node $v_1$ and node $v_2$, one is the shortest path $l_{v_1v_2}$, the other is a secondary short path, which consists of links $l_{v_1y_1}, l_{y_1y_2}$ and $l_{y_2v_2}$. When $t=T_0$, the frequency slot occupation of the four links is shown in Figure 1 (b). It can be seen that the total FS required for coming services that reach the links $l_{v_1y_1}, l_{y_1y_2}$, $l_{v_2y_2}$ and $l_{y_2v_2}$ at this time are 16, 9, 11, and 4, respectively. Therefore, the congestion degree of the link $l_{v_1v_2}$ can be calculated according to formula (1) as follows:

$$\lambda_{v_1v_2} = \frac{\sum_{i=1}^{16} R^{f}_{i} + \sum_{j=1}^{17} R^{w}_{j}}{|X|} = \frac{(1+2+4+1+2+4)+16}{24} = 1.25$$

Similarly, $\lambda_{v_1y_1} = 0.96$, $\lambda_{y_1y_2} = 1.04$, and $\lambda_{y_2v_2} = 0.75$, respectively. Since there are two paths between nodes $v_1$ and $v_2$, according to formula (2), we can get

$$\mu_{v_1v_2} = \min \left( \lambda_{v_1v_2}, \max (\lambda_{v_1y_1}, \lambda_{y_1y_2}, \lambda_{y_2v_2}) \right) = 1.04$$

Because $\mu_{v_1v_2} > 1$, it indicates that the path between $v_1$ and $v_2$ is in a congested state.

### 2.3. Structure of model and model training

The congestion degree $\lambda$ defined in the previous section will be predicted by deep learning models. Furthermore, the prediction RNN model used in this paper consists of an encoder, a decoder, and a Dense Layer. The encoder and decoder are composed of RNNs, which act as both encoding pairs and decoding pairs. The encoder maps a variable-length source sequence to a fixed-length vector, and the decoder converts the vector back into a variable-length target sequence. As for Dense Layer, it is equivalent to a fully connected layer. The activation function of the input layer uses tanh, and the hidden layer uses Relu [7]. We set the number of residual channels to 32, which is the “number of channels” or “filters” of the remaining output. Then, the number of skip channels is also set to 32. During training, set the batch size to 128, the learning rate to 0.01, the number of training steps to 2000, and the optimizer to use Adam.

### 2.4. Traffic Prediction-Dijkstra (TPD) algorithm

As mentioned above, most traditional algorithms only approximate to local optimum at any specific time and cannot forecast possible changes in traffic load or avoid potential congestion risks in future. The improved algorithm proposed in this paper can adjust the routing selection method according to the prediction result of the congestion degree. By applying RNN, the congestion of the whole network can be reduced or even avoided.

For any path between source node and destination node (here we considered the shortest path and the second shortest path), when the average value of $\mu$ is greater than 1 within a typical period, e.g. 1 minute, it is defined as severe congestion. In our network model for predicting, that the start time point of this period is denoted as $T_0$. At this time, the routing table is being searched in advance to find a sub-short path whose $\mu_{v_1v_2} < 1$ during this time, denoted as $l$. When a new service arrives, the
shortest path and the second shortest path will be found, then it should be determined whether the path blocking degree $\mu$ is greater than 1 and whether $t=T_0$. If the above conditions are met, then it is determined whether there is a path and this path will be added to the candidate path when it exists; if it is not met, spectrum allocation is performed directly.

![Algorithm 1 Traffic Prediction-Dijkstra](image)

3. Simulation and Discussion

3.1. Dataset designing
A dataset for training and testing is established in this paper, which contains information about the network status including: number of links, volume of the incoming service requests, the number of frequency slots occupied by current service and the calculated link congestion degree $\lambda$, etc. These data are records obtained by simulating the business through NSFNET (14 nodes, 21 links) within 72 hours. Each link has 320 FSs with protected bandwidth. In the simulation, the transmission rate of network services is evenly distributed between 10 and 400 Gb / s. The above dataset is divided into a training set (0 to 60 hours) and a test set (61 to 72 hours).

3.2. Simulation results
As shown in Figure 3, the blue line is the record of network congestion degree within 3 hours in the test set, and the red line is the prediction result during this period. From the perspective of the trend, the two curves fit quite well. From the details, when the congestion degree exceeds the threshold, the model used in this paper can accurately predict the trend.

![Figure 3. Training results](image)  ![Figure 4. Network blocking rate performance](image)

Figure 4 shows the simulation results of TPD algorithm in NSFNET. It can be seen that with the increase of the network load, the blocking rate of both the improved TPD + FF and the typical D + FF algorithm increased accordingly. When the network load is light, the impact of the two algorithms on the network blocking rate is not much different. But as the network load increases, this difference
becomes more obvious. When the network load is 200 Erl, the difference of network blocking rate between two algorithms is only about 0.1%; while when the network load is 500 Erl, the difference is 1.8%; and when the network load is 600 Erl, the difference is 2.5%. In terms of numbers, compared with the D + FF algorithm, the improved TPD+ FF algorithm can improve the network blocking rate, especially when the network load is large.

4. Conclusion
This paper proposes an improved routing strategy for EON based on AI assisted traffic prediction. First, the network congestion degree is predicted by RNN, and then more low-congestion-degree links are allocated in advance to the links which congestion degree is higher than the threshold according to the prediction results. Simulation results show that the proposed algorithm can effectively reduce the network blocking rate especially under heavy network load.

References
[1] Olszewski, I. (2017) Improved dynamic routing algorithms in elastic optical networks. Photon Network Communications, 34: 323-333.
[2] Szostak, D., Walkowiak, K. (2019) Machine learning methods for traffic prediction in dynamic optical networks with service chains. In: 21st International Conference on Transparent Optical Networks. Angers.
[3] Polson, N. G., Sokolov, V. O. (2017) Deep learning for short-term traffic flow prediction. Transportation Research Part C: Emerging Technologies, 79:1-17.
[4] Moura, P. M., Fonseca, N., Scaraficci, R. A. (2014) Fragmentation aware routing and spectrum assignment algorithm. In: IEEE International Conference on Communications. Paris. pp. 1137-1142.
[5] Jia, W., Xu, Z., Ding, Z., Wang, K. (2017) Routing and spectrum assignment algorithm with prediction for elastic optical networks under self-similar traffic. In: International Conference on Optical Communications and Networks. Hangzhou. pp. 23-25.
[6] Wang, Y., Zhang, J., Zhao, Y., Wang, J., Gu, W. (2013) ACO-based routing and spectrum allocation in flexible bandwidth networks. Photonic Network Communications, 25: 135-143.
[7] He, K., Zhang, X., Ren, S., Sun, J. (2016) Identity Mappings in Deep Residual Networks. In: European Conference on Computer Vision. Amsterdam. pp. 630-645.