IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 17, 2020, accepted February 26, 2020, date of publication March 2, 2020, date of current version March 13, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2977612

Link State Aware Dynamic Routing and Spectrum Allocation Strategy in Elastic Optical Networks

YANG ZHOU, QIANG SUN, AND SIYU LIN
School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China

Corresponding author: Siyu Lin (sylin@bjtu.edu.cn)

This work was supported in part by the National Nature Science Foundation of China under Grant U1534201, and in part by the Fundamental Research Funds for the Central Universities under Grant 2019JBM005.

ABSTRACT Compared with traditional wavelength division optical network, elastic optical network (EON) divides the network spectrum into smaller spectrum slots to improve the spectrum utilization, but the high-quality spectrum division also complicates the routing and spectrum allocation (RSA) problem. Various strategies are proposed for reducing the RSA complexity and improving system traffic bearing capacity. However, previous RSA strategies do not consider the changing physical layer impairments that will also impact signal quality and even lead to violation of quality of transmission (QoT), the data cannot be transmitted correctly if the link state is degraded. Therefore, cross-layer optimization is desired, which means that different layer information is taken into account in the RSA strategy. In this paper, we propose a new link state-aware (LSA) RSA strategy to guarantee the QoT requirements under different link states. At first, the link state is evaluated based on chromatic dispersion (CD) and optical signal-to-noise ratio (OSNR), and a LightGBM model is exploited for CD and OSNR estimation. In LSA-RSA strategy, the link state is considered as a metric for qualified routing paths finding, and the link capacity is calculated based on the link state and used in spectrum allocation. Simulation results show that the average CD and OSNR estimation errors of the LightGBM model are 0.28ps/nm and 0.68dB, respectively. Under different link states and traffic loads, the LSA-RSA strategy can reduce traffic failure probability by more than 20%, and traffic load can increase 40Erlang when the bandwidth blocking probability equals 10%.

INDEX TERMS Cross-layer optimization, elastic optical network, machine learning, optical performance monitoring, routing and spectrum allocation.

I. INTRODUCTION

The continuous growth of network traffic volume requires higher network capacity and more flexible network management. The optical networks which carry the most traffic data are supposed to provide various bandwidth resources to meet current traffic requirements. Suitable network management allocates available resources in an optimal method, reduces the wastage of the resources and increases actual network throughput [1]. Nevertheless, the traditional wavelength division multiplexing (WDM) networks use fixed-sized spectrum grid as minimum resource granularity (usually 50GHz or 100GHz), it is hard to adapt different granularity service requests, an entire wavelength is required in WDM networks even the traffic request has a low bit rate, so the spectrum resource utilization is limited. To overcome the problem mentioned above, elastic optical network (EON) which is based on orthogonal frequency-division multiplexing (OFDM) is proposed [2], [3]. EON divides spectrum resources into finer spectrum slots with the bandwidth of 12.5GHz or smaller, narrower spectrum slots could be allocated to lower bit rate traffic, thereby improving the spectrum resource utilization. In the meantime, to reduce the cost of the wavelength conversion devices, the data transmitted from the source node to the destination node in EON should use same spectrum slots, which is also referred to as the spectral continuity constraint. Correspondingly, routing and spectrum allocation (RSA) becomes the core problem in EON, which is responsible to establish the lightpath for traffic bearing, where the lightpath refers to a combination of the routing path and spectrum slots assigned to the traffic request.

The associate editor coordinating the review of this manuscript and approving it for publication was Longxiang Gao.

TABLE 1. The main variables and parameters.

Variables	Description	Variables	Description
x_i	ith training sample	$p_r(i)$	Probability that the traffic needs i slots
h_i	Amplitude histogram result converted from x_i	$T(v_s, v_d, bw)$	Traffic from node v_s to v_d and requires bw bandwidth
$\hat{y_i}$	Estimated value	$LS(LS_D, LS_O)$	Link state, LS_D/LS_O represents the CD/OSNR value
$\Delta\tau$	Maximum allowed dispersion value	α	Multiplication factor used in routing paths finding
E_b	Energy per bit	P_T	Routing paths for traffic T
N_o	Spectral noise density	SS_T	Available spectrum slots for traffic T
R_s	Symbol rate	R^2	Coefficient of determination
$G(V, E)$	Network topology, V and E represent node and link set	bf_T	Traffic blocking flag for traffic T
$M(e, j)$	Slot occupancy state flag of jth slot in link e	ls_T	Link state degradation flag for traffic T
$C_{sb}(k)$	Capacity of the spectrum block with k slots	$\lvert SS_{e,k}\rvert$	Slot number of the kth spectrum block in link e
$P(k, i)$	Probability that only i slots can be used in k slots block	sf_k	Spectrum block fragmentation flag for the kth block

At present, most existing works on RSA mainly focus on improving the spectrum utilization [4]– [6]. However, the absence of optical-electrical-optical in EON results in the accumulation of the optical impairments, that is to say, the optical signal is more susceptible to physical layer impairments, such as dispersion, amplifier noise, and crosstalk. Some traffic interruptions even start with link quality degradation caused by serious physical impairments [7]. Therefore, the lightpath establishment and the quality of transmission (QoT) depend not only on sufficient spectrum slots assigned by RSA strategy but also on the link or network quality [8]–[10]. At this point, to improve resource utilization and actual data throughput, it is desired to exploit dynamic cross-layer network optimization in EON, which considers information from different network layers. Related research about cross-layer optimization has been carried out in optical networks recently [11]–[13]. In EON, the cross-layer RSA strategy should be adjusted based on physical layer information to meet the QoT requirements, the links with poor quality are limited for less traffic bearing, thereby avoiding adverse effects on network reliability. Furthermore, load balancing in networks could be realized if spectrum occupancy information in the physical layer is considered. Therefore, for the implementation of cross-layer optimization, physical layer impairments and states should be monitored firstly, which is also known as optical performance monitoring (OPM) technique [14], [15]. Recently, artificial intelligence (AI) technique gets lots of attention in optical networks and provides a prospective method for OPM [16], [17], several parameters can be derived by analyzing physical signal features through AI-based models. Then if the link state is evaluated based on OPM results and treated as a reference in RSA, cross-layer optimization could be implemented. However, research work combining OPM and cross-layer optimization is still relatively rare, and further research is needed for better network management.

Based on the above mentioned considerations, we investigate the impacts of different link states on physical layer signals in this paper. Then the link state is represented by optical parameters, including chromatic dispersion (CD) and optical signal to noise ratio (OSNR) which are estimated via the LightGBM model. Based on the link state evaluation results, we proposed a link state-aware (LSA) RSA strategy to mitigate the adverse effects of link state degradations. In the routing phase, LSA routing algorithm is proposed to search qualified routing paths that satisfy the QoT requirements. Traffic carried by the link is restricted when the link state is degraded. The load balancing issue is also considered in the routing phase. In the spectrum allocation process, fragmentation reducing (FR) algorithm is proposed to allocate suitable spectrum resources for traffic. The effectiveness of the proposed LSA-RSA strategy is validated by simulation, and results show that the LSA-RSA strategy improves the network throughput with the QoT requirements, especially when the link state is degraded.

The rest of this paper is organized as follows: related works are presented in Section II. In Section III, we introduce the system model, which enables cross-layer network optimization. Our link state evaluation method is presented in Section IV, and the LSA-RSA strategy is presented in Section V. Simulation results are given in Section VI. In the end, we briefly summarize this paper in Section VII.

For the convenience of presentation, we summarize the main variables and parameters used in this paper in Table 1.

II. RELATED WORKS

Cross-layer optimization in EON has recently attracted lots of research to improve network performance, researchers consider information from different network layers when making network resource allocation. The joint optimization of delay-bandwidth and fragmentation is proposed for RSA when the physical layer impairments are considered [18], and the extension work considers the bit loading and guarantees the end-to-end BER for spectrum allocation [19]. However, the above works are carried out under static traffic scenario and exploit off-line RSA strategy. In real network operation, the requests will arrive and tear down in randomness. To solve the dynamic RSA problem, several spectrum allocation schemes are proposed, such as first-fit (FF) [20], subcarrier-slot partition with first-last-fit (FLF) [21], spectrum partition policy with FLF [22], traffic-based fragmentation-aware spectrum allocation (TFSA) [23], and access blocking probability (ABP) based algorithm [24]. The performance of FF is compared with random fit and exact fit [25], and last fit

scheme is also included as a benchmark [26]. Based on these strategies, dynamic and cross-layer optimization works in EON are studied. Beyranvand *et al.* analyze the impairments and nonlinear effects of fibers in the physical layer to estimate the OSNR and the transmission distance under different modulation formats [27]. To reduce the dispersion effect, a dispersion-adaptive spectrum allocation scheme is proposed to decrease the traffic blocking probability, the spectrum slots with lower dispersion are assigned to longer lightpath [28]. Dharmaweera *et al.* propose an impairment-driven guard-band allocation scheme to ensures the transmission quality and improves the network throughput [29]. Fontinele *et al.* consider QoT degradation caused by new establishing traffic, then select suitable routing path, modulation format and spectrum resources for request aiming to reduce traffic blocking probability [30].

However, these existing works only consider the theoretical impairments existing in the physical layer, without considering the more significant real-time optical impairments. These changing impairments may cause link state degradation and could be represented by optical parameters. Therefore, the parameter estimation is the basis to achieve real-time cross-layer optimization. Dedicated equipment could monitor several physical layer parameters, besides, the development of machine learning (ML) in recent years provides another promising and cost-effective way. Extensive works are carried out via various ML techniques for CD and OSNR estimation, since these two parameters are directly related to signal transmission quality. Thrane *et al.* exploit neural network (NN) to estimate OSNR values under different modulation formats [31]. For multiple parameters estimation, Huang *et al.* investigate an artificial neural network (ANN) model for OSNR and CD estimation [32]. Wang *et al.* [33] use constellation data and convolution neural network (CNN) to estimate OSNR, but the estimation accuracy varies under different OSNR values. However, accurate clock synchronization is required in previous works since the eye or constellation diagram is involved. To extract signal features more effectively, asynchronous amplitude histograms (AHs) is introduced [34]. Sun *et al.* select AH as input data, and combine particle swarm optimization and deep neural network (DNN) to implement OSNR estimation [35]. Besides, asynchronous sampling is used to preprocess the data, then the data is inputted to the convolutional neural network (CNN) for OSNR estimation [36], [37]. To improve the robustness of parameter estimation, Tanimura *et al.* propose an OSNR estimation model based on DNN, while assessing the current accuracy and providing the uncertainty information [38]. Besides these various neural network models, LightGBM [39] and XGBoost [40] models can also achieve the parameters estimation. These two models are both based on gradient boosting decision tree which combines many weak decision trees to form a complicated and accurate decision tree. As mentioned above, some works about parameters monitoring have been studied, but the work about link state evaluation based on monitoring results and regarding results as indicators in network management is relatively less and still needs further research, it is self-evident that the combination of cross-layer and dynamic RSA can provide better performance when the link state is degraded.

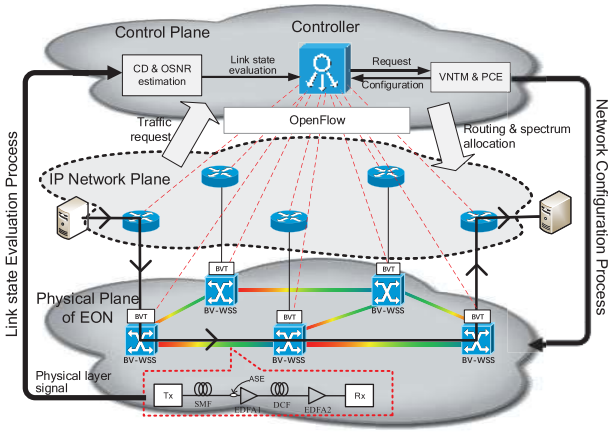


FIGURE 1. Link state-aware system model based on EON.

Based on existing research works, in this paper, we exploit the LightGBM model for both CD and OSNR estimation, then evaluate the link state for RSA and cross-layer optimization. The network throughput with different link state conditions is measured to validate the effectiveness of our proposed strategy.

III. SYSTEM MODEL

The LSA-RSA strategy enables the cross-layer optimization, routing path and spectrum resources are allocated according to the evaluated link state, and the system model is shown in Fig. 1. The model includes three planes: physical plane of EON, IP network plane and control plane. The physical plane is based on EON that provides spectrum resources and link state information for lightpath establishment and link state evaluation, respectively, it consists of two main components: bandwidth variable transponder (BVT) and bandwidth variable wavelength selective switch (BV-WSS). BVT assigns traffic to the suitable central frequency with enough spectrum resources, and BV-WSS performs cross-connect in EON, the incoming signal at particular central frequency is switched to specified destination fiber or node. The IP network plane is responsible for traffic access and collects the request information. Firstly, it receives traffic requests and aggregates the data from different clients. Then, the request information, including the source, destination node addresses, and the required bandwidth, is forwarded to the control plane. The IP network plane coordinates with the physical plane for data transmission, the traffic is accessed by the IP network plane and carried in the physical plane of EON. On the other hand, the IP network plane coordinates with control plane for network management, the system is controlled by the control plane through the particular protocol, such as OpenFlow. And the control plane can be seen as the brain of the entire system, which is responsible for controlling the network, evaluating

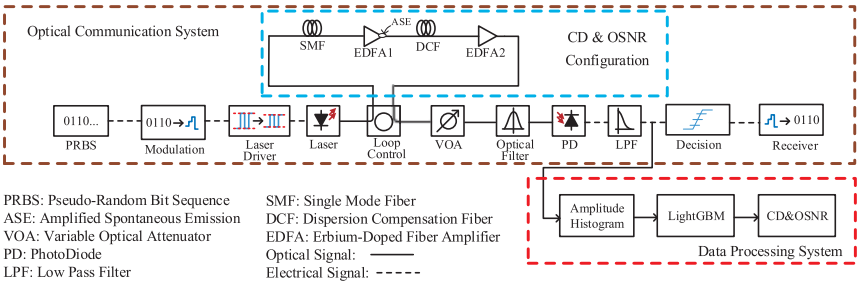


FIGURE 2. Schematic diagram of parameters estimation model.

the link state, routing traffic requests, and allocating spectrum resources. To make the resource allocation decision, the control plane works with the other two planes and performs two key processes: link state evaluation process and network configuration process. In the link state evaluation process, the physical signal is collected periodically and then used to estimate optical domain parameters, including CD and OSNR, then the optical link state is evaluated according to these two parameters. For the network configuration process, the controller module allocates routing path and spectrum resources based on link state, then the network configuration information including routing and spectrum allocation results are provided to the IP network plane and EON for traffic bearing. Two processes compose an evaluation-action cycle, and then the link state based resource allocation in EON could be triggered.

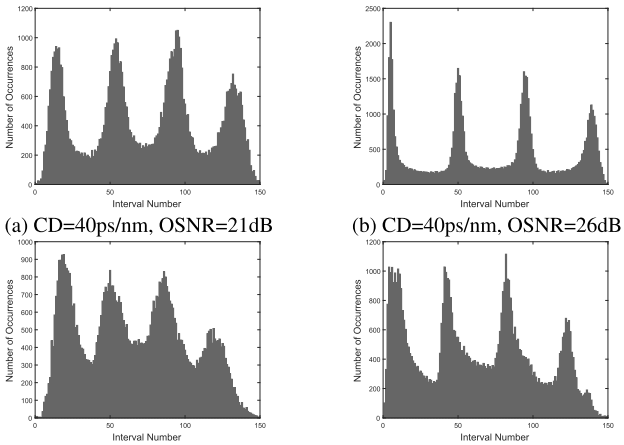
In the control plane, three modules work together to implement link state evaluation and network control processes. Signals in the physical plane are affected by physical layer impairments, and the impairments information is also reserved due to the usage of high-speed sampling rate devices in receiver. Therefore, it is possible to estimate related parameters from physical layer signal. The controller module evaluates the link state based on the CD and OSNR estimation results and performs LSA-RSA strategy. The CD & OSNR estimation module receives physical layer signals and implements parameters estimation. The virtual network topology manager and path computation element (VNTM & PCE) model is invoked to perform routing and resources allocation process. The main function of VNTM is to maintain network topology, VNTM can simplify the routing problem in optical networks since it ignores connection details in lower planes. Also, the spectrum occupancy status of EON is recorded in VNTM. PCE is used to find routing paths and assign spectrum resources for requests based on link state information. While the controller getting routing path and resources allocation scheme, the network configuration is triggered, then the corresponding lightpath is established in EON for traffic bearing. VNTM & PCE model avoids the excessive workload of nodes and centralized controller, which only need to focus on data transmission and making decisions, respectively. Therefore, the overall performance of the network can be improved.

The network states, including both the resource utilization state and link state, are changing over time and are essential for network management. Taking lower plane information into account in network management, performance decreasing introduced by link state degradation could be compensated. We evaluate the link state based on CD and OSNR, and then perform LSA-RSA strategy in the controller for better network performance. The controller connects the different planes and makes the global decisions, both link state and spectrum resource information are considered. Through the LSA-RSA strategy, almost all established lightpaths in EON could meet the quality requirements, the transmission quality is guaranteed and the network resource utilization is also improved.

IV. LINK STATE EVALUATION

In this section, we propose a link state evaluation method based on CD and OSNR estimation values. In optical communication systems, the accumulation of link impairments, including CD and amplified spontaneous emission (ASE) noise, degrade the signal transmission quality. CD leads to ISI and ASE reduces the signal OSNR, both of them lead the signal degradation and severe bit errors at the receiver. In our experiment, the link impairments are presented as CD and OSNR pairs. The OSNR can be used to represent the link noise level and is intuitive to indicate the link state, so we regard the OSNR as one link parameter. Under different CD and OSNR pair configurations, we first collect the received signal and then estimate CD and OSNR values. Two key processes are invoked in this section: CD and OSNR estimation process and link state evaluation process.

In CD and OSNR estimation process, we build a model consisting of an optical communication system [41] and a data processing system, as shown in Fig. 2. In the optical communication system, various link state conditions can be configured and the signal can be transmitted. Pseudo-random bit sequence (PRBS) is generated as an information source, then the PRBS is modulated to a specified format and converted to optical signal. During signal transmission, the dispersion compensation fiber (DCF) compensates the dispersion effect caused by single mode fiber (SMF), so the transmission distance can be extended. The erbium-doped fiber amplifier (EDFA) amplifies the attenuated optical signal, but ASE noise is also introduced into signal and reduces the OSNR. At the receiver, the optical signal is converted into electrical signal after photodiode (PD) and low



(a) CD=40ps/nm, OSNR=21dB (b) CD=40ps/nm, OSNR=26dB

(c) CD=120ps/nm, OSNR=21dB (d) CD=120ps/nm, OSNR=26dB

FIGURE 3. AHs under different CD and OSNR pairs, interval number of AH is equal to 150, and the CD value is residual dispersion that has been compensated by DCF.

pass filter (LPF). In addition to the signal decision, part of electrical signal is forwarded to the data processing system, whose main role is to estimate the CD and OSNR values. The key of estimation is to obtain the probability distribution characteristics of the signal, so the received continuous signal is sampled as discrete signal firstly, and the amplitude range of the signal is divided into several intervals. Then each sampled signal will fall into a specific amplitude interval. We count the number of each amplitude interval, and the probability distribution characteristics of signal amplitude can be obtained and shown as amplitude histogram (AH). In our experiment, we adjust the dispersion coefficient of SMF and the noise figure of EDFA to configure CD and received OSNR, respectively. Under different CD and OSNR pair configurations, AHs have significantly different characteristics, which are shown in Fig. 3. It can be seen that the AH of the better quality signal has sharper and more independent peaks than the degraded quality signal. Due to different distribution characteristics of AHs containing CD and OSNR information, it can be used to estimate the two parameters by the LightGBM model.

The LightGBM model in the data processing system is trained to learn the relationship between AHs and CD as well as that between AHs and OSNR. In our experiment, several sets of signal are collected under different CD and OSNR pair configurations to construct the training set. The ith training sample is expressed as $x_i = [h(i), \max(s(i))]$, $h(i)$ is the AH result, and $\max(s(i))$ is the maximum value of received signal $s(i)$. Then the training set can be expressed as $X = [x_1; x_2; \ldots; x_m]$. Correspondingly, the training set is labeled as $Y = [y_{CD_1}, y_{OSNR_1}; y_{CD_2}, y_{OSNR_2}; \ldots; y_{CD_m}, y_{OSNR_m}]$. To obtain a better LightGBM model and estimate CD and OSNR values more accurately, the training process aims to minimize the objective function, which is calculated as

$$\begin{aligned} Obj(\Theta) &= L(\Theta) + \Omega(\Theta) \\ &= \sum_{i=1}^{n} \frac{1}{2}(y_i - \hat{y}_i)^2 + \sum_{k=1}^{K}(\gamma T_k + \frac{1}{2}\sum_{j=1}^{T_k} w_j^2). \end{aligned} \quad (1)$$

where $L(\Theta)$ is loss function and can be represented as a mean square error, y_i and $\hat{y}_i$ are the true and estimation values of CD or OSNR, respectively. $\Omega(\Theta)$ represents the regularization term which is used to prevent overfitting. The regularization term includes constraint on the number of leaves T_k, and L2 norm of the leaf weight w_j^2. Compared with XGBoost model [40], LightGBM uses histogram optimization to convert the feature values into histogram data before training, thereby accelerating the split point finding. In our work, the kth feature of training samples is the kth column of training set $X[:, k]$, and it can be seen as a $m \times 1$ vector, then the vector is converted into a $n \times 1$vector through histogram technique, where n is the number of bins in histogram. The training set dimension changes from m rows to n rows, denoted by X_H. During the training process, the decision tree will split at the point which has the largest variance gain, and variance gain of kth feature at split point d is defined as

$$V_k(d) = \frac{1}{n}\left(\frac{\left(\sum_{\{x_i \in X_H : x_{i,k} \le d\}} g_i\right)^2}{n_l^k(d)} + \frac{\left(\sum_{\{x_i \in X_H : x_{i,k} > d\}} g_i\right)^2}{n_r^k(d)}\right). \quad (2)$$

where g_i is negative gradient of loss function and calculated as $-\frac{\partial L(\Theta)}{\partial \hat{y}_i} = (y_i - \hat{y}_i)$, it is also known as residual error. $x_{i,k}$ is the kth feature of the ith training sample in processed training set X_H. $n_l^k(d) = \sum I(x_i \in X_H : x_{i,k} < d)$ and $I()$ is the indicator function. In one iteration, feature values of all samples are traversed to split decision tree at d_k^*, where $d_k^* = argmax(V_k(d))$. In the later iteration, the next decision tree is generated to fit residual errors until the model accuracy meets requirements or the number of iterations reaches preset maximum number. In the end, the output of the LightGBM model is expressed as $\hat{y}_i = \sum_{l=1}^{L_T} f_l(x)$, where L_T is total number of iteration, i.e., the number of decision trees, and $f_l(x)$ is the output of the lth decision tree. Once we get a trained LightGBM model, CD and OSNR can be inferred from AH data by the model.

In the link state evaluation process, the state is evaluated according to CD and OSNR estimation results. Links with different link states should offer acceptable transmission quality for traffic bearing. Dispersion will cause ISI, to meet the QoT requirement, the maximum allowed dispersion $\Delta\tau$ should be less than one-fourth of the symbol pulse duration [8], which can be calculated as

$$\Delta\tau < \frac{T_S}{4}, \quad \Delta\tau = LD\Delta\lambda. \quad (3)$$

where T_S is the duration time of one symbol, L and D are the fiber length and dispersion coefficient, respectively. $\Delta\lambda$ is the spectral width of the laser. For example, the pulse duration of one symbol equals 50ps at 20GSymbol/s (40Gbit/s at 4-PAM modulation format), so the dispersion delay must be less than 12.5ps. In this case, if a laser with 0.1nm spectral width, CD of the link cannot exceed 125ps/nm for the QoT requirement. OSNR is the other important parameter that has a direct relationship with bit error rate (BER) [9]. To ensure

TABLE 2. Required parameters at different bit rates.

	20Gbps	40Gbps	100Gbps
CD requirement	<250ps/nm	<125ps/nm	<50ps/nm
OSNR requirement	>20dB	>23dB	>27dB

the data transmission quality, we should keep OSNR above the threshold, and OSNR can be derived from BER. Therefore, we set the BER threshold firstly and then calculate the corresponding OSNR threshold. For M-PAM modulation formats, the BER can be calculated as

$$BER = \frac{M-1}{M\log_2 M} erfc\left(\sqrt{\frac{3\log_2 M}{k\left(M^2-1\right)}\frac{E_b}{N_o}}\right). \quad (4)$$

$$\left.\frac{E_b}{N_o}\right|_{dB} = OSNR|_{dB} - 10\log_{10}\frac{R_s\log_2 M}{B_{ref}}. \quad (5)$$

where E_b and N_o represent energy per bit and spectral noise density, respectively. The variance k indicates the detection method used in receiver, $k=1$ means that coherent detection is used and $k=2$ means direct detection is used. $erfc()$ is the complementary error function. R_s corresponds to the symbol rate, and B_{ref} means the optical measurement bandwidth and is usually equal to 12.5GHz. In light of the above, different bit rates require different CD and OSNR pairs. We take 40Gbps bit rate with 4PAM modulation format and coherent detection as an example. When the OSNR equals 22dB, E_b/N_o equals 16.9dB, then BER can be derived as 2×10^{-10}. With 1dB OSNR redundancy and ensuring BER is below 10^{-9}, we calculate several CD and OSNR requirements based on [8], [9], and the results are listed in Table 2.

The CD may change due to the variation of fiber aging and other external environments. OSNR may vary due to changes in signal power and noise levels. We evaluate the link state based on CD and OSNR pairs. Once the CD or OSNR cannot meet the requirements, the data transmission should be stopped due to QoT violation. The maximum traffic bit rate in EON should also be restricted according to the link state, and the number of available contiguous spectrum slots, i.e., the maximum bandwidth is restricted accordingly. In optical networks, a lightpath may contain several links or spans, so the total CD should be the accumulated value of multiple links, and OSNR can be expressed as

$$OSNR|_{dB} = P_L - Loss - N_F - 10\log_{10}N_S + 58. \quad (6)$$

where P_L is launch power at the transmitter, $Loss$ and N_F are the loss and noise figure of each span, N_S is the number of spans. To represent the worst case, we use the largest $Loss$ and N_F among the lightpath to calculate OSNR.

V. LINK STATE-AWARE ROUTING AND SPECTRUM ALLOCATION

Based on the link state evaluation results, we propose LSA routing and FR spectrum allocation algorithms. For the dynamic network operation, the source node, destination node, and duration time of the request are unpredictable, so the routing path and corresponding spectrum should be allocated based on the current network state and the information of the coming traffic request. In the routing phase, the LSA algorithm guarantees that the traffic can be transmitted with required QoT. Load balancing is also taking into account in LSA. In the spectrum allocation phase, the spectrum block and link capacity are calculated based on link state, and then the proposed FR algorithm allocates spectrum resources to carry traffic with minimum link capacity loss. In this section, we present the link capacity calculation method, then the details of LSA and FR algorithms are designed. For convenience, we first introduce several notations and variables used to represent EON. $V=\{v_1,v_2,\ldots,v_n\}$ and $E=\{e_1,e_2,\ldots,e_m\}$ represent nodes and links in the network, respectively, then EON can be represented as $G(V,E)$. The number of spectrum slots in one link is represented as N, and the occupancy of slots in EON is represented as a matrix M, where $M(e,j)=1$ indicates that the jth slot in link e has been occupied, and 0 otherwise. $T(v_s,v_d,bw)$ means that a traffic request starts at node v_s and ends at v_d with requirement of bandwidth bw. P_T and SS_T represent the routing path and allocated spectrum for traffic T, respectively.

A. LINK STATE BASED CAPACITY CALCULATION

The changing link state and spectrum slot occupancy both lead to the link capacity changing, so the capacity calculation should be as simple as possible to achieve rapid capacity update. In this part, we utilize theory of the compositions of number to calculate spectrum block capacity $C_{sb}(k)$, where k means the number of adjacent and unoccupied spectrum slots in one spectrum block. Link capacity is expressed as $C_{link}=\sum C_{sb}$, and $C_{sb}(k)$ is calculated as

$$C_{sb}(k)=\sum_{i=1}^{k} P(k,i)\cdot i. \quad (7)$$

where $P(k,i)$ means the probability that only i slots can be used in a spectrum block which has k slots, that is to say, the next traffic request needs more than $k-i$ slots while i spectrum slots have been occupied. For example, if $k=4$ and $i=3$, there are 4 cases, as shown in Fig. 4. **i).** One request with 3 slots. **ii).** Two requests and the first one requires 1 slot, the second requires 2 slots. **iii).** Two requests and first requires 2 slots, second requires 1 slots, this case is different from the **ii)** case. **iv).** Three requests and each requires 1 slot. Then next traffic request needs more than one slots, so $P(4,3)$ can be expressed as

$$P(4,3)=(p_r(3)+p_r(1)p_r(2)+p_r(2)p_r(1)+p_r^3(1))\cdot p_r(>1). \quad (8)$$

where $p_r(i)$ means the probability that a traffic needs i slots, and $p_r(>1)$ means that a traffic needs more than one slot.

In the network operation, different traffic requests require different bandwidths, we consider that the required bandwidth follows a uniform distribution, that is to say, the proba-

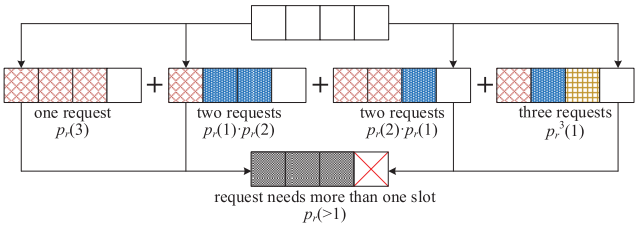


FIGURE 4. The example of $P(4, 3)$ calculation.

bility of requiring different number of slots is equal. For one spectrum block which has i spectrum slots, we can find that if T_n traffic requests occupy the block, it is equivalent to divide spectrum block into T_n sub blocks, there are total $\frac{(i-1)!}{(T_n-1)!(i-T_n)!}$ cases. Therefore, the capacity of the spectrum block can be calculated as

$$C_{sb}(k)=\sum_{i=1}^{k}\sum_{j=1}^{i}\frac{(i-1)!}{(j-1)!(i-j)!}\cdot p_r^j\cdot p_r(>k-i)\cdot i. \quad (9)$$

Currently, the highest line rate in EON is generally 100Gbps, and a maximum bandwidth of 100GHz is required, so we limit the maximum number of slots that traffic required to 8 when a slot bandwidth is 12.5GHz. If one spectrum block has more than 8 slots, we divided it into multiple sub blocks, and the capacity is calculated separately. For example, $C_{sb}(17) = 2\times C_{sb}(8) + C_{sb}(1)$. Besides, the link capacity is determined by the link state, if the link state is degraded, the maximum number of continuous slots in the link is reduced. Taking the CD increasing as an example, if the CD value is more than 125ps/nm, the link can only carry traffic with the maximum bit rate of 40Gbps. At that time, the link capacity is decreasing even if the link has enough unoccupied spectrum slots, the maximum number of continuous slots is limited at 4 (i.e., 50GHz), a spectrum block will be divided into several smaller blocks, and the block capacity changes correspondingly. For example, $C_{sb}(17)$ is calculated as $4\times C_{sb}(4) + C_{sb}(1)$.

B. LINK STATE-AWARE ROUTING ALGORITHM

In this subsection, we present the details of the LSA routing algorithm in Alg. 1, which is designed to search k-available routing paths (k-ASP) when a traffic request arrives. LSA ensures that all the found paths are competent to carry arriving traffic, which means that the CD and OSNR of the paths could meet the requirements and the paths can transmit data with QoT provisioning. As shown in Alg. 1, the link state is evaluated after CD and OSNR estimation, then the controller obtains CD and OSNR values of each link, which are represented as LS_D and LS_O, respectively. Link weight is determined by link capacity as $weight = N/C_{link}$, where N is the total number of spectrum slots in one link, and the involvement of link weight can achieve load balancing in the routing phase. One link has a small weight if it has a larger free capacity, and the lower weight path is the preferred routing path for traffic bearing. After finding a path, the weight coefficients of all links in this path is multiplied by a factor α (line7). The total routing path CD value is calculated as the cumulative value of all links (line9), the final OSNR is determined by the worst link state (line10), CD and OSNR requirements are presented in section IV. The available spectrum slots in the path are searched (line12-17), the available slot means this slot is unoccupied in all links among the path. Then whether the path p_t meets the requirements is determined. Only when there are enough available slots in the path, the CD and OSNR meet the requirements, the path can be recorded as a qualified path (line18-22). All the qualified paths are represented by P_T. SS_T indicates the index of unoccupied spectrum blocks in all paths. The number of iterations which is defined as *Iteration* is used to avoid endless loop in the path searching phase, and the LSA algorithm return an empty P_T if none suitable path can be found after several iterations.

Algorithm 1 LSA Routing Algorithm

Input: Network $G(V, E)$; Traffic request $T(v_s, v_d, bw)$; Slot state M; Link State $LS(LS_D, LS_O)$; Link Weight $weight$; CD Requirement $R_D(T)$; OSNR Requirement $R_O(T)$

Output: Routing Path P_T; Assigned Slot SS_T for traffic T

1: **function** findpath(G, T, M, k, LS)
2: $Iteration \leftarrow 1$; $sumM \leftarrow 0$; $i \leftarrow 1$
3: $disps \leftarrow 0$; $osnr \leftarrow \infty$
4: **while** $i < k$ **and** $Iteration < 2\times k$ **do**
5: $p_t \leftarrow shortestpath(G, v_s, v_d, weight)$
6: **for** each link e **in** p_t **do**
7: $weight(e) \leftarrow weight(e)\times\alpha$
8: $sumM \leftarrow sumM + M(e, :)$
9: $disps \leftarrow disps + LS_D(e)$
10: $osnr \leftarrow \min\{osnr, LS_O(e) - 10\log(||p_t||)\}$
11: **end for**
12: $fs \leftarrow find(sumM == 0)$
13: **for** $j = 1$ **to** length(fs) $- bw + 1$ **do**
14: **if** $fs(j + bw - 1) == fs(j) + bw - 1$ **then**
15: $avaslot \leftarrow avaslot$.append($fs(j)$)
16: **end if**
17: **end for**
18: **if** $avaslot \neq \varnothing$ **and** $disps \leq R_D$ **and** $osnr \geq R_O$ **then**
19: $P_T(i, :) \leftarrow p_t$
20: $SS_T(i, :) \leftarrow avaslot$
21: $i \leftarrow i + 1$
22: **end if**
23: $Iteration \leftarrow Iteration + 1$
24: **end while**
25: **return** P_T, SS_T
26: **end function**

C. FRAGMENTATION REDUCING SPECTRUM ALLOCATION ALGORITHM

Spectrum slots are allocated for lightpath establishment once routing paths are obtained. The spectrum allocation algorithm is implemented to select spectrum slots for traffic bearing.

In our work, we use link capacity reduction as a metric to evaluate the performance of different spectrum allocation solutions. The solution with the lowest capacity reduction will be selected as the final RSA solution to carry traffic. The link capacity calculation method has been shown in Section V-A. And the detail of our LSA-RSA strategy is designed as Alg. 2, which contains the details of the LSA routing algorithm and FR spectrum allocation algorithm.

While a traffic request arrives, routing paths are allocated by Alg. 1 (line2-3). The capacity of each link is calculated under the current network state and spectrum occupation state. Then, for each routing path, all the spectrum allocation schemes will cause link capacity reduction, and all the reduction values are calculated (line7-12), the scheme which owns minimum capacity reduction is reserved as the final scheme to allocate spectrum resources (line13-16). Once spectrum slots are occupied or the traffic is released, the spectrum state matrix is updated accordingly (line23-25). If the LSA routing algorithm cannot find any available routing path, the traffic request is blocked (line21).

Algorithm 2 LSA-RSA Strategy

Input: Same as Alg. 1

Output: Path Matrix Record PM; Spectrum Slot Matrix Record SSM for each traffic

1: Get the state of all links LS in network $G(V, E)$
2: For incoming traffic $T(v_s, v_d, bw)$
3: $[P_T, SS_T]$ =FINDPATH(G, T, M, LS)
4: $LC_{before} \leftarrow LinkCapacity(G, M, LS)$
5: $\Delta LC \leftarrow \infty$
6: **if** $P_t \neq \varnothing$ **then**
7: **for** each $P_T(i)$ **in** P_T **do**
8: **for** each $SS_T(i, j)$ in $SS_T(i, :)$ **do**
9: $M' \leftarrow M$
10: $M'(P_T(i), SS_T(i, j)) \leftarrow 1$
11: $LC_{after} \leftarrow LinkCapacity(G, M', QL)$
12: $\Delta C_{tem} \leftarrow LC_{before} - LC_{after}$
13: **if** $\Delta C_{tem} < \Delta LC$ **then**
14: $\Delta LC \leftarrow \Delta C_{tem}$
15: $PM \leftarrow P_T(i)$
16: $SSM \leftarrow SS_T(i, j)$
17: **end if**
18: **end for**
19: **end for**
20: **else**
21: traffic T is blocking
22: **end if**
23: $M(PM, SSM) \leftarrow 1$
24: **while** traffic T is releasing
25: $M(PM, SSM) \leftarrow 0$

The time complexity of LSA-RSA strategy can be divided into two parts, the LSA routing algorithm finds k routing paths, the complexity can be determined as $\mathbf{O}(k \times n^2)$, where n is the number of nodes in the network. In the FR algorithm, all available spectrum blocks in k routing paths are detected, the complexity is $\mathbf{O}(k \times N)$. Then, the capacity of each block is calculated, and all spectrum allocation schemes are compared with each other, the complexity is $\mathbf{O}(s)$, where s is the number of spectrum blocks. Obviously, s is smaller than N, so the complexity of spectrum allocation process can be determined as $\mathbf{O}(k \times N)$.

TABLE 3. Simulation parameters of communication system shown in Fig. 2.

Parameters	Values
Modulation format	4PAM
Size of PRBS	1024bit
Number of Loop	1
Length of SMF, DCF	60km, 30km
Attenuation of SMF, DCF	0.2dB/km, 0.6dB/km
Dispersion coefficient of SMF	16.5, 17, 17,5,...,21ps/(nm·km)
Dispersion coefficient of DCF	-33ps/(nm·km)
Gain of EDFA1, EDFA2	12dB, 18dB
Noise figure of EDDA1 and EDFA2	2, 2.5, 3,...,6.5dB

VI. SIMULATION AND ANALYSIS

To evaluate the performance of the parameter estimation model and LSA-RSA strategy, simulations are carried out in this section with Intel i5 CPU@3.30GHz and 8GB memory. In the part of the optical parameter estimation, we build the LightGBM model by using the Python programming language, then compare its performance with XGBoost model. The simulation of the proposed LSA-RSA strategy is carried out by MATLAB. Then the performance of the proposed LSA-RSA strategy is compared with several existing RSA strategies.

A. CD AND OSNR ESTIMATION RESULTS

In the part of CD and OSNR estimation, we first evaluate the impact of the training set size on estimation accuracy. The parameters of the optical communication system (shown in Fig. 2) are listed in Table 3. The CD values are between 990ps/nm and 1260ps/nm with the interval of 30ps/nm, and DCF could provide 990ps/nm dispersion compensation, so the residual CD ranges from 0ps/nm to 270ps/nm. OSNR values are between 30dB and 39dB, and the interval is 1dB. The training sample $s(i)$ is sampled as a 1×32768 vector and converted to AH record with 150 interval bins. In the end, each training sample consists of the AH and the maximum value of $s(i)$, expressed as a 1×151 vector. We collect 4000 training samples under different CD and OSNR pairs. Then different numbers of training samples are selected randomly to form training sets. After the training process, another 200 samples are used as the test set. Estimation accuracy is evaluated by the coefficient of determination [42], which is determined by R^2 and calculated as

$$R^2 = 1 - \frac{\sum_{i=1}^{K} \left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{K} \left(y_i - \bar{y}\right)^2}. \tag{10}$$

TABLE 4. Training time with different numbers of training samples (Unit: Second).

Model \ Training Samples	1000	1500	2000	2500	3000	3500	4000
LightGBM, CD Estimation	1.87	22.88	3.06	3.16	3.22	3.33	3.37
LightGBM, OSNR Estimation	1.66	2.56	2.89	2.92	3.05	3.08	3.25
XGBoost, CD Estimation	6.6	8.89	11.1	13.4	15.3	16.9	18.3
XGBoost, OSNR Estimation	6.99	9.49	11.7	13.9	16.1	18.1	19.2

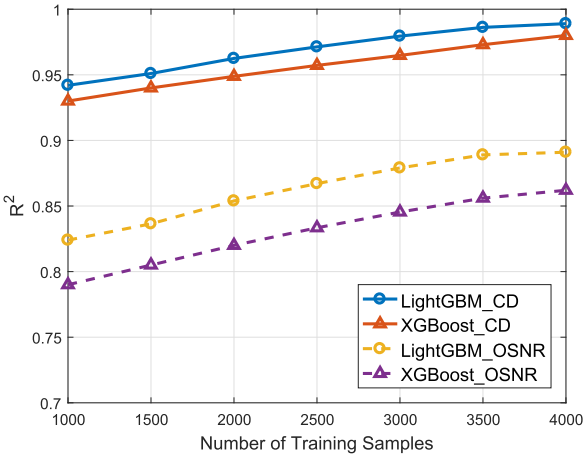


FIGURE 5. R^2 values under different numbers of training samples, where 'LightGBM_CD' indicates that CD is estimated by LightGBM model. Correspondingly, 'XGBoost_OSNR' indicates that OSNR value is estimated by XGBoost model.

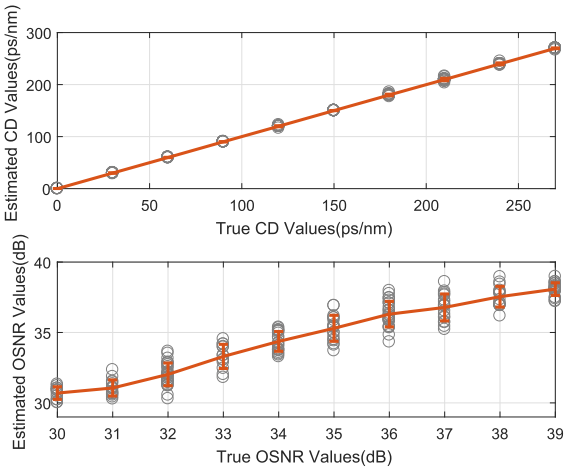


FIGURE 6. Errors of CD and OSNR estimation, grey circles represent estimated results of all test samples, and the error bar shows the mean values and the standard deviation values.

where K is the number of test samples, y_i and $\hat{y}_i$ represent true and estimated CD or OSNR values of the ith sample, respectively, and $\bar{y}$ is mean value of all samples in the test set.

To evaluate the estimation accuracy of the LightGBM model, we select XGBoost as a baseline, since it outperforms the other gradient boosting decision tree based models [40]. The models' parameter settings are introduced as follows. The numbers of estimators of two models are set to 400, the leaves number of LightGBM equals 63, while the maximum depth of the decision tree in the XGBoost model equals 6. The R^2 values with different numbers of training samples are shown in Fig. 5. As the number of training samples increases, the R^2 value increases, which means that the parameter estimation accuracy improves. The LightGBM model outperforms the XGBoost model in both CD and OSNR estimation. Under 4000 training samples, R^2 of the LightGBM model for CD and OSNR estimations equal 0.99 and 0.89, respectively. Besides, the results of R^2 show that LightGBM model converges faster, it remains stable when the number of sample size reaches 3500.

To show the estimation results more intuitively, the CD and OSNR estimation errors of the LightGBM model are plotted as Fig. 6, the error bar (plotted in solid line) shows the mean values and standard deviation values of estimated results, and grey circles represent the estimated values. It can be seen that the estimation error of CD is smaller than OSNR, the maximum average errors of CD and OSNR estimation are 0.28ps/nm (at 270ps/nm) and 0.68dB (at 30dB), respectively, where the average error is calculated as $\sum_{i=1}^{K} |y_i - \hat{y}_i|/K$.

The training time of LightGBM and XGBoost models are listed in Table 4, we disable the early stopping of two models in the training process. The results indicate that the training speed of the LightGBM model is 3-4 times faster than the XGBoost model under different numbers of samples. Based on the above simulation results, LightGBM gets higher estimation accuracy with lower training time and is more suitable for parameter estimation.

B. LSA-RSA PERFORMANCE

In this part, we evaluate the performance of proposed LSA-RSA strategy through NSF network topology [43], which has 14 nodes and 21 links. The traffic requests are generated where the arriving time follows the Poisson Process. λ traffic requests arrive per time unit, and duration time of each traffic follows a negative exponential distribution with a mean value of μ, so the traffic load can be expressed as $\lambda \cdot \mu$ Erlang. The number of required slots is denoted by $NS \in \{1, 2, \ldots, 8\}$, and the required slots of different traffics follow uniform distribution, i.e., $p_r(NS) = 0.125$. The number of spectrum slots of each link is set to 400. The LSA-RSA performance is evaluated after 12000 traffic requests arrive during the simulation. We first compare traffic failure probability (TFP) under different link state conditions. Under normal link state condition, which means that CD is small enough and OSNR is large enough, bandwidth blocking probability (BBP) of

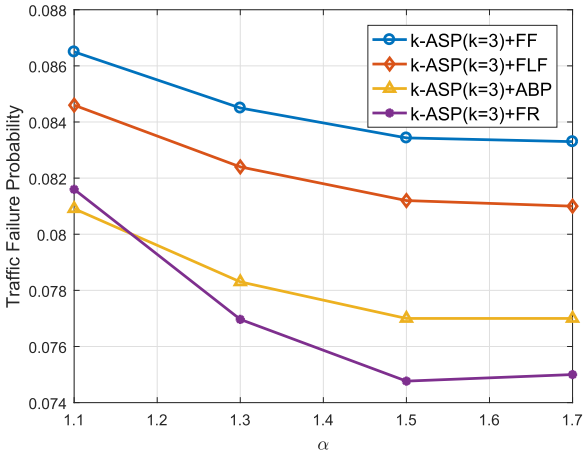


FIGURE 7. The TFP results of different RSA strategies with different α values.

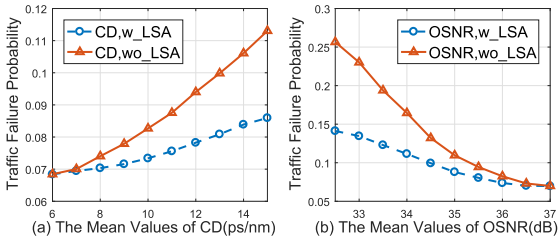


FIGURE 8. The TFP results under different link state conditions. (a) shows the TFP under different CD values and (b) shows TFP under different OSNR values. 'w_LSA' means LSA is used, and 'wo_LSA' means LSA is not considered in the routing algorithm.

different RSA strategies is compared. Spectrum fragmentation ratio (SFR) is measured under all link state conditions. In the end, we use standard deviation (SD) [44] to evaluate the performance of load balancing.

1) TFP: In EON, the prerequisite of the establishment of a lightpath is that there must be enough available spectrum resources and the state of all links meet the requirements. Otherwise, the traffic request will be blocked or traffic data cannot be transmitted correctly. In both two cases, traffic is failed, so we defined TFP as the probability of failed traffic transmission and calculate it as follows

$$TFP = \frac{\sum_T bw_T \times (bf_T || ls_T)}{\sum_T bw_T}. \tag{11}$$

where bw_T indicates the number of spectrum slots required by the traffic request T, and bf_T is traffic blocking flag, $bf_T = 1$ means request T is blocked because there are not enough spectrum slots, and 0 otherwise. Similarly, $ls_T = 1$ indicates that the routing path cannot support traffic due to the degradation of the link state, and ls_T=0 means traffic data can be transmitted correctly on the current path. The lower TFP means higher network throughput and more traffic can be transmitted.

The effect of α values on TFP in different RSA strategies is analyzed firstly, and the results are shown in Fig. 7, where α is used in the LSA routing algorithm to find multiple routing paths. The traffic load is set to 600Erlang, k-ASP (k=3) are found via LSA algorithm, FF, FLF, ABP based, and the proposed FR methods are used for spectrum allocation. We set α to 1.5 during our simulation since all strategies could almost obtain the best performance in this setting. The TFP with different k values is also measured (results are not shown in this paper), we find that TFP decreases from 0.08 (k=1) to about 0.07 (k=3) in normal link state, and then remains stable even if k increases. Therefore, we set k equals 3 in the LSA routing algorithm.

We compare TFP under different link state conditions. Firstly, the CD values of all links are set to follow normal distribution with mean values ranging from 6ps/nm to 15ps/nm, the variance always equals 25, and OSNR of all links are 37dB. Secondly, OSNR values follow normal distribution with mean values ranging from 32.5dB to 37dB, the variance equals 4, CD of all links are less than 5ps/nm. The routing algorithm without considering the link state is selected as the baseline, in which only the sufficient spectrum resources are considered when searching the routing path. FR spectrum allocation algorithm is always used to allocate spectrum slots. The results are shown in Fig. 8. When the link state is degraded, the LSA routing algorithm can provide better network throughput performance. Compared with the routing algorithm without LSA, the LSA algorithm can reduce TFP by up to 24% under different CD settings and up to 45% under different OSNR settings.

Then, the TFP results under different traffic loads are measured. We select three link state conditions: **i).** The CD values of all links follow normal distribution with the mean value of 20ps/nm and the variance of 25, and OSNR values are equal to 37dB. **ii).** The CD of all links are less than 5ps/nm, and OSNR values follow normal distribution with the mean value of 33dB and the variance of 4. **iii).** The CD and OSNR both follow normal distribution with mean values of 15ps/nm, 34dB and variances of 25, 4. The results are shown in Fig. 9. It can be seen that the LSA algorithm provides a better TFP performance, the impact of link state degradation could be compensated and we can get a performance closer to the normal link state. In the first two conditions, the LSA routing algorithm decreases TFP by about 58% and 52%. For the third link state condition, in which the link state degradation is slighter, the LSA algorithm can still decrease TFP by around 30%.

2) BBP: When the network operates normally, i.e., all links are qualified to bear any traffic request, the LSA algorithm only considers whether there are available spectrum resources when searching k-ASP. In this scenario, ls_T always equals 0, so TFP can be represented as BBP, which means the probability of blocking traffic. To compare the performance of LSA and FR algorithms, we select shortest path (SP), k-shortest path (k-SP, k=3) and one-available shortest path (k-ASP, k=1) as routing path finding algorithms, where the weight of links are calculated based on link capacity and load balancing is considered in all routing strategies. FF, FLF, and ABP based

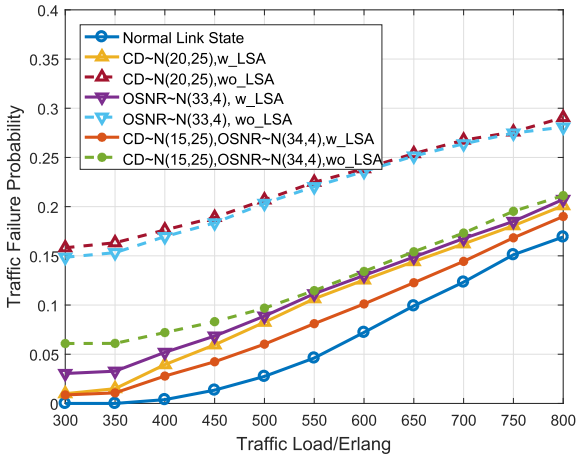


FIGURE 9. The TFP results with different traffic loads, $N(m, v)$ in the legend means the normal distribution with mean value of m and variance of v.

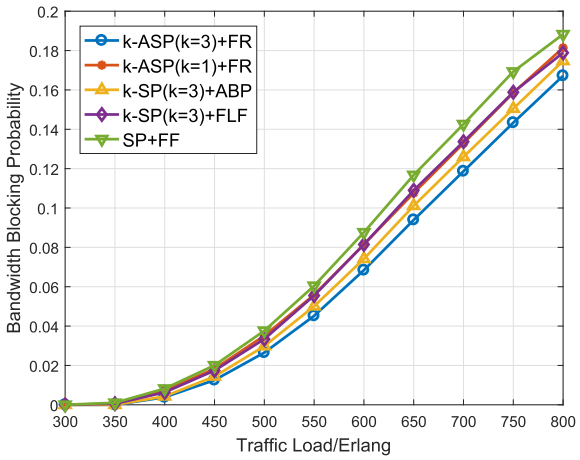


FIGURE 10. The BBP results of different RSA strategies with different traffic loads.

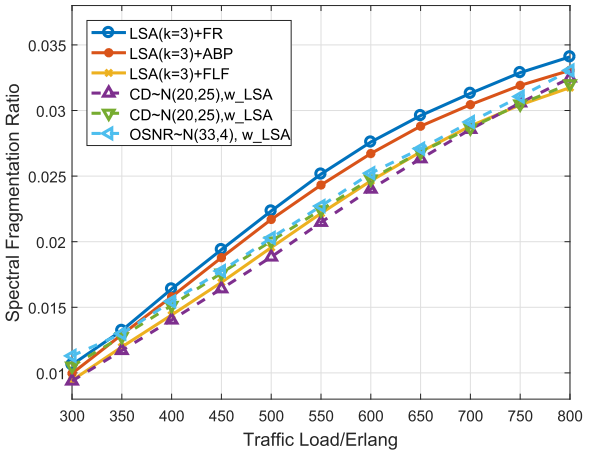


FIGURE 11. The SFR results with different traffic loads and link state conditions, FR algorithm is used for spectrum slots allocation.

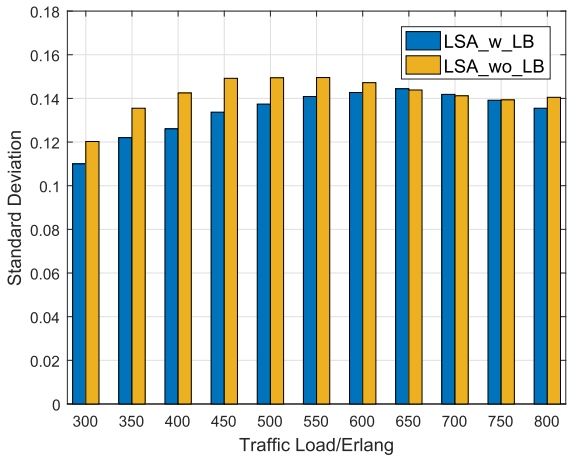


FIGURE 12. The SD results with different traffic loads. 'LSA_w_LB' means load balancing is used and 'LSA_wo_LB' means load balancing is not used.

methods are used for spectrum allocation. The BBP results are shown in Fig. 10. The proposed RSA strategy, which is denoted as k-ASP (k=3)+FR, performs best among all the RSA strategies, compared to SP+FF, it can increase the traffic load of 40Erlang while the BBP equals 10%.

3) SFR: The SFR indicates the proportion of the fragmentation spectrum in the network. In our work, the fragmentation spectrum means that the number of spectrum slots is less than 3 in a spectrum block, i.e., the bandwidth of this block is less than 37.5GHz. The SFR is calculated as

$$SFR = \frac{\sum_{e\in E}\sum_{k}|SS_{e,k}| \times sf_k}{|E|N}. \tag{12}$$

where $|SS_{e,k}|$ is the slots number of kth spectrum block SS_k in link e, and if the number of slots is less than 3 in SS_k, we set sf_k equals 1. $|E|$ is the number of links in the network, and N is the total number of slots in one link. The SFR results with different traffic loads and link state conditions are shown as Fig. 11. Considering the normal link state at first, the FR algorithm has a higher SFR than ABP and FLF. However, we should notice that FR provides the highest throughput, more traffic requests are established in EON, so it brings more fragmentary spectrum blocks. Secondly, SFR is lower when the link state is degraded, since fewer traffic requests are established in EON to guarantee the transmission quality, so there can be more free spectrum blocks which have more than 3 spectrum slots, and the number of fragmentary spectrum blocks decreases.

4) SD: We use SD as an indicator to determine the uniformity of traffic distribution, which can be measured based on the uniformity of spectrum occupation status and calculated as

$$SD = \sqrt{\frac{1}{|E|N^2}\sum_{e\in E}\left(\sum_{j=1}^{N}M(e,j) - \overline{M}\right)^2}. \tag{13}$$

$$\overline{M} = \frac{\sum_{e\in E}\sum_{j=1}^{N}M(e,j)}{|E|}. \tag{14}$$

TABLE 5. Consuming time under different RSA strategies and traffic loads (Unit: Second).

Strategies \ Traffic loads	300	350	400	450	500	550	600	650	700	750	800
LSA-RSA, normal link state	214.5	216.7	215.7	214.4	223.8	224.7	233	242.7	247.9	249	250.2
LSA-RSA, CD~N(20,25)	237.6	248.8	256	251	248.1	250.2	251.3	245.9	242	241	245.9
LSA-RSA, OSNR~N(33,4)	255.6	255.7	260	253.9	252.1	252.4	254.6	251.6	246.8	249.8	249
SP+FF	81.3	95.2	108.3	115.9	119.1	121.1	126.7	125.3	130.9	136.8	139.4
k-SP+FLF	180.6	185	186.8	188.6	193.7	204.1	207.7	211.2	222.2	221.5	229.2
k-SP+ABP	217.1	222.5	224.8	233.6	235.9	238.8	235.3	236.8	238.4	246.9	247

where $M(e, j)$ represents the occupation status of the jth slot in link e. The SD results are shown in Fig. 12. As we can see, our load balancing method decreases the SD in various traffic loads, indicating that the traffic distributes more uniformly in the network.

In the end, we measure the consuming time of different RSA strategies and traffic loads, the results are listed in Table 5. Under the high traffic load, the LSA-RSA strategy consumes almost the same time under different link state conditions, and the time is about the same as k-SP+FLF and k-SP+ABP.

VII. CONCLUSION

To improve the network throughput performance when the network state is degraded, resources should be allocated based on network and link state. In this paper, we proposed an LSA-RSA strategy. The link state was evaluated based on CD and OSNR pairs, which were estimated by the LightGBM model, and the link state was used as a metric in RSA. In the routing phase, we proposed the LSA routing algorithm aiming to find k-available routing paths, where the available path means that the path has sufficient spectrum resources and is capable to bear current traffic request. Then we proposed the FR algorithm to allocate spectrum slots aiming to minimize the link capacity reduction. The simulation results showed that the LightGBM could provide accurate CD and OSNR estimation results, and our LSA-RSA strategy could improve the network throughput especially when the link state is degraded.

REFERENCES

[1] H. Rastegarfar, L. A. Rusch, and A. Leon-Garcia, "Optical load-balancing tradeoffs in wavelength-routing cloud data centers," *J. Opt. Commun. Netw.*, vol. 7, no. 4, pp. 286–300, Mar. 2015.

[2] M. Jinno, H. Takara, B. Kozicki, Y. Tsukishima, Y. Sone, and S. Matsuoka, "Spectrum-efficient and scalable elastic optical path network: Architecture, benefits, and enabling technologies," *IEEE Commun. Mag.*, vol. 47, no. 11, pp. 66–73, Nov. 2009.

[3] I. Tomkos, S. Azodolmolky, J. Sole-Pareta, D. Careglio, and E. Palkopoulou, "A tutorial on the flexible optical networking paradigm: State of the art, trends, and research challenges," *Proc. IEEE*, vol. 102, no. 9, pp. 1317–1337, Sep. 2014.

[4] B. C. Chatterjee, N. Sarma, and E. Oki, "Routing and spectrum allocation in elastic optical networks: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1776–1800, 3rd Quart., 2015.

[5] F. S. Abkenar and A. G. Rahbar, "Study and analysis of routing and spectrum allocation (RSA) and routing, modulation and spectrum allocation (RMSA) algorithms in elastic optical networks (EONs)," *Opt. Switching Netw.*, vol. 23, no. 1, pp. 5–39, Jan. 2017.

[6] B. C. Chatterjee, S. Ba, and E. Oki, "Fragmentation problems and management approaches in elastic optical networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 183–210, 1st Quart., 2018.

[7] A. P. Vela, B. Shariati, M. Ruiz, F. Cugini, A. Castro, H. Lu, R. Proietti, J. Comellas, P. Castoldi, S. J. B. Yoo, and L. Velasco, "Soft failure localization during commissioning testing and lightpath operation," *J. Opt. Commun. Netw.*, vol. 10, no. 1, pp. A27–A36, Nov. 2017.

[8] G. Chauvel, "Dispersion in optical fibers," Anritsu, Kanagawa, Japan, Tech. Rep. MBCDPMD-WP01-0801-A4, 2008.

[9] Y. Wang, "Study of the impact of nonlinearities on advanced modulation formats in optical systems and networks," M.S. thesis, Coll. Opt. Sci., Arizona Univ., Tucson, AZ, USA, 2017.

[10] N. Dong-Nhat, M. A. Elsherif, and A. Malekmohammadi, "Investigations of high-speed optical transmission systems employing absolute added correlative coding (AACC)," *Opt. Fiber Technol.*, vol. 30, pp. 23–31, Jul. 2016.

[11] Y. Pointurier, M. Brandt-Pearce, S. Subramaniam, and B. Xu, "Cross-layer adaptive routing and wavelength assignment in all-optical networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 6, pp. 32–44, Aug. 2008.

[12] I. Sartzetakis, K. Christodoulopoulos, and E. Varvarigos, "Cross-layer adaptive elastic optical networks," *J. Opt. Commun. Netw.*, vol. 10, no. 2, pp. A154–A164, Jan. 2018.

[13] X. Cao, N. Yoshikane, I. Popescu, T. Tsuritani, and I. Morita, "Software-defined optical networks and network abstraction with functional service design," *J. Opt. Commun. Netw.*, vol. 9, no. 4, pp. C65–C75, Apr. 2017.

[14] Z. Dong, F. N. Khan, Q. Sui, K. Zhong, C. Lu, and A. P. T. Lau, "Optical performance monitoring: A review of current and future technologies," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 525–543, Jan. 15, 2016.

[15] M. S. Faruk, Y. Mori, and K. Kikuchi, "In-band estimation of optical signal-to-noise ratio from equalized signals in digital coherent receivers," *IEEE Photon. J.*, vol. 6, no. 1, pp. 1–9, Feb. 2014.

[16] J. Mata, I. de Miguel, R. J. Durán, N. Merayo, S. K. Singh, A. Jukan, and M. Chamania, "Artificial intelligence (AI) methods in optical networks: A comprehensive survey," *Opt. Switching Netw.*, vol. 28, pp. 43–57, Apr. 2018.

[17] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An overview on application of machine learning techniques in optical networks," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1383–1408, 2nd Quart., 2019.

[18] S. Behera, J. George, and G. Das, "Effect of transmission impairments in CO-OFDM based elastic optical network design," *Comput. Netw.*, vol. 144, pp. 242–253, Oct. 2018.

[19] S. Behera, A. Deb, G. Das, and B. Mukherjee, "Impairment aware routing, bit loading, and spectrum allocation in elastic optical networks," *J. Lightw. Technol.*, vol. 37, no. 13, pp. 3009–3020, Jul. 1, 2019.

[20] A. Rosa, C. Cavdar, S. Carvalho, J. Costa, and L. Wosinska, "Spectrum allocation policy modeling for elastic optical networks," in *Proc. High Capacity Opt. Netw. Emerg./Enabling Technol.*, Istanbul, Turkey, Dec. 2012, p. 242.

[21] W. Fadini, B. C. Chatterjee, and E. Oki, "A subcarrier-slot partition scheme with first-last fit spectrum allocation for elastic optical networks," *Comput. Netw.*, vol. 91, pp. 700–711, Nov. 2015.

[22] H.-L. Liu, L. Lv, Y. Chen, and C. Wei, "Fragmentation-avoiding spectrum assignment strategy based on spectrum partition for elastic optical networks," *IEEE Photon. J.*, vol. 9, no. 5, pp. 1–13, Oct. 2017.

[23] X. Chen, J. Li, P. Zhu, R. Tang, Z. Chen, and Y. He, "Fragmentation-aware routing and spectrum allocation scheme based on distribution of traffic bandwidth in elastic optical networks," *J. Opt. Commun. Netw.*, vol. 7, no. 11, pp. 1064–1074, Oct. 2015.

[24] D. Amar, E. Le Rouzic, N. Brochier, J.-L. Auge, C. Lepers, and N. Perrot, "Spectrum fragmentation issue in flexible optical networks: Analysis and good practices," *Photonic Netw. Commun.*, vol. 29, no. 3, pp. 230–243, Mar. 2015.
[25] D. Sharma and S. Kumar, "Evaluation of network blocking probability and network utilization efficiency on proposed elastic optical network using RSA algorithms," *J. Opt. Commun.*, to be published.
[26] D. Sharma and S. Kumar, "Network blocking probability-based evaluation of proposed spectrum assignment strategy for a designed elastic optical network link," *J. Opt.*, vol. 47, no. 4, pp. 496–503, Aug. 2018.
[27] H. Beyranvand and J. A. Salehi, "A quality-of-transmission aware dynamic routing and spectrum assignment scheme for future elastic optical networks," *J. Lightw. Technol.*, vol. 31, no. 18, pp. 3043–3054, Sep. 2013.
[28] B. C. Chatterjee and E. Oki, "Dispersion-adaptive first–last fit spectrum allocation scheme for elastic optical networks," *IEEE Commun. Lett.*, vol. 20, no. 4, pp. 696–699, Apr. 2016.
[29] N. Dharmaweera, L. Yan, M. Karlsson, and E. Agrell, "An impairment-aware resource allocation scheme for dynamic elastic optical networks," in *Proc. Opt. Fiber Commun. Conf.*, Los Angeles, CA, USA, 2017, pp. 1–3.
[30] A. Fontinele, I. Santos, J. N. Neto, D. R. Campelo, and A. Soares, "An efficient IA-RMLSA algorithm for transparent elastic optical networks," *Comput. Netw.*, vol. 118, pp. 1–14, May 2017.
[31] J. Thrane, J. Wass, M. Piels, J. C. M. Diniz, R. Jones, and D. Zibar, "Machine learning techniques for optical performance monitoring from directly detected PDM-QAM signals," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 868–875, Feb. 15, 2017.
[32] Y. Huang, Y. Chen, and J. Yu, "Optical performance monitoring of 56 Gbps optical PAM 4 signal using artificial neural networks," in *Proc. Asia Commun. Photon. Conf.*, 2017, pp. 1–3.
[33] D. Wang, M. Zhang, J. Li, Z. Li, J. Li, C. Song, and X. Chen, "Intelligent constellation diagram analyzer using convolutional neural network-based deep learning," *Opt. Express*, vol. 25, no. 15, pp. 17150–17166, Jul. 2017.
[34] L. Guesmi and M. Menif, "Modulation formats recognition technique using artificial neural networks for radio over fiber systems," in *Proc. 17th Int. Conf. Transparent Opt. Netw. (ICTON)*, Jul. 2015, pp. 1–4.
[35] X. Sun, S. Su, J. Wei, X. Guo, and X. Tan, "Monitoring of OSNR using an improved binary particle swarm optimization and deep neural network in coherent optical systems," *Photonics*, vol. 6, no. 4, p. 111, Oct. 2019.
[36] T. Tanimura, T. Hoshida, T. Kato, S. Watanabe, and H. Morikawa, "Data-analytics-based optical performance monitoring technique for optical transport networks," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2018, pp. 1–3.
[37] T. Tanimura, T. Hoshida, T. Kato, S. Watanabe, and H. Morikawa, "Convolutional neural network-based optical performance monitoring for optical transport networks," *J. Opt. Commun. Netw.*, vol. 11, no. 1, pp. A52–A59, Oct. 2018.
[38] T. Tanimura, T. Hoshida, T. Kato, and S. Watanabe, "OSNR estimation providing self-confidence level as auxiliary output from neural networks," *J. Lightw. Technol.*, vol. 37, no. 7, pp. 1717–1723, Apr. 1, 2019.
[39] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T. Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 3146–3154.
[40] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2016, pp. 785–794.
[41] F. Da Ros, V. Cristofori, O. Ozolins, M. E. Chaibi, X. Pang, G. Jacobsen, S. Popov, M. Galili, L. K. Oxenløwe, and C. Peucheret, "4-PAM dispersion-uncompensated transmission with micro-ring resonator enhanced 1.55-μm DML," in *Proc. Conf. Lasers Electro-Optics*, Oct. 2017, pp. 1–2.
[42] N. R. Draper and H. Smith, "The general regression situation," in *Applied Regression Analysis*, 3rd ed. New York, NY, USA: Wiley, 1998, pp. 138–141.
[43] D. S. Yadav, S. Rana, and S. Prakash, "Hybrid connection algorithm: A strategy for efficient restoration in WDM optical networks," *Opt. Fiber Technol.*, vol. 16, no. 2, pp. 90–99, Mar. 2010.
[44] D. Batham, D. S. Yadav, and S. Prakash, "Least loaded and route fragmentation aware RSA strategies for elastic optical networks," *Opt. Fiber Technol.*, vol. 39, pp. 95–108, Dec. 2017.

YANG ZHOU was born in 1993. He received the B.S. degree in communication engineering from the Qingdao University of Science and Technology, in 2016. He is currently pursuing the Ph.D. degree in communication and information systems with Beijing Jiaotong University, Beijing, China. His major study focuses on optical transport networks and resource allocation.

QIANG SUN was born in 1959. He received the M.S. degree in telecommunication automation and the Ph.D. degree in electromagnetic field, microwave technology, and microwave communication engineering from the Beijing University of Posts and Telecommunications, in 1978 and 1986, respectively. He is currently a Professor with the Railway Network Joint Laboratory, Institute of Cyber Security, Telecommunications, and Information System, Beijing Jiaotong University, Beijing, China. His major research interests include optical fiber communication, optical fiber sensor, and optical network technology.

SIYU LIN received the B.E. and Ph.D. degrees in electronics engineering from Beijing Jiaotong University, China, in 2007 and 2013, respectively. From 2009 to 2010, he was an Exchange Student with the Universidad Politénica de Madrid, Madrid, Spain. From 2011 to 2012, he was a Visiting Student with the University of Victoria, BC, Canada. Since 2016, he has been with Beijing Jiaotong University, where he is currently an Associate Professor. His main research interests include wireless communication networks and railway mobile communications. He received the First Class Award of Science and Technology in Railway, in 2017. He has served as a Track Co-Chair of the IEEE VTC2020-Fall.