IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 28, 2021, accepted February 7, 2021, date of publication February 19, 2021, date of current version March 2, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3060375

Temporal and Spectral 2D Fragmentation-Aware RMSA Algorithm for Advance Reservation Requests in EONs

YAN LIU, RONGXI HE, (Member, IEEE), SHICHENG WANG, AND CUNQIAN YU
College of Information Science and Technology, Dalian Maritime University, Dalian 116026, China

Corresponding authors: Rongxi He (hrx@dlmu.edu.cn) and Cunqian Yu (yucunqian@dlmu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61371091, Grant 61801074, and Grant 61971083; in part by the Dalian Science and Technology Innovation Fund under Grant 2019J11CY015; in part by the Science General Foundation of Chinese Postdoctoral under Grant 2019M661074; and in part by the Liaoning Natural Science Foundation under Grant 2019-BS-021.

ABSTRACT In this paper, we explore the problem of routing, modulation and spectrum assignment (RMSA) for advance reservation (AR) requests in elastic optical networks (EONs), with the objective to further reduce the blocking probability and improve the spectrum utilization. On the one hand, we propose a minimum resource consumption routing policy (MRC) for leaving more free resources for future requests. On the other hand, we introduce a two-dimensional (2D) path resource model, in which each AR request can be transformed into a rectangle block whose length and width are related to request bandwidth and request duration respectively. Therefore, the spectrum assignment problem of AR requests is simplified to the two-dimensional rectangle packing problem. Furthermore, we design several factors to evaluate the two-dimensional fragmentation caused by a rectangle packing operation, where we consider the fragmentation situation in the rectangle block neighborhood for the first time. Based on these factors, we propose a two-dimensional fragmentation-aware spectrum assignment strategy (2D-FA). Through the combination of MRC and 2D-FA, we propose a novel RMSA algorithm named MRC-2D-FA for dynamically provisioning AR requests in EONs. Simulation results show that MRC-2D-FA can achieve lower bandwidth blocking probability and higher spectrum utilization compared with several well-performed benchmark algorithms.

INDEX TERMS Elastic optical networks (EONs), routing, modulation and spectrum assignment (RMSA), advance reservation (AR), two-dimensional (2D) fragmentation, rectangle block.

I. INTRODUCTION

Today, with the emergence of various new applications such as 5G mobile services, Internet of things (IoT), high-definition videos, and grid computing, Internet traffic is growing exponentially and urgently needs optical networks to provide larger network capacity. Owing to the use of rigid modulation format and fixed coarse-grained wavelength channels, conventional wavelength division multiplexing (WDM) networks [1] result in a serious waste of spectrum resources. To alleviate this problem, elastic optical networks (EONs) [2] have been proposed to support tremendous traffic and heterogeneous requests in recent years. Based on Orthogonal Frequency Division Multiplexing (OFDM) and distance-adaptive technology [3], the EONs can assign just enough spectrum resources to meet the bandwidth requirement of each incoming request. This greatly reduces the spectrum consumption and improves the utilization of spectrum resources. Therefore, it is attractive for EONs to overcome the shortcomings of WDM networks.

While bringing great benefits, the EONs are also encountering some new issues and challenges. Among them, one of the most important issues is routing, modulation and spectrum assignment (RMSA) problem [4]. When modulation is not considered, RMSA is simplified to routing and spectrum assignment (RSA) problem. What RMSA does for each connection request is to find an optimal path, configure an appropriate modulation format, and allocate sufficient contiguous spectrum slots along the path. In the RMSA process, each request needs to follow the spectrum contiguity and

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

continuity constraints [5]. The spectrum contiguity constraint requires that the reserved spectrum slots must neighbor to each other on each link of the selected path, while the spectrum continuity constraint requires that the reserved spectrum slots must be aligned along the selected path.

However, when an EON considers dynamic traffic scenarios, the traffic demands arrive and leave the network randomly. As time goes by, the free spectrum resources will become more and more fragmented. When the free contiguous slots are insufficient to meet bandwidth requirements or cannot be aligned along the path, these slots will be unusable, which is called fragmentation problems [6]. When the number of fragments in the spectrum is large enough, it may cause a large number of requests to find no available spectrum resources, resulting in higher blocking probability and lower spectrum utilization.

In order to reduce fragmentation in dynamic scenarios, two methods have been extensively studied. The first method is defragmentation [7], which is used to remedy the spectrum after fragmentation occurs, but this method may cause connection interruption. The second method is fragmentation awareness [8], which is dedicated to avoiding fragmentation generation before each connection is established.

Traditionally, most of the researches on fragmentation-aware RMSA algorithms in EONs only involve immediate reservation (IR) requests [8]–[12]. However, in recent years, some new applications, such as cloud computing, video conferencing, data backups, and so on, are driving advance reservation (AR) services [13] to become an increasingly important part of the network. The biggest difference between AR requests and IR requests is that when they arrive at the network, IR requests need to be served immediately, while AR requests allow a specified time delay before obtain services. This feature of AR requests brings two benefits: one is that some applications supporting initial delay tolerance (such as data backup) can be planned more effectively in the future; the other is that some urgent real-time applications (such as video conference) can be arranged in advance. Therefore, we expect that it is necessary for the EONs to support AR requests well.

However, the scheduling and planning of AR requests involve two dimensions: time and frequency. This makes the service provision of related applications more complicated. In particular, IR requests only generate spectrum fragments, while AR requests introduce temporal and spectral two-dimensional (2D) fragmentation [14] into the network, which deteriorates network performance seriously. Therefore, it is crucial for us to develop a 2D fragmentation-aware RMSA algorithm to provision AR requests efficiently.

Previously, for the purpose of provisioning AR requests in EONs, the researchers have proposed some RMSA algorithms [14]–[20], which will be discussed in detail in related work of this paper. These algorithms can reduce 2D fragmentation, but they only care about the fragmentation condition along the frequency axis or time axis, which is not comprehensive enough. To the best of our knowledge, the fragmentation status of the neighboring area surrounding the pre-assigned resources is also crucial to assess the quality of an RMSA solution, but none of the previous studies has taken it into consideration. Therefore, we will conduct an in-depth study for this essential factor. Furthermore, we propose a novel RMSA algorithm named MRC-2D-FA to further decrease 2D fragmentation, reduce blocking probability, and improve spectrum utilization. The main contributions of this paper can be summarized as follows.

1) On the one hand, we propose a minimum resource consumption routing policy (MRC) as well as combining the advantages of minimum hop routing and shortest distance routing. This policy has low time complexity and helps reduce time-frequency resource consumption.
2) On the other hand, we introduce a two-dimensional link resource model, which can describe the spectrum availability along the time axis. Based on this model, we can easily derive the two-dimensional path resource model, which reflects the availability of time-frequency resources along a routing path. Then, for each AR request, we can map the request demands (bandwidth and duration) into a rectangle block, and then the problem of searching for feasible RMSA solutions on the path resource model can be transformed into a 2D rectangle packing problem. To evaluate the performance of an available rectangle block, we introduce three factors, which are described as follows.
    a. *The increased number of runs*. This factor reflects the change of fragmentation status along the frequency axis and time axis after a rectangle block resource is reserved. By minimizing the increased number of runs, we can optimize the fragmentation condition on the rows and columns where the rectangle block is located.
    b. *The neighborhood matching degree*. This factor reflects the fragmentation status of the time-frequency resources neighboring the rectangle block. By minimizing the neighborhood matching degree, we can make the occupied time-frequency resources more contiguous, thus further reducing 2D fragmentation.
    c. *The distance to the frequency boundary*. The reason for introducing this factor is that we expect the selected rectangle block to be as close as possible to the two boundaries of the frequency axis. As mentioned in [18], the benefit of the factor is that the free resources on each link are concentrated in the center area on the frequency axis. In the long run, this helps alleviate spectrum continuity constraint.

   Based on these three factors, we design a two-dimensional fragmentation-aware spectrum assignment strategy (2D-FA). Note that MRC and 2D-FA are the routing part and spectrum assignment part of the MRC-2D-FA algorithm, respectively.

3) Lastly, through simulation, we demonstrate that the proposed MRC-2D-FA algorithm can further reduce bandwidth blocking probability and improve spectrum utilization compared with several benchmarks.

The rest of this paper is organized as follows. Section II reviews the related work. In Section III, we describe the network model and the resource model, and then introduce three factors to evaluate each spectrum assignment scheme. To improve spectrum utilization and reduce blocking probability, we propose a dynamic MRC-2D-FA RMSA algorithm for provisioning AR requests in EONs in Section IV. Section V shows the simulation results. Finally, we summarize this paper in Section VI.

II. RELATED WORK

A. CLASSIFICATION OF AR REQUESTS

In 2012, Charbonneau *et al.* in [13] divided AR requests into two types: known duration and unknown duration. With known duration, AR requests can be further divided into three types: STSD-fixed, STSD-flexible, and UTSD. For STSD-fixed requests, the start time is fixed and cannot be earlier or later. Compared with STSD-fixed requests, the start time of STSD-flexible requests is flexible and can slide in a pre-specified time window. The scope of the time window can be determined by declaring the minimum and maximum start time. UTSD requests are similar to STSD-flexible requests, but do not explicitly specify the start time. UTSD requests specify a deadline, as long as the required tasks can be completed before the deadline.

B. RSA/RMSA ALGORITHMS FOR AR REQUESTS

In EONs, to provision AR requests, the researchers have proposed a few RSA/RMSA algorithms in recent years. In 2013, Lu *et al.* first investigated the issue of AR service provisioning in EONs [14]. To improve spectrum efficiency and reduce request blocking, they designed several assignment strategies for deadline-driven AR requests (UTSD). Before long they realized a revenue-driven AR provisioning policy in software-defined elastic optical networks (SD-EONs) [15].

To reduce 2D fragmentation, Wang *et al.* in [16] designed a 2D fragmentation metric by considering the request duration information. Based on this metric, they developed a two-phase RMSA algorithm named the max volume selectivity algorithm (MVS) for hybrid IR/AR requests. In [17], Chen *et al.* introduced a notion called available time-frequency consecutiveness (TSC) to evaluate fragmentation degree along time axis and spectrum axis, then based on TSC, they proposed several online RSA algorithms to decrease blocking probability for deadline-driven AR requests. Besides, a farsighted RMSA algorithm [18] named Min-RDDR was proposed to reduce AR blocking in EONs, in which a novel metric was introduced to measure the pros and cons of all possible assignment schemes. Min-RDDR can decrease 2D fragmentation and reduce AR blocking effectively by selecting the scheme with the smallest metric value.

Lu *et al.* designed a proactive AR scheduling algorithm [19] to coordinate service provisioning of IR and AR requests. Zhu *et al.* in [20] proposed a dynamic time and spectrum fragmentation-aware RMSA algorithm for IR requests and AR requests. In addition, Sugihara *et al.* proposed a novel spectrum partitioning method [21] to reduce spectrum fragmentation and mitigate IR/AR service conflicts in EONs. The method partitions the spectrum into multiple prioritized areas, and each area is used to accommodate such requests that require the same number of spectrum slots. This is beneficial to reduce 2D fragmentation. Furthermore, each priority area is divided into two sub-areas, one is dedicated to IR requests, and another is shared by IR and AR requests. This helps reduce IR service degradation caused by AR requests.

Some researchers have discussed static provisioning of AR requests in EONs. Chen *et al.* in [22] first established an integer linear program (ILP) model for the problem, then proposed a heuristic RSA algorithm for AR static provisioning. Besides, Wang *et al.* in [23] proposed a re-provisioning optimization algorithm (RPO) with three policies to dynamically re-provision AR requests. In [24], Afsharlar *et al.* proposed a novel RSA technology called delayed spectrum assignment (DSA) to provision AR requests in EONs. Unlike traditional RSA, DSA allows to delay the allocation of resources as long as the specified start time has not been reached.

However, the above algorithms still have a few shortcomings in reducing 2D fragmentation. Some of them [16], [17] can reduce two-dimensional fragments to a certain extent by considering all time-frequency resources on the selected path, but the corresponding computational complexity is too high. Although other algorithms [19], [20] consider the cutting degree of free resources along frequency axis and time axis caused by each spectrum allocation, this method is difficult to evaluate such assignment schemes without causing cutting of time-frequency resources.

C. DIFFERENCES FROM PREVIOUS WORK

In fact, the purpose of reducing 2D fragmentation is to make the free time-frequency resources more contiguous and continuous to accommodate more traffic demands. In this paper, based on this goal, we will investigate the approach for further reducing 2D fragmentation. Notably, the process of reserving time-frequency resources for an AR request is similar to filling a rectangle block in the two-dimensional plane, thus the problem of reducing 2D fragmentation can be seen as how to leave more free large-area zones for future AR requests in the two-dimensional plane.

Moreover, in the spectrum assignment process, the free time-frequency resources closest to the pre-assigned resources suffer the most serious spectrum contiguity loss, just as the bomb explosion always causes the most serious damage to the nearest locations of the bomb. Based on the above inspiration, we propose two feasible suggestions to reduce 2D fragmentation occurrence. On the one hand we consider preferential selection of rectangle blocks adjacent to less free time-frequency resources, which helps to

alleviate the damage of neighboring free resources around the rectangle blocks caused by spectrum assignment; and on the other hand, we hope that neighboring free resources around the rectangle blocks are sufficiently contiguous so that fewer 2D fragments occur after the rectangle block is occupied. Motivated by the above considerations, we propose a novel RMSA algorithm for dynamically provisioning AR requests in EONs. As shown later in this paper, our proposal is beneficial to the decline of 2D fragmentation, so the blocking probability of AR requests can be reduced further.

III. PROBLEM DESCRIPTION

A. NETWORK MODEL

The EON network topology can be modeled as a directed graph $G(V, E)$, where V is the node set, E is the set of fiber links. On each fiber link, the frequency resources can be divided into F spectrum slots. The capacity of each spectrum slot is 12.5 GHz, denoted by C_{slot}.

For a pending AR request, we denote it as a tuple $AR(s, d, t_a, H, C, w)$, where s and d are the source and destination nodes, t_a represents the arrival time, H is the holding time, C denotes that the line rate is C Gb/s, and w is the start time window that ranges from the earliest start time t_{es} to the latest start time t_{ls}. When the value of t_{es} is equal to t_{ls}, it denotes that the start time is a fixed value and the request is a STSD-fixed AR request, such as video conferencing, etc. On the other hand, when the value of t_{es} is smaller than t_{ls}, the request is a STSD-flexible AR request, such as large file transfer, etc.

To provision an AR request, we need to convert its transmission rate C Gb/s into N contiguous spectrum slots. In this conversion process, we adopt a distance-adaptive strategy [3] to improve the spectrum efficiency of the network. Based on this strategy, we can configure the appropriate modulation format for each request according to path distance, so as to reduce the number of required spectrum slots. Once the modulation format is determined, we can obtain the required number of spectrum slots N, which can be calculated as follows [14].

$$N = \left\lceil \frac{C}{M_P \cdot C_{slot}} \right\rceil + N_g, \tag{1}$$

where M_P denotes the modulation format of path P, N_g is the number of guard-band spectrum slots, and $\lceil \cdot \rceil$ means the ceiling integer. Note that for the different routing paths, the modulation format M_P may be different. In this paper, $M_P = 1,2,3,4$ means that the modulation formats are BPSK, QPSK, 8QAM and 16QAM, respectively.

B. RESOURCE MODEL

In this subsection, to reflect the spectrum condition in the time domain, we establish a two-dimensional link resource model to describe the time-frequency resources on each fiber link. Based on the link resource model, we can easily get the two-dimensional path resource model of a routing path. In each candidate path, we map the required time-frequency resources (bandwidth and duration) of an AR request into a fixed-size rectangle block on the path resource model.

Figure 1 visually shows a two-dimensional link resource model, which demonstrates the spectrum occupancy along the time axis on a fiber link. The horizontal and vertical axes represent frequency and time, respectively. On the frequency axis, the frequency resources are quantized as F spectrum slots, which are numbered as 1, 2, ……, F in sequence. Note that, we assume that the EONs work in a time-discrete manner [13]. Therefore, on the time axis, the time resources on each link can be divided into T time slots, which are numbered as $1, 2, \ldots\ldots, T$. In which, T denotes the maximum number of look-ahead time slots that can be observed by the network operator. For simplicity, we assume that AR requests only can be scheduled at the beginning of a time slot.

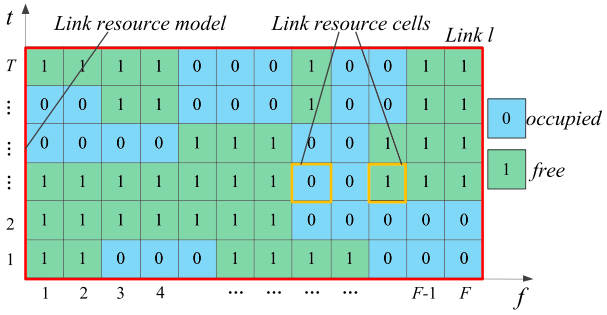


FIGURE 1. An example of a two-dimensional link resource model.

To further describe the link resource model, we define each small grid on the link resource model as a "*link resource cell*", which is the minimum unit of time-frequency resources on each link. Each link resource cell has two attributes, which appear as a spectrum slot along the frequency axis and a time slot along the time axis. In addition, each link resource cell has two states, namely free and occupied. When a link resource cell is free, it is available for the incoming requests. On the contrary, an occupied link resource cell cannot be used for any incoming request. In Fig. 1, the free and occupied link resource cells are shown in the green and blue areas, respectively. The state of a link resource cell can be expressed as

$$u_{t,f}^{l} = \begin{cases} 0, & occupied \\ 1, & free, \end{cases} \tag{2}$$

where variable $u_{t,f}^{l}$ indicates the state of a link resource cell at time slot t and spectrum slot f on link l. The value of $u_{t,f}^{l}$ equal to 0 indicates that the link resource cell is occupied, while the value of $u_{t,f}^{l}$ equal to 1 indicates that the link resource cell is free. Furthermore, we can obtain the state matrix of the link resource model, which can be expressed as follows.

$$U_{T,F}^{l} = \left[u_{t,f}^{l} \right]_{T \times F}, \tag{3}$$

where $T \times F$ indicates that the matrix has T row time slots and F column spectrum slots.

EONs require that the available resources allocated for requests must be aligned along the candidate path. Therefore, it is necessary for us to determine the path resource model, which can reflect the resource occupation along the selected path. Suppose we find a routing path P for an AR request. Then we can easily determine the path resource model by the following two Formulas.

$$U^P_{T,F} = \left[u^P_{t,f}\right]_{T\times F}, \tag{4}$$

$$u^P_{t,f} = \prod_{l\in P} u^l_{t,f}. \tag{5}$$

In Formula (4), $U^P_{T,F}$ represents the state matrix of the path resource model. Note that, being similar to the concept of link resource cells, we name each element of the path resource model as a "*path resource cell*", denoted by $u^P_{t,f}$. The state value of $u^P_{t,f}$ can be determined by Formula (5). A path resource cell is free only when the value of $u^P_{t,f}$ is equal to 1. It means that the free link resource cells at the same position on each link of path P can be aligned with each other. When the value of $u^P_{t,f}$ is equal to 0, the corresponding path resource cell is occupied and cannot be used for future requests.

Figure 2 depicts how to determine the path resource model of a candidate path. Figure 2(a) shows a routing path $P^1_{A,C}$, which includes two links $l_{A,B}$ and $l_{B,C}$. Figure 2(b) and Figure 2(c) illustrate the link resource model of link $l_{A,B}$ and link $l_{B,C}$, respectively. Based on these two link resource models, we can determine the path resource model of path $P^1_{A,C}$ by Formulas (4)-(5), as illustrated in Fig. 2(d). In which, free path resource cells are shown in green, and occupied path resource cells are shown in blue.

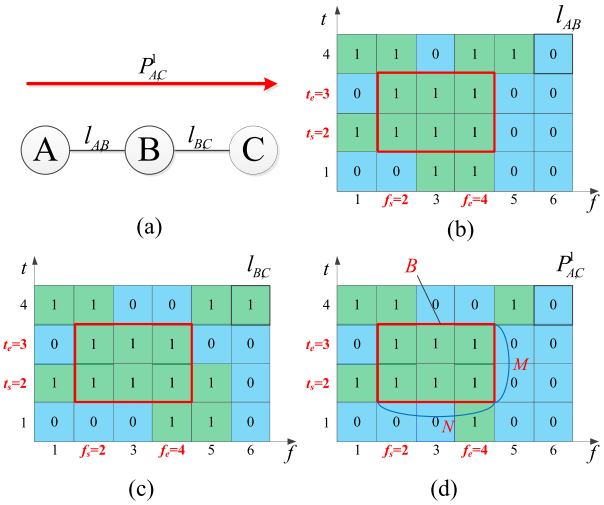


FIGURE 2. An example showing how to determine the path resource model of a path. (a) The routing path. (b-c) The link resource models. (d) The path resource model.

In this paper, we use the path resource model to schedule the available resources for AR requests. For a pending AR request $AR(s, d, t_a, H, C, w)$, its required resources can be mapped into a rectangle block of size $M \times N$. In detail, the length of the rectangle block is denoted by N, which means that the required line rate C is mapped into N spectrum slots in the frequency dimension. In different candidate paths, the length N of the rectangle blocks may be different according to Formula (1). The width of the rectangle block is represented by M, which indicates that the required holding time H is mapped into M time slots in the time dimension. Note that, N should be greater than or equal to the number of required spectrum slots for actual data transmission, and M time slots should be greater than or equal to the actual required duration. Each candidate rectangle block can be denoted as follows.

$$B = \left(P^k_{s,d}, f_s, f_e, t_s, t_e\right). \tag{6}$$

In which, $P^k_{s,d}$ denotes the k-th candidate path from source node s to destination node d. $P^k_{s,d}$ is the path where block B is located. The index k is taken in the interval $[1, K]$, and K is the number of candidate paths. On the path resource model of path $P^k_{s,d}$, f_s and t_s are the indexes of the start spectrum slot and the start time slot of block B, respectively. Then according the block size $M \times N$, the indexes of the end spectrum slot and end time slot of the block can be calculated as $f_e = f_s + N - 1$ and $t_e = t_s + M - 1$, respectively. Note that, the candidate rectangle block B must respect the contiguity and continuity constraints. The contiguity constraint demands that M time slots and N spectrum slots of a rectangle block remain contiguous along the time axis and the frequency axis, respectively. The continuity constraint demands that $M \times N$ free link resource cells on each candidate link requires to maintain alignment along path $P^k_{s,d}$.

Moreover, the spectrum assignment problem for an AR request can be transformed into a 2D rectangle packing problem. In other words, we can search for appropriate rectangle blocks to accommodate each incoming AR request on the path resource model. For example, we suppose that an AR request requires a rectangle block of size 2×3 on path $P^1_{A,C}$ in Figure 2. In Fig. 2(d), there is a candidate rectangle block B that can satisfies the request requirement. On the block, the start spectrum slot index $f_s = 2$, the end spectrum slot index $f_e = 4$, the start time slot index $t_s = 2$, and the end time slot index $t_e = 3$, respectively. Obviously, the candidate rectangle block B follow the contiguity and continuity constraints.

C. THE METRICS FOR RECTANGLE BLOCKS

In this subsection, we first introduce three factors to describe the quality of the candidate rectangle block, which are 1) *the increased number of runs*; 2) *the neighborhood matching degree*; 3) *the distance to the frequency boundary*. Then based on these three factors, we design two metrics to find the optimal rectangle blocks for AR requests during spectrum assignment. Next, we will describe the above three factors in detail.

The first factor is *the increased number of runs*. First, we introduce the concept of runs. In fact, many commonly

used criteria for the statistical analysis involve the concept of runs, i.e., uninterrupted sequences of alike elements bordered at each end by other types of elements or by the beginning or by the end of the complete sequence [25]. The number of elements contained in a run is called the run length. For example, in a binary sequence 011001111000, there are five runs 0, 11, 00, 1111 and 000, and their run lengths are 1, 2, 2, 2, 4, respectively. In this paper, we expand the concept of runs into the two-dimensional link resource model. On the two-dimensional link resource model, a run is defined as a sequence of contiguous link resource cells with same status (free or occupied) in a horizontal (or vertical) direction, as shown in Fig. 3. Then we give a more formal definition of runs.

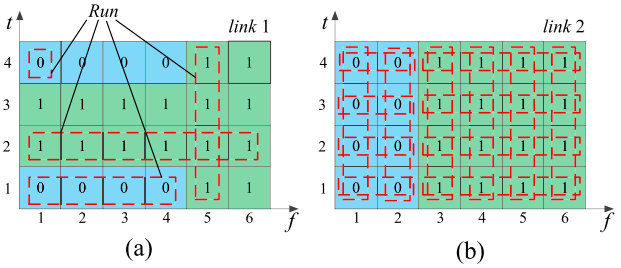


FIGURE 3. An example showing runs in the two-dimensional link resource model.

Definition (*Run*): Let $u^l_{t,f}$ denote the link resource cell in the t-th row and f-th column on the two-dimensional link resource matrix. On one hand, when t is fixed and the following formula is satisfied,

$$u^l_{t,f-1} \neq u^l_{t,f} = u^l_{t,f+1} = \cdots = u^l_{t,f+x-1} \neq u^l_{t,f+x}, f \geq 1, \tag{7}$$

the sequence $u^l_{t,f}, u^l_{t,f+1}, \cdots, u^l_{t,f+x-1}$ forms a run (in a horizontal direction) on the link resource matrix, and x denotes the run length. On the other hand, when f is fixed and the following formula is satisfied,

$$u^l_{t-1,f} \neq u^l_{t,f} = u^l_{t+1,f} = \cdots = u^l_{t+x-1,f} \neq u^l_{t+x,f}, t \geq 1, \tag{8}$$

the sequence $u^l_{t,f}, u^l_{t+1,f}, \cdots, u^l_{t+x-1,f}$ comprises a run (in a vertical direction) on the link resource matrix.

In the two-dimensional link resource model, when the number of free link resource cells is the same, less total number of runs means that the free time-frequency resources are more concentrated and contiguous, which is helpful for subsequent AR requests to find available rectangle blocks. For example, the resource models of *link* 1 and *link* 2 are shown in Fig. 3(a) and Fig. 3(b), respectively. In Fig. 3(b), each run is surrounded by a red dashed box, so by counting the number of red dashed boxes, we can get that the total number of runs of *link* 2 is 14. Similarly, we can obtain that the number of runs of *link* 1 is 20. Obviously, although the number of free link resource cells on *link* 1 and *link* 2 are both 16, the run number of *link* 2 is smaller, so the free resources on *link* 2 are more centralized and available.

Therefore, in order to improve the contiguity and availability of remaining free time-frequency resources, we always hope that less number of runs are introduced into each link resource model along the candidate paths. However, the process of reserving a candidate rectangle block may lead to an increase in the total number of runs on the selected path, which makes the free time-frequency resources become more and more fragmented. Fortunately, the increase in the number of runs caused by a candidate rectangle block is measurable, and we can judge how good a candidate rectangle block is by calculating the increased number of runs caused by the block.

Specifically, on each link of a candidate path, we introduce $x^l_{t,f}$ to measure the change in the number of runs on each row or column where a rectangle block is located. $x^l_{t,f}$ is calculated as follows.

$$x^l_{t,f} = \begin{cases} -2, & u^l_{t,f_s-1} + u^l_{t,f_e+1} = 0 \text{ or } u^l_{t_s-1,f} + u^l_{t_e+1,f} = 0 \\ 2, & u^l_{t,f_s-1} + u^l_{t,f_e+1} = 2 \text{ or } u^l_{t_s-1,f} + u^l_{t_e+1,f} = 2 \\ 0, & otherwise. \end{cases} \tag{9}$$

In Formula (9), $x^l_{t,f} = -2$ means that two runs are decreased in the two-dimensional link resource model, $x^l_{t,f} = 2$ means that two runs are increased in the two-dimensional link resource model, and $x^l_{t,f} = 0$ means that the number of runs does not change in the link resource model. Based on $x^l_{t,f}$, we traverse the time range $[t_s, t_e]$ and frequency range $[f_s, f_e]$ of the rectangle block along the candidate path $P^k_{s,d}$, then we will obtain the increased number of runs caused by the rectangle block. The increased number of runs can be calculated as follows.

$$R_n = \sum_{l \in P^k_{s,d}} \left(\sum_{t=t_s}^{t_e} x^l_{t,f} + \sum_{f=f_s}^{f_e} x^l_{t,f} \right). \tag{10}$$

When we evaluate each candidate rectangle block on a candidate path, a block with a smaller R_n means that a smaller number of runs are introduced during spectrum assignment process, which is beneficial to reducing the total number of runs on the selected path and maintaining the contiguity of remaining free time-frequency resources, thereby decreasing 2D fragmentation and improving the successful probability of resource assignment for the pending requests.

Figure 4 is an example to illustrate the calculation of the factor R_n to find the best rectangle block. For simplicity, we assume that there is only *link* 3 on the candidate path. The two-dimensional link resource model of *link* 3 is shown in Fig. 4(a). The occupied link resource cells are marked in blue and the free ones are marked in green. In this example, two candidate rectangle blocks (B_1 and B_2) are marked out in Fig. 4(a) and Fig. 4(b), respectively. For each block, we can get the increased number of runs by Formula (10). The calculation results are $R_n(B_1) = 0 - 2 = -2$ and $R_n(B_2) = 4 + 2 = 6$. Obviously, we prefer to assign block B_1

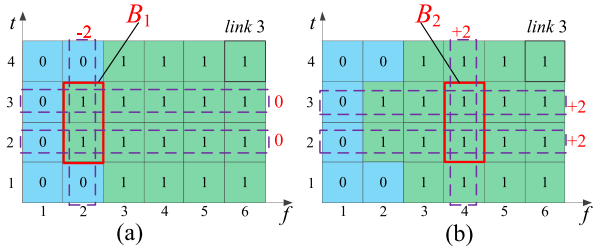


FIGURE 4. An example to show how to calculate the increased number of runs caused by candidate rectangle blocks.

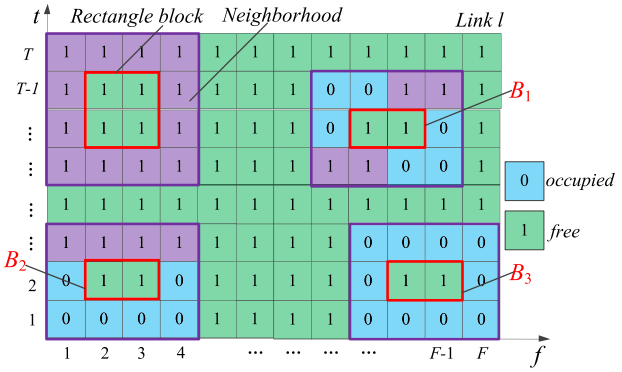


FIGURE 5. An example to describe the block neighborhood and the neighborhood matching degree of candidate rectangle blocks.

to the request because block B_1 corresponds to less increase in the number of runs. As a result, we reduce 2D fragmentation and maintain the contiguity of free time-frequency resources.

The second factor is *the neighborhood matching degree*. On the two-dimensional link resource model, each candidate rectangle block has an adjacent area that is composed of neighboring link resource cells. We refer to the area as the neighborhood of the rectangle block, as shown in Fig. 5. Note that, when we assign a rectangle block to an AR request, the contiguity of free resources in the block neighborhood will be damaged seriously. As a result, the free link resource cells in the block neighborhood may be unavailable for future requests due to insufficient contiguousness. In order to alleviate this negative impact, we propose two novel viewpoints: 1) the less free link resource cells adjacent to the rectangle block, the more helpful to maintain the contiguity and continuity of free time-frequency resources on the selected path; 2) the less number of free areas on the block neighborhood, the more contiguous and continuous of the occupied resources on the selected path. Based on these two points, we introduce the second factor——The *neighborhood matching degree*. The neighborhood matching degree refers to the matching degree between the rectangle block and its neighborhood, which can be calculated as follows.

$$C_n^l = O_n^l + B_n^l, \tag{11}$$

$$C_n = \sum_{l \in P_{s,d}^k} C_n^l. \tag{12}$$

In Formula (11), for a rectangle block on link l, O_n^l represents the number of neighboring free link resource cells, and B_n^l represents the number of free areas on the neighborhood. C_n^l denotes the neighborhood matching degree on link l. In (12), C_n is the sum of C_n^l on each link of the candidate path $P_{s,d}^k$. A smaller C_n^l means a higher neighborhood matching degree and less damage to the contiguity of remaining free link resource cells. This makes it more possible for future requests to obtain available time-frequency resources in the network. When C_n^l is equal to the minimum value (0), the rectangle block matches its neighborhood perfectly.

Figure 5 give an example to illustrate the calculation for the neighborhood matching degree of candidate rectangle blocks. For ease of description, we assume only link l on the candidate path. For an AR request, there are three candidate rectangle blocks on link l, namely B_1, B_2 and B_3. For block B_1, there are six free link resource cells in its neighborhood, which form two free areas. Thus we get that $O_n^l(B_1) = 4$, $B_n^l(B_1) = 2$ and $C_n(B_1) = C_n^l(B_1) = 4 + 2 = 6$. Similarly, we can get $C_n(B_2) = C_n^l(B_2) = 4 + 1 = 5$ and $C_n(B_3) = C_n^l(B_3) = 0+0 = 0$. Obviously, the block B_3 has the smallest C_n, thus its neighborhood matching degree is highest. Therefore, we prefer to select block B_3 to assign to the request.

On the other hand, when we assign a rectangle block to a request, we expect the block to be as close to the frequency boundary as possible. The frequency boundary refers to the position of $f = 1$ or $f = F$ on the frequency axis. For this purpose, we introduce the third factor——*The distance to the frequency boundary.* The distance between the rectangle block and the frequency boundary can be calculated as follows [18].

$$D_F = \min\{D_L, D_R\}. \tag{13}$$

In which, $D_L = f_s - 1$ and $D_R = F - f_e$. f_s and f_e are the start and end spectrum slots of the rectangle block, respectively. We prefer to choose a rectangle block with a small D_F, so that the occupied resources on the link can be concentrated near the frequency boundary. Conversely, the free time-frequency resources on each link are concentrated in the center area on the frequency axis, which helps alleviate spectrum continuity constraint.

To consider the above three factors (i.e., R_n, C_n and D_F) jointly, we put them into two metrics for evaluating the quality of an assignment scheme (i.e., a candidate rectangle block) comprehensively. The primary metric is calculated as follows.

$$W_1 = C_n + D_F, \tag{14}$$

and the secondary metric is calculated as follows.

$$W_2 = R_n. \tag{15}$$

The reason why W_1 is the primary metric is that W_1 has a more comprehensive investigation for each candidate rectangle block, while W_2 only can evaluate such rectangle blocks that cause the change of the number of runs. Specifically, to serve an AR request, we first try to find the rectangle

block(s) with minimum W_1. If there exist multiple rectangle blocks, we choose the one with the minimum W_2.

IV. THE PROPOSED MRC-2D-FA ALGORITHM

In this paper, in order to improve the spectral utilization and reduce the blocking probability of AR requests, we propose a dynamic RMSA algorithm named MRC-2D-FA. The proposed algorithm can be divided into two stages: routing selection and spectrum assignment.

A. ROUTING SELECTION POLICY

In this subsection, we propose a minimum resource consumption routing policy (MRC) to improve spectrum efficiency. The existing RSA algorithms mainly choose the k-shortest path (KSP) algorithm [26] as the routing strategy, and most of these algorithms take the hop number as the routing weight. The minimum hop routing focuses on mitigating spectrum continuity constraint. However, the routes with shorter hops may also correspond to longer path distances, which means that the request may require more free contiguous spectrum slots because it can only be configured with a lower modulation format. This case will aggravate the consumption of spectrum resources, and make the request face more stringent spectrum contiguity constraint. In addition, path distance is often used as the routing weight. The shortest distance routing focuses on alleviating the spectrum contiguity constraint. Nevertheless, the routes with shorter path distance may also correspond to longer hop number, which usually means that the rectangle blocks allocated for the request need to be aligned along more links on the candidate path. In other words, the request will face more serious spectrum continuity constraint due to the excessive number of hops in the candidate path. It can be seen that too large hop number or too long distance will cause excessive consumption of time-frequency resources.

In order to improve spectrum efficiency, our routing algorithm pays more attention to the resources consumed by AR requests, and combines the advantages of minimum hop routing and shortest distance routing algorithms. We notice that the less time-frequency resources the current request consumes, the more free resources are left for future requests. Therefore, we set the routing weight as the consumed time-frequency resources (i.e., total consumed link resource cells), which can be calculated as follows.

$$W_P = M \cdot N \cdot h_P, \tag{16}$$

where h_p is the hop number of path P, M and N represent the number of required time slots and required spectrum slots, respectively. Based on W_P, the process of MRC can be described as the following two phases.

The first phase is static, where all computation operations are performed in advance. In this phase, for each s-d pair in the network topology $G(V, E)$, we pre-calculate K shortest hop paths, and we also pre-calculate K shortest distance paths. Then, we put all calculated paths into the prepared routing set R_P. Note that there may be duplicate paths in the calculated paths, which are de-duplicated through the set R_P.

The second phase is dynamic and starts to execute when an AR request arrives to the network. When an AR request $AR(s, d, t_a, H, C, w)$ arrives, the routing algorithm will extract all routing paths from s to d in the routing set R_P. Then K candidate paths are selected from the extracted paths according to the ascending order of routing weight W_P. Finally, K candidate paths are put into the candidate routing set $P_{s,d}$.

B. SPECTRUM ASSIGNMENT POLICY

For a pending AR request $AR(s, d, t_a, H, C, w)$, when we get K candidate paths for the request from the routing set $P_{s,d}$, the proposed algorithm enters the spectrum assignment stage. This stage can also be divided into two phases: rectangle block search and rectangle block selection.

In the rectangle block search phase, for each candidate path, we obtain the link resource model of each link on the path, then we obtain the state matrix of the path resource model according to Formulas (5)-(6). After that, we calculate the required rectangle block size $M \times N$ for the request, and we find all holding time $[t_s, t_e]$ cases according to the sliding start time window w. Next, for each $[t_s, t_e]$ in the path resource model, we look for all such rectangle blocks, which must be adjacent to the occupied resources along the frequency axis. Then we repeat the above process on K candidate paths to obtain all available rectangle blocks for the request. If no rectangle block is found, we block the request. Otherwise, the algorithm enters the rectangle block selection phase.

In the rectangle block selection phase, we use the primary metric W_1 and secondary metric W_2 designed in section III to evaluate each candidate rectangle block. Specifically, we calculate the metrics W_1 and W_2 for each rectangle block by Formulas (14)-(15). If there is only one block with minimum W_1, we will assign the block and the corresponding path to the request. Otherwise, we will assign the block with minimum W_2 and the corresponding path to the request.

C. OVERALL ALGORITHM

Algorithm 1 shows the detailed procedures of the proposed dynamic RMSA algorithm MRC-2D-FA for provisioning AR requests in EONs.

Lines 1 to 5 present the process of pre-calculating routes for all s-d pairs during the network initialization phase, and Lines 6 to 21 illustrate the dynamic provisioning procedure for an AR request. To provision an AR request, Line 6 indicates that K candidate paths are selected from the pre-calculated path set R_P by Formula (16). Lines 7 to 21 describe the process of spectrum assignment. Lines 7 to 10 illustrate how to find all available rectangle blocks on K candidate paths. Then, in Lines 11 to 13, the algorithm calculates the primary metric W_1 and the second metric W_2 for each feasible block by Formulas (14)-(15). Lines 17 to 21 denote that if candidate rectangle blocks exists, the proposed algorithm will determine the optimal rectangle block

Algorithm 1 The proposed MRC-2D-FA algorithm

Input: An AR request $AR(s, d, t_a, H, C, w)$ and network topology $G(V, E)$;
Output: The best candidate rectangle block B and the routing path where the block is located;
Phase I: **Network Initialization**
1: **for** all s-d pairs in $G(V, E)$ **do**
2: calculate K shortest paths based on hop;
3: calculate K shortest paths based on path distance;
4: record the above paths to set R_p after removing the duplicate items;
5: **end for**
Phase II: **Dynamic Service Provision**
6: choose K candidate paths from R_p based on the routing selection policy and record them in set $P_{s,d}$;
7: **for** each $[t_s, t_e]$ **do**
8: **for** each path $P^k_{s,d}$ **do**
9: get the path resource matrix $U^P_{T,F}$ for path $P^k_{s,d}$ by Formula (4);
10: find all available rectangle blocks within $[t_s, t_e]$ on matrix $U^P_{T,F}$;
11: **for** each rectangle block B **do**
12: calculate W_1 with Formula (14);
13: calculate W_2 with Formula (15);
14: **end for**
15: **end for**
16: **end for**
17: **if** candidate rectangle blocks exist **then**
18: determine the best rectangle block $(P^k_{s,d}, f_s, f_e, t_s, t_e)$ based on the primary metric W_1 and secondary metric W_2, assign the block and the corresponding path to the request;
19: **else**
20: block the request;
21: **end if**

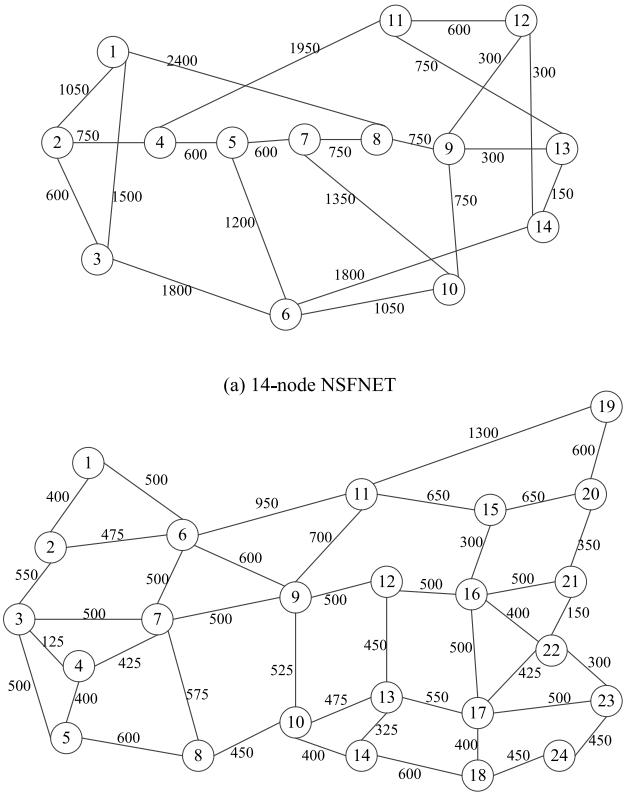
through the metrics W_1 and W_2, and assign the block and the corresponding path to the request. Otherwise, the request will be blocked.

D. COMPLEXITY ANALYSIS

The computational complexity of provisioning an AR request can be derived from the procedures of *Algorithm 1*. The time complexity of routing selection phase (Line 6) is $O(K)$, where K is the number of candidate paths. The worst time complexity of the rectangle block search phase (Lines 7 to 10) is $O(K \cdot |V| \cdot F \cdot |w| \cdot M)$, where $|V|$ is the total number of nodes, and the number of links on the candidate path is less than $|V|$, $|w|$ is the number of optional start times. M is the number of required time slots. In the rectangle block selection phase (Lines 11 to 21), the algorithm needs to calculate the metrics W_1 and W_2 for all candidate rectangle blocks. For each block, the time complexity of calculating W_1 is $O(2 \cdot (M + N) + 4)$ in Line 12. The time complexity of calculating W_2 is $O(M + N)$ in Line 13. Then the time complexity of rectangle block selection phase can be represented as $O(I \cdot |V| \cdot (M + N))$, where I is the total number of candidate rectangle blocks. Hence, the worst time complexity of Algorithm 1 is approximately $O(|V| \cdot (K \cdot F \cdot |w| \cdot M + I \cdot M + I \cdot N))$.

V. PERFORMANCE EVALUATION

In this section, we evaluate the MRC-2D-FA algorithm through simulation using the 14-node NSFNET and 24-node USFNET as shown in Fig. 6. Then we compare the MRC-2D-FA algorithm with three relevant benchmarks in term of bandwidth blocking probability (BBP) and spectrum utilization (SU).



(a) 14-node NSFNET

(b) 24-node USFNET

FIGURE 6. Network topology.

A. SIMULATION SETTINGS

We assume that the elastic optical network is configured in the C-band, in which the bandwidth of each fiber link is 4.475THz. Based on O-OFDM technology, the size of each spectrum slot is set as 12.5 GHz. Therefore, each fiber link can accommodate 358 spectrum slots. For time resources, we assume that the maximum number of look-ahead time slots that can be observed by the network operator is set as $T = 300$ time slots, and the duration of each time slot is 10 minutes. In order to improve the spectral efficiency, we adopt distance-adaptive modulation in the

proposed algorithm. The optional modulation formats include BPSK, QPSK, 8QAM and 16QAM. Their bits per symbol are 1, 2, 3 and 4, and the corresponding maximum transmission distances are 9600km, 4800km, 2400km and 1200km, respectively [3].

In the simulations, AR requests are generated with the Poisson traffic model, i.e., the requests arrive with an average rate of λ per time slot, and the duration of each request follows a negative exponential distribution with an average value of $1/\mu$. Then the AR traffic load can be quantified with λ/μ in Erlang. For each AR request, the *s-d* node pair is randomly chosen, the line rate is uniformly distributed over {40 Gb/s, 80 Gb/s, 100 Gb/s, 120 Gb/s, 150 Gb/s, 180 Gb/s, 200 Gb/s, 240 Gb/s, 300 Gb/s, 400 Gb/s}. Meanwhile, the earliest start time is randomly chosen from [1, 30] time slots, and the size of start time window is randomly chosen from [1, 20] time slots. The main simulation parameters are listed in Table 1.

TABLE 1. Main simulation parameters.

Parameters	Values
Network topology	NSFNET, USFNET
Number of spectrum slots on each link	358
Number of look-ahead time slots	300
Size of each spectrum slot	12.5 GHz
Size of each time slot	10 minutes
Transmission distance of BPSK (M_p=1)	9600 km
Transmission distance of QPSK (M_p=2)	4800 km
Transmission distance of 8QAM (M_p=3)	2400 km
Transmission distance of 16QAM (M_p=4)	1200 km
K, Number of candidate paths	5
Average duration of AR requests	10 time slots
Arrive rate of AR requests	Poisson distribution
Earliest start time of AR requests	[1,30] time slots
Size of flexible window for the start time	[1,20] time slots

All the simulation results are averaged over 10 runs for different random seeds with the 95% confidence interval. During each simulation, 10^5 AR requests are generated.

B. PERFORMANCE INDEXES

To better evaluate the proposed algorithm, we consider several performance indexes as follows.

1) Bandwidth blocking probability (BBP): Since the service provisioning of AR requests introduces time dimension into the EON, being similar to [14], we define the BBP of AR requests as follows.

$$P_B = \frac{\sum_{r \in R_B} C_r \cdot H_r}{\sum_{r \in R} C_r \cdot H_r}, \tag{17}$$

where R is the set of total requests and R_B is the set of the blocked requests. C_r and H_r represent the line rate and duration of request r, respectively. A lower BBP means that fewer requests are blocked in the network and more data transfer demands are scheduled successfully.

2) Spectrum Utilization (SU): Being similar to [16], spectrum utilization is defined as the ratio of total occupied link resource cells to total link resource cells during each simulation. A higher spectrum utilization generally indicates that less 2D fragmentation and time-frequency resources are used more effectively.

C. COMPARISON WITH BENCHMARKS

In order to evaluate the performance of the proposed MRC-2D-FA algorithm, we choose three well-performed benchmark algorithms for comparison, namely the First-Fit (FF) algorithm, the Min-RDDR algorithm [18] and the Proactive AR algorithm [19].

Figure 7 compares the performance of our proposal with other benchmarks in terms of BBP in the NSFNET topology. On the whole, the BBP of the four algorithms gradually becomes higher with the increase of traffic load. When the same traffic load is considered, MRC-2D-FA achieves the lowest BBP compared with the other three benchmarks. When the traffic load is 700 Erlang, the BBP performance of MRC-2D-FA algorithm will decrease 57% approximately compared with FF. The reason is that MRC-2D-FA can minimize the generation of 2D fragments to improve the access probability of the arriving requests. The Proactive AR algorithm considers reducing fragmentation along the frequency axis and time axis, while MRC-2D-FA not only considers this factor, but also focus on the fragmentation condition in the neighborhood of each candidate rectangle block. Therefore, MRC-2D-FA has a lower BBP than Proactive AR algorithm.

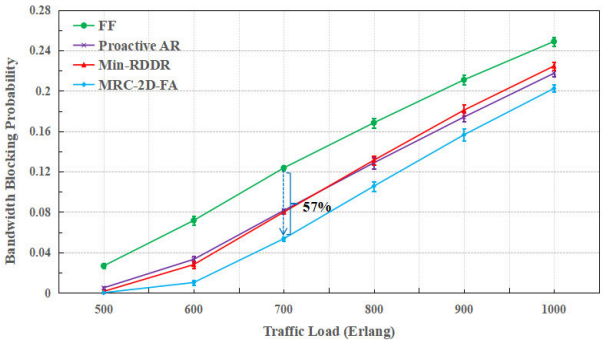


FIGURE 7. Bandwidth blocking probability in the NSFNET.

In addition, Min-RDDR attempts to schedule AR requests by using the rectangle blocks as close as possible to the time and frequency boundaries, which reserves more contiguous free resources near the central area for subsequent requests. However, a potential risk of Min-RDDR is that spectrum resources close to the time boundary may be overused, which may cause related AR requests to be blocked. Different from Min-RDDR, MRC-2D-FA considers planning AR requests closer to the frequency boundary without setting any constraints on the time dimension, which helps to balance the spectrum occupation on each time slot, thereby alleviating the request blocking caused by insufficient spectrum resources on some time slots. Hence, MRC-2D-FA outperforms the Min-RDDR algorithm in terms of BBP.

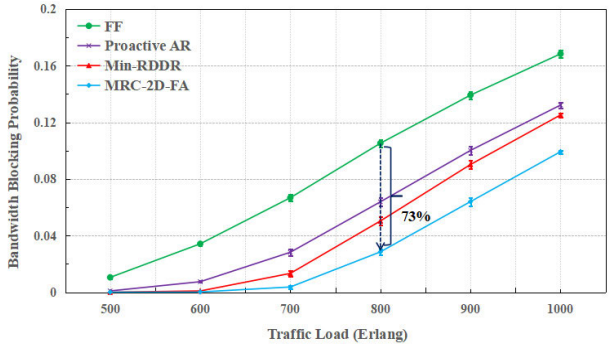


FIGURE 8. Bandwidth blocking probability in the USFNET.

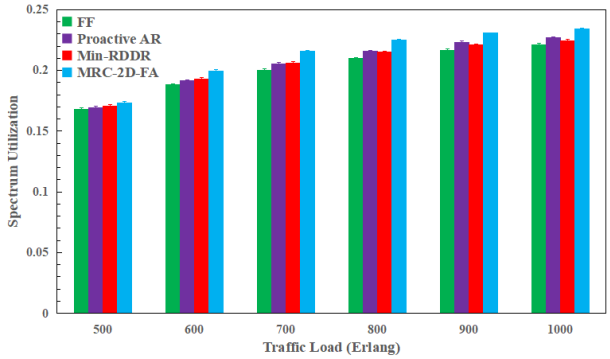


FIGURE 9. Spectrum utilization in the NSFNET.

In Figure 8, we use a larger topology USFNET to evaluate the performance of the proposed algorithm in terms of BBP. By comparing Fig. 7 and Fig. 8, we find that the BBP of each algorithm in the USFNET follows a similar trend with the results in the NSFNET. However, it is worth noting that the proposed algorithm can achieve higher performance improvement than other three algorithms in the USFNET topology. Specifically, the fragmentation problem of USFNET is more serious than that of NSFNET because the former has more links and nodes. This means that AR requests need to satisfy more stringent contiguity and continuity constraints in the USFNET. In this case, MRC-2D-FA can reduce fragmentation more effectively than other benchmarks, so the performance improvement is more obvious.

Figure 9 compares the performance of MRC-2D-FA with the benchmarks in terms of spectrum utilization in the NSFNET topology. It can be clearly seen that the spectrum utilization of these four algorithms gradually increases with the increase of traffic load. When considering the same traffic load, the proposed algorithm has the highest spectrum utilization compared with FF, Proactive AR and Min-RDDR. This is because the proposed algorithm reduces massive 2D fragments, so that free time-frequency resources of candidate paths become more contiguous and satisfy more AR requests. Therefore MRC-2D-FA has higher spectrum utilization compared with other algorithms in the NSFNET topology.

Figure 10 shows the spectrum utilization of the above four algorithms in the USFNET. Obviously, the performance

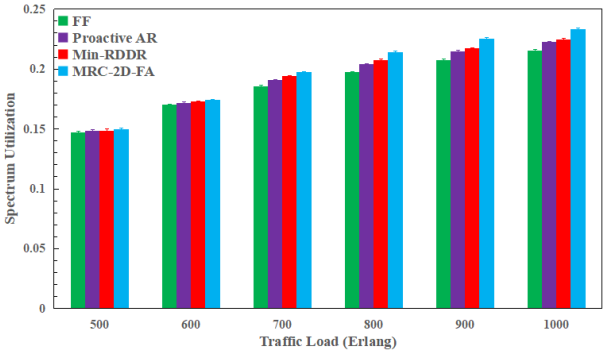


FIGURE 10. Spectrum utilization in the USFNET.

of the proposed algorithm is still optimal in terms of spectrum utilization. The results in Fig. 9 and Fig. 10 indicate that MRC-2D-FA algorithm can achieve higher spectrum utilization in different network topologies, which indirectly highlights that the proposed algorithm can further reduce 2D fragmentation and efficiently schedule time-frequency resources to serve more AR requests.

VI. CONCLUSION

In this article, we propose an RMSA algorithm called MRC-2D-FA, which aims to optimize the provision of AR requests in EONs by reducing resource consumption and decreasing 2D fragmentation. In the routing phase, by considering the information of path hops, path distance and resource consumption comprehensively, MRC-2D-FA can improve the spectrum efficiency, and balance the spectrum contiguity and continuity constraints with a low time complexity. In the spectrum assignment phase, we design several factors to thoroughly evaluate the pros and cons of an AR provision scheme. Based on these factors, MRC-2D-FA can further reduce 2D fragments and improve the contiguity and continuity of free resources on candidate paths, which helps subsequent requests to find available time-frequency resources. Furthermore, MRC-2D-FA can also balance the spectrum occupancy on each time slot, thereby alleviating traffic congestion caused by excessive load on some time slots. Simulation results show that compared with the three benchmarks, the proposed algorithm can further reduce the bandwidth blocking probability and improve the spectrum utilization. It is worth noting that the flexibility of the start time of STSD-fixed and STSD-flexible AR requests is different, leading to great unfairness in these two types of requests in terms of blocking probability. Therefore, in our future works, we look forward to designing some efficient methods to improve the fairness between STSD-fixed and STSD-flexible AR services in EONs.

REFERENCES

[1] C. A. Brackett, "Dense wavelength division multiplexing networks: Principles and applications," *IEEE J. Sel. Areas Commun.*, vol. 8, no. 6, pp. 948–964, Aug. 1990.

[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] M. Jinno, B. Kozicki, H. Takara, A. Watanabe, Y. Sone, T. Tanaka, and A. Hirano, "Distance-adaptive spectrum resource allocation in spectrum-sliced elastic optical path network," *IEEE Commun. Mag.*, vol. 48, no. 8, pp. 138–145, Aug. 2010.
[4] 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, pp. 5–39, Jan. 2017.
[5] 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.
[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] S. K. Singh and A. Jukan, "Efficient spectrum defragmentation with holding-time awareness in elastic optical networks," *J. Opt. Commun. Netw.*, vol. 9, no. 3, p. B78, 2017.
[8] Y. Yin, H. Zhang, M. Zhang, M. Xia, Z. Zhu, S. Dahlfort, and S. J. B. Yoo, "Spectral and spatial 2D fragmentation-aware routing and spectrum assignment algorithms in elastic optical networks [invited]," *J. Opt. Commun. Netw.*, vol. 5, no. 10, p. A100, Oct. 2013.
[9] D. Adhikari, D. Datta, and R. Datta, "Impact of BER in fragmentation-aware routing and spectrum assignment in elastic optical networks," *Comput. Netw.*, vol. 172, May 2020, Art. no. 107167.
[10] H. Liu, H. Hu, Y. Chen, and L. Du, "Fragmentation-avoiding spectrum assignment algorithm based on time-spectrum partition for elastic optical networks," *Opt. Fiber Technol.*, vol. 53, Dec. 2019, Art. no. 102020.
[11] S. Trindade and N. L. S. da Fonseca, "Proactive fragmentation-aware routing, modulation format, core, and spectrum allocation in EON-SDM," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.
[12] F. Pederzolli, D. Siracusa, A. Zanardi, G. Galimberti, D. La Fauci, and G. Martinelli, "Path-based fragmentation metric and RSA algorithms for elastic optical networks," *J. Opt. Commun. Netw.*, vol. 11, no. 3, pp. 15–25, 2019.
[13] N. Charbonneau and V. M. Vokkarane, "A survey of advance reservation routing and wavelength assignment in wavelength-routed WDM networks," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 4, pp. 1037–1064, 4th Quart., 2012.
[14] W. Lu and Z. Zhu, "Dynamic service provisioning of advance reservation requests in elastic optical networks," *J. Lightw. Technol.*, vol. 31, no. 10, pp. 1621–1627, May 2013.
[15] W. Lu, S. Ma, C. Chen, X. Chen, and Z. Zhu, "Implementation and demonstration of revenue-driven provisioning for advance reservation requests in OpenFlow-controlled SD-EONs," *IEEE Commun. Lett.*, vol. 18, no. 10, pp. 1727–1730, Oct. 2014.
[16] N. Wang, J. P. Jue, X. Wang, Q. Zhang, H. C. Cankaya, and M. Sekiya, "Holding-time-aware scheduling for immediate and advance reservation in elastic optical networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 5180–5185.
[17] H. Chen, Y. Zhao, J. Zhang, R. He, W. Wang, J. Wu, Y. Wang, Y. Ji, H. Zheng, Y. Lin, and B. Hou, "Time-spectrum consecutiveness based scheduling with advance reservation in elastic optical networks," *IEEE Commun. Lett.*, vol. 19, no. 1, pp. 70–73, Jan. 2015.
[18] X. Li, J. Yuan, Q. Zhang, Z. Ren, and L. Yang, "Farsighted spectrum resource assignment method for advance reservation requests in elastic optical networks," *IEEE Access*, vol. 7, pp. 167836–167846, 2019.
[19] W. Lu, Z. Zhu, and B. Mukherjee, "On hybrid IR and AR service provisioning in elastic optical networks," *J. Lightw. Technol.*, vol. 33, no. 22, pp. 4659–4670, Nov. 15, 2015.
[20] R. Zhu, Y. Zhao, H. Yang, X. Yu, J. Zhang, A. Yousefpour, N. Wang, and J. P. Jue, "Dynamic time and spectrum fragmentation-aware service provisioning in elastic optical networks with multi-path routing," *Opt. Fiber Technol.*, vol. 32, pp. 13–22, Dec. 2016.
[21] S. Sugihara, Y. Hirota, S. Fujii, H. Tode, and T. Watanabe, "Dynamic resource allocation for immediate and advance reservation in space-division-multiplexing-based elastic optical networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 9, no. 3, pp. 183–197, Mar. 2017.
[22] H. Chen, Y. Zhao, J. Zhang, W. Wang, and R. Zhu, "Static routing and spectrum assignment for deadline-driven bulk-data transfer in elastic optical networks," *IEEE Access*, vol. 5, pp. 13645–13653, 2017.
[23] W. Wang, Y. Zhao, H. Chen, J. Zhang, H. Zheng, Y. Lin, and Y. Lee, "Re-provisioning of advance reservation applications in elastic optical networks," *IEEE Access*, vol. 5, pp. 10959–10967, 2017.
[24] P. Afsharlar, A. Deylamsalehi, J. M. Plante, J. Zhao, and V. M. Vokkarane, "Routing and spectrum assignment with delayed allocation in elastic optical networks," *J. Opt. Commun. Netw.*, vol. 9, no. 3, p. B101, Mar. 2017.
[25] D. L. Antzoulakos, S. Bersimis, and M. V. Koutras, "On the distribution of the total number of run lengths," *Ann. Inst. Stat. Math.*, vol. 55, no. 4, pp. 865–884, Dec. 2003.
[26] J. Y. Yen, "Finding the K shortest loopless paths in a network," *Manage. Sci.*, vol. 17, pp. 712–716, Jul. 1971.

YAN LIU received the B.S. degree in electronic information engineering from Northeast Petroleum University, Daqing, China, in 2018. He is currently pursuing the M.S. degree in information and communication engineering with Dalian Maritime University, Dalian, China. His research interests include soft-defined networking and elastic optical networks.

RONGXI HE (Member, IEEE) received the B.S. and M.S. degrees from Dalian Maritime University, in 1992 and 1995, respectively, and the Ph.D. degree from the University of Electronic Science and Technology of China, in 2002.

From 2002 to 2004, he was a Postdoctoral Fellow with the Network and Communication Center, Northeastern University. From 2006 to 2007, he was a Visiting Scholar with the Broadband Communications Research (BBCR) Laboratory, University of Waterloo, ON, Canada. He is currently a Professor with Dalian Maritime University. He is the author of two books and more than 100 articles. His research interests include wireless networks and optical networks. He is a Senior Member of the China Computer Federation.

SHICHENG WANG received the B.S. degree in communication engineering from Dalian Maritime University, Dalian, China, in 2019, where he is currently pursuing the M.S. degree in information and communication engineering. His research interest includes elastic optical networks.

CUNQIAN YU received the Ph.D. degree in communication and information systems from Northeastern University, Shenyang, China, in 2017. He is currently an Associate Professor with the College of Computer Science Technology, Dalian Maritime University. His research interests include optical data center networks, industrial internet, and maritime broadband communication. He has authored more than 50 international journal and conference papers in the above fields.