IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 11, 2017, accepted January 25, 2018, date of publication January 30, 2018, date of current version March 9, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2799867

Distributed Sub-Tree-Based Optical Multicasting Scheme in Elastic Optical Data Center Networks

XIN LI, LU ZHANG, YING TANG, JUNFENG GUO, AND SHANGUO HUANG, (Member, IEEE)
State Key Laboratory of Information Photonic and Optical Communication, Beijing University of Posts and Telecommunications, Beijing 100876, China

Corresponding author: Xin Li (xinli@bupt.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61701039, Grant 61601054, and Grant 61571058, and in part by the National Science Foundation for Outstanding Youth Scholars of China under Grant 61622102.

ABSTRACT The sub-tree-based optical multicasting scheme provides a spectrum-efficient approach to providing emerging multicast services in optical data center networks. Moreover, multicast services are usually replicated and maintained in multiple geographically distributed data centers to improve its access efficiency and reliability. Therefore, the source data centers of all constructed sub-trees for a multicast demand are not confined to a common data center and can be independently determined by the requested distribution. In this paper, we study the problem of multicast service provisions while leveraging multicast service backups among multiple geographically distributed data centers. A novel distributed sub-tree-based optical multicasting (DST-OM) scheme is proposed. An integer linear program model is developed for the DST-OM scheme with the aim of minimizing the total spectrum consumption of all multicast demands in elastic optical data center networks. We also define the minimum spectrum sub-tree (MSST) problem for the DST-OM scheme. Two modulation-level-aware heuristic algorithms are developed to address the MSST problem. Numerical results show that the DST-OM scheme achieves higher spectrum efficiency and lower blocking probability than the conventional common source sub-tree-based optical multicasting scheme and the single-tree-based optical multicasting scheme.

INDEX TERMS Elastic optical data center networks (EO-DCNs), multicast service, optical multicasting, distributed sub-tree (DST).

I. INTRODUCTION

Recently, data centers have gained high speed development. The traffic among data centers is booming, and more services are emerging, such as cloud computing, big data, augmented reality, and virtual reality [1]. To cope with the booming traffic among data centers, the optical interconnection has been widely applied to data center networking [2]. It has advantages of large capacity, high speed and low energy. Moreover, an elastic optical data center network (EO-DCN) that supports the adaptive and fine-grained spectrum allocation is a promising candidate for the next generation of optical data center networks [3], [4]. Moreover, multicast services, such as data backup, scientific computing, ultra-high-definition TV delivery, and others, are gaining popularity and momentum [5]. It requires the same data to be transmitted from a source data center to a set of destination users. When serving multicast demands in optical data center networks, the light-tree-based optical multicasting scheme is more spectrum-efficient and power-efficient than the IP multicasting scheme [6]. In [7], a light-tree-based optical data center network was proposed to optimize network resource usage while serving unicast and multicast services. However, the coverage area of a light-tree is limited due to signal attenuation, light-splitting, amplified spontaneous emission (ASE), and nonlinear impairments. When a large number of geographically scattered users request a common multicast service, it is impractical to construct a single light-tree that can cover all requesters since the length of its longest branch may exceed the transmission reach of an optical signal. Even if a single large-size light-tree can be constructed, it will consume too many spectrum resources since the low modulation level usually needs to be adopted. To address this problem, a sub-tree-based optical multicasting scheme is proposed to provide a spectrum-efficient approach to serve multicast demands with a large number of requesters [8]–[13]. It uses multiple small-size sub-trees to jointly serve one multicast demand and conducts the distance-adaptive routing and spectrum allocation for each sub-tree.

In recent years, natural disasters such as earthquakes, hurricanes, tsunamis, and human-made intentional attacks have been frequent. These can cause severe service disruptions that affect the survivability of optical data center networks and are therefore attracting more attention. Multicast services are usually replicated and maintained in multiple geographically distributed data centers to improve its access efficiency and reliability. Yao *et al.* [14] studied the fast and coordinated data backup for cloud services among multiple geographically distributed data centers. They tried to minimize the time duration with the joint optimization of backup site selection and data-transfer paths. Boru *et al.* [15] designed multiple models for the energy consumption and bandwidth demand of database access in cloud computing data centers. Based on these models, they proposed an energy-efficient replication strategy that can improve the quality of service (QoS) of cloud computing with reduced communication delays. A new concept of content connectivity that was defined as the reachability of content from any point of a network was proposed in [16]. It does not merely ensure the connectivity between source and destination nodes but guarantees the connectivity between users and the required content. A disaster-aware dynamic content placement algorithm was developed to minimize the number of replications of content while satisfying the content connectivity requirement [17]. Furthermore, the concept of k-node (edge) content connectivity was proposed to quantitatively measure content connectivity and provide protection for different types of content [18]. When a multicast service is hosted in multiple geographically distributed data centers, a requester can access any reachable data center to obtain the desired multicast service. For the sub-tree-based optical multicasting scheme, the source of a sub-tree is no longer confined to one particular data center, and any data center that hosts the required multicast service can be designated the source data center. In other words, it is possible to construct multiple sub-trees with different source data centers to jointly serve one multicast demand.

By leveraging multicast service backups among multiple geographically distributed data centers, we propose a novel distributed sub-tree-based optical multicasting (DST-OM) scheme in EO-DCNs. The DST-OM scheme uses multiple distributed sub-trees to jointly serve one multicast demand and conducts distance-adaptive routing and spectrum allocation for each sub-tree. An integer linear program (ILP) model is developed for the DST-OM scheme with the aim of minimizing the total spectrum consumption of all multicast demands. Moreover, we define the minimum-spectrum sub-tree (MSST) problem that requires the use of the minimum spectrum resource to serve a multicast demand. Two heuristic algorithms that explore the modulation-level-aware strategy are developed to address the MSST problem. The rest of this paper is organized as follows. Section II discusses related works and our contributions. Section III elaborates the DST-OM scheme. The ILP formulation is developed in Section IV. We define the MSST problem and develop two heuristic algorithms in Section V. Numerical results are presented and analyzed in Section VI. Finally, Section VII concludes this paper.

II. RELATED WORKS AND OUR CONTRIBUTIONS

In this section, we summarize the current status and research progress in the light-tree-based optical multicasting scheme and the conventional common source sub-tree-based optical multicasting (CSST-OM) scheme in wavelength switched optical networks, elastic optical networks, and optical data center networks. Moreover, the contributions of this paper are summarized and elaborated.

A. THE LIGHT-TREE-BASED OPTICAL MULTICASTING SCHEME

The light-tree-based optical multicasting scheme has been extensively researched in both wavelength-switched optical networks [19]–[24] and elastic optical networks [25]–[29]. Lee *et al.* [19] compared the performance of the light-path scheme with that of the light-tree scheme in minimizing network blocking probability for dynamic multicast traffic grooming in WDM networks. The light-tree scheme presented good performance when IP demands became high. In [20], a dynamic light-tree-based multicast grooming algorithm was proposed to address the online multicast traffic grooming problem in WDM networks. In [21], a cost-effective multicast-capable optical node with the tap-and-continue and tap-and-binary-split functionalities was developed in WDM networks. Andrei *et al.* [22] studied the problem of provisioning dynamic multicast data-distribution requests with flexible scheduling over optical WDM networks. Zhu *et al.* [23] addressed the 3R regenerator placement problem along light-trees in WDM networks. Zhou *et al.* [24] derived two approximate optimal multicast light-tree computation algorithms (the reroute-to-source algorithm and the member-only algorithm) for both unweighted and non-equally weighted WDM networks. Wang and Chen [25] developed a heuristic shortest path tree-based first-fit (SPT-FF) algorithm and a heuristic minimal spanning tree-based first-fit (MST-FF) algorithm to conduct light-tree construction and spectrum allocation in elastic optical networks. Gong *et al.* [26] studied the multicast-capable routing, modulation level, spectrum allocation (RMLSA) problem for static and dynamic multicast services in elastic optical networks. Walkowiak *et al.* [27] proposed two ILP formulations and a heuristic method to optimize multicast traffic with the distance-adaptive transmission. Li *et al.* [28] addressed the service scheduling and the RMLSA problem of the advance reserved multicast requests in software-defined elastic optical networks. Cai *et al.* [29] conducted the distance-adaptive spectrum resource allocation with shared protection for multicast sessions in elastic optical networks. Each link of a light-tree is protected by a backup path that is link-disjointed to the path from the source to each destination on the primary tree.

Since increasing numbers of multicast services are replicated and maintained in data centers, the light-tree-based optical multicasting scheme has a more meaningful application and has led to extensive studies in optical data center networks [30]–[33]. Gifre *et al.* [30] demonstrated an orchestrated inter-data center multicast connectivity for Ethernet services. An ABNO-driven workflow was experimentally validated to provide multicasting connectivity over a multilayer Ethernet over flex-grid network. Tokas *et al.* [31] demonstrated a scalable optical data center architecture with a multicasting capability. In [32], an efficient routing and forwarding algorithm was proposed for multicast flows in large-scale optical data center networks. This algorithm could achieve remarkable performance in terms of memory consumption, processing time, hardware costs, and delivery accuracy. Zeng *et al.* [33] used the network function virtualization (NFV) to orchestrate the multicast-oriented NFV trees in inter-DC elastic optical networks. Thus, it can be seen that the light-tree-based optical multicasting scheme provides a promising method to serve multicast services both in wavelength-switched optical networks and elastic optical networks.

B. THE CSST-OM SCHEME

To improve the spectrum efficiency of optical data center networks while serving multicast demands, the CSST-OM scheme that uses multiple small-size sub-trees with a common data center to jointly serve one multicast demand is proposed. It is more flexible and spectrum-efficient than the single-tree-based optical multicasting (ST-OM) scheme. The existing studies involve the light-forest scheme [8]–[10] and the routing, modulation level, and spectrum allocation (RMLSA) problem [11]–[13]. In [8], a light-forest scheme was proposed to support group multicast sessions in mesh-based optical grid networks. The minimum cost heuristic light-tree construction method and the linear light-tree construction method were proposed to construct distributed concurrent light-trees for a group multicast request in wavelength-switched optical networks. In [9], a light-forest with rateless network coding scheme was proposed to design efficient all-optical multicast schemes for elastic optical networks. The authors incorporated the rateless network coding into the sub-tree construction. Zhu *et al.* [10] established the minimum light-forest problem for optimizing a light-forest for all-optical multicasting in EO-DCNs. They proposed several time-efficient heuristic algorithms to address this problem. To address the RMLSA problem of sub-trees in flex-grid optical networks, the ILP model and heuristic algorithms were developed in [11] and [12]. Ruiz and Velasco [13] proposed a dynamic RMLSA algorithm for all-optical multicasting using sub-trees. Each sub-tree was constructed based on their sharing ratio on all constructed sub-trees. Zhu, and Jue [34] proposed a light-tree-based multi-class multicast flow aggregation scheme in IP over WDM networks. This paper focused on multi-class multicast flow aggregation and each sub-tree was constructed according to the classes of multicast services.

C. OUR CONTRIBUTIONS

In this paper, we focus on the DST-OM scheme in EO-DCNs. It uses multiple distributed small-size sub-trees to jointly serve one multicast demand and conducts the distance-adaptive routing and spectrum allocation for each sub-tree. Multicast service backup will require extra spectrum resource to realize data synchronization among multiple data centers. The required extra spectrum resource for service backup can be allocated when the traffic flow is at low density. This paper focuses on sub-tree construction for multicast demands and the synchronization technology is out the scope of this paper. Our contributions consist of three aspects.

1) To serve multicast demands with a large number of geographically scattered requesters, a novel DST-OM scheme is proposed by using multiple distributed sub-trees.
2) An ILP model is developed for the DST-OM scheme with the objective of minimizing the total spectrum consumption of all multicast demands. It jointly conducts sub-tree construction, modulation-level selection, and spectrum allocation.
3) The MSST problem is defined for the DST-OM scheme. Two heuristic algorithms are developed to address the MSST problem. It conducts the modulation-level-aware requester grouping and the distance-adaptive routing, modulation level, and spectrum allocation for each multicast demand.

III. PROBLEM FORMULATION

In this section, we first describe the constraints of light-tree construction in EO-DCNs. Then, the DST-OM scheme is elaborated upon and contrasted to the CSST-OM scheme and the ST-OM scheme in a small-scale EO-DCN.

A. THE LIGHT-TREE CONSTRUCTION

In EO-DCNs, the whole spectrum in each fiber link is divided into a set of continuous frequency slots (FSs) with a constant small spectrum step. The transponder at each optical node is assumed to be tunable so that any light-tree can use different sets of contiguous FSs. The spectrum allocated on each link of a light-tree must satisfy spectrum contiguity and spectrum continuity requirements that ensure that every link uses the same set of contiguous FSs [35]. The coverage area of a light-tree is limited due to signal attenuation, light-splitting, ASE, nonlinear impairments, and other factors. Sambo *et al.* [36] demonstrated the light-tree-based optical multicasting scheme through the broadcast and select node architecture. The results showed that the length of the longest branch of a light-tree is close to the reach of the transmission of an optical signal along a point-to-point light-path with the help of amplifier and equalizer. Yang *et al.* [9] provided Eq. (1) to describe the relationship among the maximum transmission reach of the adopted modulation level, the transmission distance of the longest branch of a light-tree, and the number of its destination nodes. In this paper, by considering a multicast demand with a large number of geographically

TABLE 1. Transmission distance of the longest branch of a light-tree.

$S_{m,n}$		n			
		1	2	3	4
m	1: BPSK	5000km	3842 km	3385 km	3120 km
	2: QPSK	2500 km	1920 km	1692 km	1560 km
	3: 8QAM	1250 km	870 km	845 km	802 km

scattered requesters and each light-tree is only implemented by light-splitters, we adopt Eq. (1) to determine the available modulation level of a sub-tree.

$$S_{m,n} = \frac{d_m}{\log_{10}(n) + 1} \tag{1}$$

In Eq. (1), $S_{m,n}$ denotes the transmission reach of the longest branch of a light-tree that adopts modulation level m and has n destination nodes, and d_m denotes the maximum transmission reach of modulation level m. Table I lists the value of $S_{m,n}$ when $m \in \{1, 2, 3\}$ and $n \in \{1, 2, 3, 4\}$. Typically, the transmission reaches of modulation levels BPSK, QPSK, and 8QAM are 5000 km, 2500 km, and 1250 km, respectively [11]. The transmission rates of an FS with modulation levels BPSK, QPSK, and 8QAM are 12.5 Gbps, 25 Gbps, and 37.5 Gbps respectively.

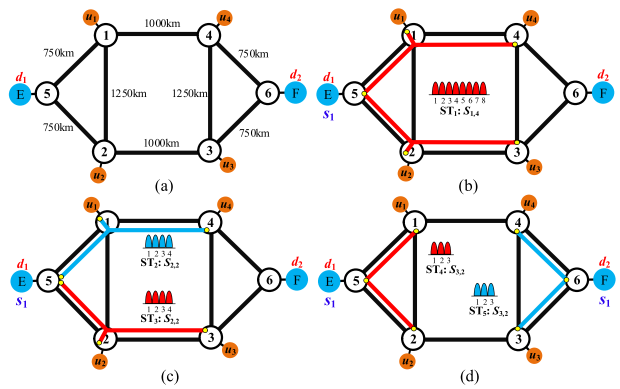


FIGURE 1. The light-tree-based optical multicasting scheme (a) n6e8, (b) the ST-OM scheme, (c) the CSST-OM scheme, and (d) the DST-OM scheme.

B. THE DST-OM SCHEME

We assume that each elastic optical node is capable of all-optical multicasting in EO-DCNs. Fig. 1 presents the spectrum consumption of the DST-OM scheme, the CSST-OM scheme, and the ST-OM scheme. In Fig. 1 (a), users u_1, u_2, u_3, and u_4 all request multicast service s_1, and the transmission rate of multicast service s_1 is assumed to be 100 Gbps. In Fig. 1 (b), multicast service s_1 is only hosted in data center d_1. For the ST-OM scheme, a single light-tree ST_1 needs to be constructed. Since the transmission distance of the longest branch of light-tree ST_1 is 1750 km, which is less than $S_{1,4}$ and greater than $S_{2,4}$, light-tree ST_1 must adopt the modulation level BPSK. Then, the number of the required FSs of multicast service s_1 is $\lceil 100/12.5 \rceil = 8$ FSs. Therefore, the total spectrum consumption of light-tree ST_1 is $4*8 = 32$ FSs. In Fig. 1 (c), sub-tree ST_2 and sub-tree ST_3 need to be constructed to minimize the spectrum consumption of the CSST-OM scheme. Since the transmission distance of the longest branch of sub-tree ST_2 or sub-tree ST_3 is also 1750 km (which is less than $S_{2,2}$ and greater than $S_{3,2}$), sub-tree ST_2 and sub-tree ST_3 can both adopt the modulation level QPSK. Then, the number of the required FSs of multicast service s_1 is $\lceil 100/25 \rceil = 4$ FSs. The total spectrum consumption of sub-tree ST_2 and sub-tree ST_3 is $2*4+2*4 = 16$ FSs.

In Fig. 1(d), multicast service s_1 is hosted in data center d_1 and data center d_2. For the DST-OM scheme, we construct sub-tree ST_4 and sub-tree ST_5, of which the source data centers are data center d_1 and data center d_2, respectively. Since the transmission distance of the longest branch of sub-tree ST_4 or sub-tree ST_5 is 750 km (which is less than $S_{3,2}$), sub-tree ST_4 and sub-tree ST_5 can both adopt the modulation level 8QAM. Then, the number of the required FSs of multicast service s_1 is $\lceil 100/37.5 \rceil = 3$ FSs. The total spectrum consumption of sub-tree ST_4 and sub-tree ST_5 is $2*3+2*3 = 12$ FSs. These results show that the DST-OM scheme has minimal spectrum consumption. Since the DST-OM scheme enables each sub-tree to adopt a high modulation level, the spectrum efficiency of the DST-OM scheme is higher than that of the ST-OM scheme and the CSST-OM scheme.

IV. THE ILP MODEL

In this section, the ILP models of the DST-OM scheme, the ST-OM scheme, and the CSST-OM scheme are developed.

A. THE ILP MODEL OF THE DST-OM SCHEME

In the developed ILP model, sub-tree construction, multicast service provisioning, modulation-level selection, and spectrum allocation are jointly considered.

1) NOTATION

$G(V, D, U, E)$: EO-DCN, where V denotes the set of optical nodes, D denotes the set of data centers, U denotes the set of users, and E denotes the set of fiber links.

$l_{(i,j)}$: the length of $link(i, j) \in E$ in kilometers.

S: the set of multicast services, $s \in S$.

b_s: the transmission rate of multicast service s.

Υ_u^s: equals one when user u requests multicast service s.

B: the base capacity of an FS with modulation level BPSK.

T: the set of constructed sub-trees, $k \in T$.

$\Lambda = \{f_1, f_2, \ldots, f_{|\Lambda|}\}$: the ordered set of FSs in each fiber link.

$M = \{1, 2, 3\}$: the set of optional modulation levels for each sub-tree, $m \in M$. Moreover, $m = 1$, 2, and 3 represents modulation levels BPSK, QPSK, and 8QAM, respectively.

$S_{m,n}$: the length of the longest branch of a sub-tree that has n destination nodes and adopts modulation level m.

Δ: a very large positive integer.

2) VARIABLES

$P^k_{(i,j)} \in \{0, 1\}$: equals one if $link(i, j)$ is used by sub-tree k, otherwise 0.

$X^k_{(i,j),f} \in \{0, 1\}$: equals one if FS f in $link(i, j)$ is used by sub-tree k, otherwise 0.

$R^k_{d,u} \in \{0, 1\}$: equals one if sub-tree k originates at data center d and terminates at user u, otherwise 0.

$f^k_{(i,j)}$: integer commodity-flow variable on the $link(i, j)$ of sub-tree k. Each user of a sub-tree needs one unit of the commodity.

$A^{s,k}_{d,u} \in \{0, 1\}$: equals one if multicast service s is transmitted through sub-tree k which originates at data center d and terminates at user u, otherwise 0.

$B^k_s \in \{0, 1\}$: equals one if multicast service s is transmitted through sub-tree k, otherwise 0.

L^k_i: integer variable that indicates the length of the longest branch that roots at node i in sub-tree k.

$n^k_i \in \{0, 1\}$: equals one if node i is occupied by sub-tree k, otherwise 0.

$Z^k_m \in \{0, 1\}$: equals one if sub-tree k adopts modulation level m, otherwise 0.

$Y^k_{s,m} \in \{0, 1\}$: equals one if multicast service s is transmitted through sub-tree k, which adopts modulation level m, otherwise 0.

$\gamma^{m,n}_k \in \{0, 1\}$: equals one if sub-tree k satisfies the specific m and n according to Eq. (1), where m is the adopted modulation level and n is the number of destination nodes, otherwise 0.

$H^s_d \in \{0, 1\}$: equals one when data center d hosts multicast service s, otherwise 0.

n_k: integer variable which denotes the number of the destination users of sub-tree k.

l_k: integer variable which denotes the length of the longest branch of sub-tree k.

3) OBJECTIVE FUNCTION

$$\min(\sum_{k\in T}\sum_{(i,j)\in E}\sum_{f\in\Lambda} X^k_{(i,j),f}) \tag{2}$$

With the objective in Eq. (2), we aim to minimize the total spectrum consumption of all constructed sub-trees.

4) SUB-TREE CONSTRUCTION CONSTRAINTS

$$\sum_{j:(i,j)\in E} f^k_{(i,j)} - \sum_{j:(j,i)\in E} f^k_{(j,i)} \begin{cases} \geq n^k_i & i \in D \\ = 0 & i \in V \\ = -n^k_i & i \in U \end{cases}, \quad \forall k \in T \tag{3}$$

$$P^k_{(i,j)} \geq f^k_{(i,j)}/\Delta, \quad \forall (i,j) \in E,\ \forall k \in T \tag{4}$$

$$P^k_{(i,j)} \leq f^k_{(i,j)}, \quad \forall (i,j) \in E,\ \forall k \in T \tag{5}$$

$$\sum_{d\in D} n^k_d \leq 1, \quad \forall k \in T \tag{6}$$

$$R^k_{d,u} \leq n^k_u, \quad \forall u \in U,\ \forall d \in D,\ \forall k \in T \tag{7}$$

$$R^k_{d,u} \leq n^k_d, \quad \forall u \in U,\ \forall d \in D,\ \forall k \in T \tag{8}$$

$$R^k_{d,u} \geq (n^k_u + n^k_d - 1), \quad \forall u \in U,\ \forall d \in D,\ \forall k \in T \tag{9}$$

Equation (3) ensures flow conservation at every intermediate node. Moreover, each user has no outgoing flows, and the source data center has no incoming flows. Equations (4)-(5) ensure that a link is occupied by a sub-tree only if there is a positive flow on this link. Equation (6) ensures that a sub-tree only has one source data center. Equations (7)-(9) ensure that sub-tree k originates at data center d and terminates at user u as long as data center d and user u both belong to sub-tree k.

5) MULTICAST SERVICE PROVISIONING CONSTRAINTS

$$\sum_{k\in T}\sum_{d\in D} A^{s,k}_{d,u} \geq \Upsilon^s_u, \quad \forall u \in U,\ \forall s \in S \tag{10}$$

$$\sum_{k\in T}\sum_{d\in D} A^{s,k}_{d,u} \leq \Delta * \Upsilon^s_u, \quad \forall u \in U,\ \forall s \in S \tag{11}$$

$$\frac{\sum_{k\in T}\sum_{u\in U} A^{s,k}_{d,u}}{\Delta} \leq H^s_d, \quad \forall d \in D,\ \forall s \in S \tag{12}$$

$$H^s_d \leq \sum_{k\in T}\sum_{u\in U} A^{s,k}_{d,u}, \quad \forall d \in D,\ \forall s \in S \tag{13}$$

$$\sum_{s\in S} B^k_s \leq 1, \quad \forall k \in T \tag{14}$$

$$A^{s,k}_{d,u} \leq B^k_s, \quad \forall u \in U,\ \forall d \in D,\ \forall s \in S,\ \forall k \in T \tag{15}$$

$$A^{s,k}_{d,u} \leq R^k_{d,u}, \quad \forall u \in U,\ \forall d \in D,\ \forall s \in S,\ \forall k \in T \tag{16}$$

$$A^{s,k}_{d,u} \geq (R^k_{d,u} + B^k_s - 1), \quad \forall u \in U,\ \forall d \in D,\ \forall s \in S,\ \forall k \in T \tag{17}$$

Equations (10)-(11) ensure that every user can only obtain the desired multicast service. Equations (12)-(13) determine the data center that should host the required multicast service. Equation (14) ensures that a sub-tree can be used to transmit only one multicast service at most. Equations (15)-(17) ensure that user u obtains multicast service s from sub-tree k that originates at data center d and terminates at user u if multicast service s is transmitted through sub-tree k and sub-tree k originates at data center d and terminates at user u.

6) LONGEST BRANCH CALCULATION CONSTRAINTS

$$L^k_i \leq \Delta * n^k_i, \quad \forall i \in (D \cup V \cup U),\ \forall k \in T \tag{18}$$

$$(L^k_i - L^k_j) \geq (l_{(i,j)} - \Delta * (1 - P^k_{(i,j)})), \quad \forall (i,j) \in E,\ \forall k \in T \tag{19}$$

$$l_k \geq \sum_{d\in D} L^k_d, \quad \forall k \in T \tag{20}$$

Equation (18) ensures that the branch can only be rooted at the node that is located on sub-tree k. Equation (19) determines the length of the branch that is rooted at node i on the sub-tree k. Equation (20) calculates the length of the longest branch of sub-tree k.

7) MODULATION-LEVEL SELECTION CONSTRAINTS

$$n_k = \sum_{u \in U} n_u^k, \quad \forall k \in T \tag{21}$$

$$\sum_{m \in M} \sum_{n \leq |U|} \gamma_k^{m,n} \leq 1, \quad \forall k \in T \tag{22}$$

$$\sum_{m \in M} \sum_{n \leq |U|} n * \gamma_k^{m,n} = n_k, \quad \forall k \in T \tag{23}$$

$$\sum_{m \in M} \sum_{n \leq |U|} S_{m,n} * \gamma_k^{m,n} \geq l_k, \quad \forall k \in T \tag{24}$$

$$\sum_{m \in M} Z_m^k \leq 1, \quad \forall k \in T \tag{25}$$

$$Z_m^k = \sum_{n \leq |U|} \gamma_k^{m,n}, \quad \forall m \in M, \forall k \in T \tag{26}$$

$$Y_{s,m}^k \leq B_s^k, \quad \forall m \in M, \forall s \in S, \forall k \in T \tag{27}$$

$$Y_{s,m}^k \leq Z_m^k, \quad \forall m \in M, \forall s \in S, \forall k \in T \tag{28}$$

$$Y_{s,m}^k \geq (B_s^k + Z_m^k - 1), \quad \forall m \in M, \forall s \in S, \quad \forall k \in T \tag{29}$$

Equation (21) calculates the number of users that sub-tree k covers. Equations (22)-(24) set the value of $\gamma_k^{m,n}$. Equation (22) ensures that a sub-tree with the specific n can at most adopts a kind of modulation level. Equation (23) determines the number of destination users for $\gamma_k^{m,n}$. Equation (24) ensures that the transmission distance of the adopted modulation level which satisfies the specific m and n according to Eq. (1), exceeds the length of the longest branch of a light-tree. Equation (25) ensures that each sub-tree can only adopt one kind of modulation level. Equation (26) determines the modulation level adopted by sub-tree k. Equations (27)-(29) set the value of $Y_{s,m}^k$. $Y_{s,m}^k$ is set to one if multicast service s is transmitted through sub-tree k and sub-tree k adopts modulation level m. Otherwise, it is set to zero.

8) SPECTRUM ALLOCATION CONSTRAINTS

$$\sum_{f \in \Lambda} X_{(i,j),f}^k \leq \sum_{s \in S} \sum_{m \in M} (Y_{s,m}^k * (\left\lceil \frac{b_s}{m * B} \right\rceil + 1)), \quad \forall (i,j) \in E, \quad \forall k \in T \tag{30}$$

$$\sum_{f \in \Lambda} X_{(i,j),f}^k \leq \Delta * P_{(i,j)}^k, \quad \forall (i,j) \in E, \forall k \in T \tag{31}$$

$$\sum_{f \in \Lambda} X_{(i,j),f}^k \geq [\Delta * (P_{(i,j)}^k - 1) + \sum_{s \in S} \sum_{m \in M} (Y_{s,m}^k * (\left\lceil \frac{b_s}{m * B} \right\rceil + 1))], \quad \forall (i,j) \in E, \quad \forall k \in T \tag{32}$$

Equation (30) calculates the number of FSs assigned to sub-tree k. Moreover, an extra free FS is allocated for each multicast service as a guard band. Equation (31) ensures that the spectrum can only be assigned to the links of a sub-tree. Equation (32) determines the number of FSs that are allocated on each link of a sub-tree.

9) SPECTRUM CONTIGUITY AND CONTINUITY CONSTRAINTS

$$\sum_{j:(j,i) \in E} f_{(j,i)}^k * X_{(j,i),f}^k = \sum_{j:(i,j) \in E} f_{(i,j)}^k * X_{(i,j),f}^k, \quad \forall i \in V, \quad \forall f \in \Lambda, \quad \forall k \in T \tag{33}$$

$$\sum_{j:(j,i) \in E} G_{(j,i),f}^k = \sum_{j:(i,j) \in E} G_{(i,j),f}^k, \quad \forall i \in V, \forall f \in \Lambda, \quad \forall k \in T \tag{34}$$

$$G_{(i,j),f}^k \leq X_{(i,j),f}^k * \Delta, \quad \forall (i,j) \in E, \forall f \in \Lambda, \quad \forall k \in T \tag{35}$$

$$G_{(i,j),f}^k \leq f_{(i,j)}^k, \quad \forall (i,j) \in E, \forall f \in \Lambda, \forall k \in T \tag{36}$$

$$G_{(i,j),f}^k \geq (X_{(i,j),f}^k - 1) * \Delta + f_{(i,j)}^k, \quad \forall (i,j) \in E, \quad \forall f \in \Lambda, \forall k \in T \tag{37}$$

$$(1 - X_{(i,j),f}^k + X_{(i,j),(f+1)}^k) \times \Delta \geq \sum_{f' \in [f+2, |\Lambda|]} X_{(i,j),f'}^k, \quad \forall (i,j) \in E, \quad \forall f \in \Lambda, \quad \forall k \in T \tag{38}$$

Equation (33) ensures the spectrum contiguity requirement at each optical node. Since Eq. (33) is nonlinear, we define a positive integer variable $G_{(j,i),f}^k = f_{(j,i)}^k * X_{(j,i),f}^k$ and use Eqs. (34)-(37) to linearize Eq. (33). Equation (38) ensures the spectrum continuity requirement on each link of sub-tree k.

10) UNIQUENESS AND CAPACITY CONSTRAINTS

$$\sum_{k \in T} X_{(i,j),f}^k \leq 1 \quad \forall (i,j) \in E, \forall f \in \Lambda \tag{39}$$

$$\sum_{k \in T} \sum_{f \in \Lambda} X_{(i,j),f}^k \leq |\Lambda| \quad \forall (i,j) \in E \tag{40}$$

Equation (39) ensures that an FS in a fiber link can only be assigned to one sub-tree. Equation (40) ensures that the total number of FSs assigned to all sub-trees through a fiber link does not exceed this link's capacity.

B. THE ILP MODEL OF THE CST-OM SCHEME

For the CSST-OM scheme, all constructed sub-trees for a multicast demand have a common source data center.

$$\sum_{d \in D} H_d^s \leq 1, \quad \forall s \in S \tag{41}$$

Equation (41) ensures that a multicast service can only be placed in one data center. Therefore, all sub-trees of a multicast demand have a common source data center.

C. THE ILP MODEL OF THE ST-OM SCHEME

For the ST-OM scheme, only one light-tree is constructed to serve a multicast demand.

$$\sum_{k \in T} B_s^k \leq 1, \quad \forall s \in S \tag{42}$$

Equation (42) ensures that a multicast service can be transmitted only through one light-tree.

V. HEURISTIC APPROACH

We define the MSST problem that requires the use of the minimum spectrum resource to serve a multicast demand. To solve the MSST problem, a heuristic approach is developed. It consists of two major steps, requester grouping and sub-tree construction. For requester grouping, we develop the highest modulation-level-aware requester grouping (HMLA-RG) algorithm and the lowest modulation-level-aware requester grouping (LMLA-RG) algorithm. The HMLA-RG algorithm groups requesters according to their highest available modulation levels. The LMLA-RG algorithm groups requesters according to their geographical positions and their lowest available modulation levels. For sub-tree construction, we adopt the distance-adaptive routing and spectrum allocation (DA-RSA) algorithm.

A. THE REQUESTER GROUPING ALGORITHM

The HMLA-RG algorithm and the LMLA-RG algorithm first divide all requesters into several big requester groups according to their distances from data centers. That means that all requesters who are closest to a common data center will be grouped into one big requester group.

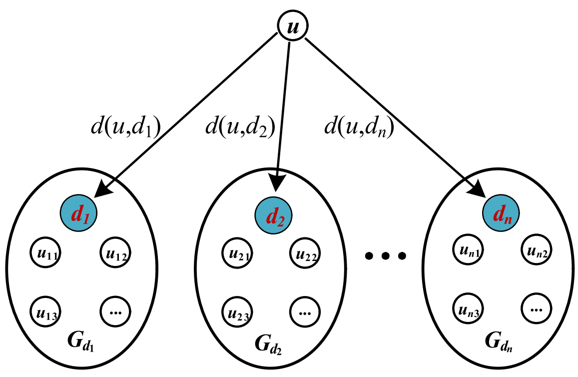


FIGURE 2. Geographical position-based requester grouping.

In Fig. 2, for requester u, the HMLA-RG algorithm and the LMLA-RG algorithm both calculate and compare distances $d(u, d_1), d(u, d_2), \ldots, d(u, d_n)$. Function $d(u, d)$ denotes the shortest distance between requester u and data center d. If requester u is closest to data center d, then u will be added into requester group G_d which denotes a big requester group in which any requester is closest to data center d. Then, the HMLA-RG algorithm groups requesters in G_d according to their highest available modulation levels. It first searches the requester that is the farthest away from the data center and adds it into an empty candidate requester group. Then, it adds the nearest requester into this candidate requester group until the added requester will reduce the highest available modulation level for this candidate requester group. The HMLA-RG algorithm conducts the above process until every requester is added into a requester group.

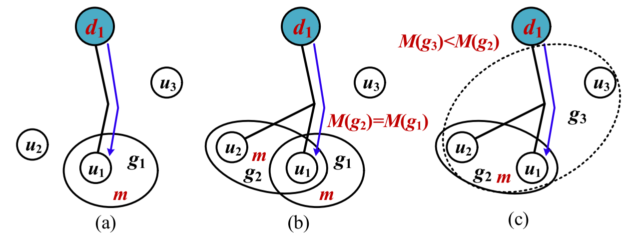


FIGURE 3. Modulation-level-aware requester grouping.

In Fig. 3(a), there exist three requesters, u_1, u_2, and u_3, such that $d(\mathrm{u}_1, d_1) > d(\mathrm{u}_2, d_1) > d(\mathrm{u}_3, d_1)$. At first, we add u_1 into g_1 and calculate the highest available modulation level $M(g_1) = m$, where $M(g)$ denotes the highest available modulation level for requester set g and $M(g)$ is determined by Eq. (1). In Fig. 3(b), u_2 is the nearest requester for light-tree $T(d_1, g_1)$, $N(u_2, T(d_1, g_1)) > \mathrm{N}(u_3, \mathrm{T}(d_1, g_1))$ and $M(u_2 \cup g_1) = M(g_1) = m$, where $N(u, T(d, g))$ denotes the shortest distance between requester u and the shortest-path tree $T(d_1, g_1)$, $N(u, T(d, g)) = \min d(u, v)$, $\forall v \in T(d, g)$. Therefore, we add u_2 into g_1, $g_2 = u_2 \cup g_1$. The HMLA-RG algorithm adds the nearest requester into the candidate requester group until the nearest requester will reduce the highest available modulation level. In Fig. 3(c), if $M(g_2 \cup u_3) < M(g_2)$, u_3 cannot be added into requester set g_2. However, the LMLA-RG algorithm divides all requesters in a big requester group into multiple small requester groups according to their lowest available modulation levels. It first searches the requester that is farthest away from the data center and adds it into an empty candidate requester group. Then, it adds the nearest requester into this candidate requester group until there does not exist an available modulation level for this candidate requester group. In Fig. 3(c), if there is not an available modulation level for requester group $g_2 \cup u_3$, u_3 cannot be added to g_2. In the following pseudo-codes, S denotes a set of multicast services, R denotes a set of multicast demands, s_r denotes the multicast service required by multicast demand r, D_{s_r} denotes a set of data centers that host multicast service s_r, U_r denotes all requesters of multicast demand r, G_d denotes the big requester group in which all requesters are closest to data center d, GD denotes all big requester groups, and GS denotes all requester groups.

The time complexity of the HMLA-RG algorithm is $\mathrm{O}(|R| * |D| * |U|^4 * |V|^2 * (|V| + |U|))$, where $|V|$ is the number of optical nodes, $|D|$ is the number of data centers, $|U|$ is the number of users, and $|R|$ is the number of multicast demands. The following analyses are under the worst case. The time complexity is $|U| * |D| * |V|^2$ from step 4 to

Algorithm 1 The HMLA-RG Algorithm

Input: $G(V, D, U, E), S, R$
Output: GS
1: set $GS = \Phi$;
2: **for** each multicast demand r in R **do**
3: search data center set D_{s_r} which hosts s_r;
4: **for** each requester u in requester set U_r **do**
5: **for** each data center d in D_{s_r} **do**
6: **if** $(d(u, d)$ is the shortest distance$)$ **do**
7: add u into G_d, and add G_d into GD;
8: **end if**
9: **end for**
10: **end for**
11: **for** each requester group G_d in GD **do**
12: search a requester r_f which is farthest away from data center d, build a new requester group g_n, add r_f into g_n, and remove r_f from G_d;
13: **for** each remaining requester r in G_d **do**
14: construct a shortest-path tree T_n from data center d to requesters g_n;
15: **for** each node o which belongs to T_n **do**
16: calculate $d(r, o)$ and find requester r_n which is nearest to T_n;
17: **end for**
18: **end for**
19: **if** $(M(r_n \cup g_n) == M(g_n))$ **do**
20: add r_n into g_n, remove r_n from G_d, go to step 13;
21: **else**
22: go to step 24;
23: **end if**
24: add g_n into GS;
25: **if** $(G_d \neq \Phi)$ **do**
26: go to step 12;
27: **end if**
28: **end for**
29: **end for**
30: **return** GS;

step 10, $|U| * (|V| + |U|) * |V|^2$ from step 15 to step 17, and $|U|^2 (|V|^2 * |U| + |U| * (|V| + |U|) * |V|^2)$ from step 13 to step 23. The total time complexity is $|R| * |D| * |U|^4 * |V|^2 * (|V| + |U|)$. It is guaranteed to run in polynomial time. For the LMLA-RG algorithm, we need to modify only step 19 of the HMLA-RG algorithm to "**if** $(M(r_n \cup g_n)$ is available$)$ **do**". The time complexity of the LMLA-RG algorithm is also $O(|R| * |D| * |U|^4 * |V|^2 * (|V| + |U|))$, which is guaranteed to run in polynomial time.

B. THE DA-RSA ALGORITHM

The DA-RSA algorithm conducts the distance-adaptive routing and spectrum allocation for each sub-tree. Here, we adopt the shortest-path tree algorithm for the sub-tree routing. It first constructs a set of spectrum window planes (SWPs) according to the adopted modulation level. Then, it scans all

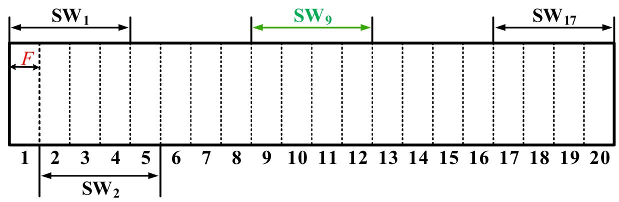


FIGURE 4. The SWs in a fiber link with 20 FSs.

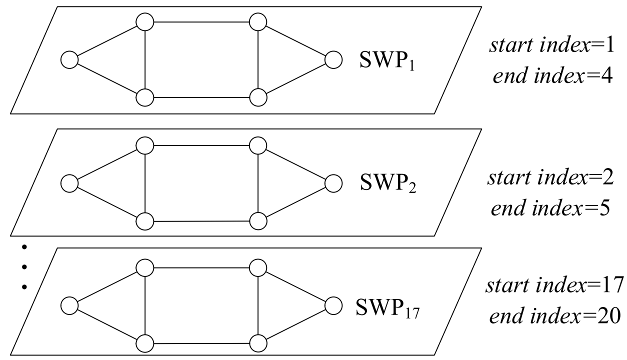


FIGURE 5. All SWPs in n6s8 with 20 FSs.

these SWPs to find a shortest-path tree and uses Eq. (1) to determine if the adopted modulation level is feasible. If there is not a shortest-path tree or the found shortest-path tree does not satisfy Eq. (1), the DA-RSA algorithm will decrease the adopted modulation level and start over until a feasible shortest-path tree is found.

The SWP is based on the spectrum window (SW) that represents a certain number of continuous FSs [37], [38]. The size of an SW is equal to the number of required FSs of a multicast service. For a multicast service with N required FSs, the fiber link with FSs set $\{f_1, f_2, \ldots, f_{|\Lambda|}\}$ contains a total of $(|\Lambda| - N + 1)$ SWs. In Fig. 4, there are 20 FSs in a fiber link. For a multicast demand with 4 required FSs, SW_1 (whose indices range from 1 to 4) is the first SW and the total number of SWs in this link is 17 from SW_1 to SW_{17}. The EO-DCN can be split into multiple SWPs, as presented in Fig. 5. In each SWP, a virtual link between a pair of nodes is connected if all FSs of the corresponding SW in this link are available. The DA-RSA algorithm scans the whole set of SWPs to find the required shortest-path tree. In the following pseudo-codes, *TS* denotes a set of sub-trees, *ST* denotes a set of SWPs, T_g denotes the shortest-path tree constructed for user group g, and T_{swp} denotes a shortest-path tree found in current SWP. Moreover, a requester group can be defined as (d_g, U_g), where d_g denotes the source data center of requester group g and U_g denotes all requesters in requester group g.

The time complexity of **Algorithm 2** is $O(|S| \times |U| \times |M| \times |\Lambda| \times (|U| + |V|)^2)$. The following analyses of time complexity are under the worst case. The time complexity is $|\Lambda| \times (|U| + |V|)^2$ from step 7 to step 16 and $|M| \times |\Lambda| \times (|U| + |V|)^2$ from step 4 to step 17. In step 2, the total number of all constructed requester groups is smaller than $|R| \times |U|$.

Algorithm 2 The DA-RSA Algorithm

Input: $G(V, D, U, E)$, GS
Output: TS
1: set $TS = \Phi$;
2: **for** each requester group g in GS **do**
3: set $n = \left|U_g\right|$, $T_g = null$;
4: **for** each modulation level $m \in M$ (from the highest to lowest level) **do**
5: $N = (\left\lceil \frac{b_s}{m*B} \right\rceil + 1)$, s is the desired multicast service for requester group g;
6: initialize ST with $(|\Lambda| - N + 1)$ SWPs;
7: **for** each spectrum window plane SWP $\in$ ST (from the lowest to highest index) **do**
8: remove all available links, search a shortest-path tree T_{swp} from d_g to U_g for requester group g;
9: **if** ($T_{swp} \neq null$ && the length of the longest branch of T_{swp} is smaller than $S_{m,n}$) **do**
10: **if** ($T_g == null$) **do**
11: $T_g \leftarrow T_{swp}$;
12: **else if** (the total hops of T_{swp} is smaller than that of T_g) **do**
13: $T_g \leftarrow T_{swp}$;
14: **end if**
15: **end if**
16: **end for**
17: **end for**
18: **if** ($T_g \neq null$) **do**
19: $TS \leftarrow (TS \cup T_g)$;
20: **end if**
21: **end for**
22: **return** TS;

Therefore, the time complexity is $|R| \times |U| \times |M| \times |\Lambda| \times (|U| + |V|)^2$ from step 2 to step 21. The **Algorithm 2** is also guaranteed to run in polynomial time.

C. BASELINE ALGORITHM

Yang *et al.* [9] incorporated the rateless network coding (R-NC) in the light-forest scheme and designed several heuristic algorithms to perform the multicast-capable routing, modulation level, and spectrum allocation. Moreover, the proposed heuristic algorithms split the traffic into multiple sub-streams and sent them over several light-trees. In this paper, we do not consider traffic-splitting for any multicast service. Zhu *et al.* [10] proposed a node-based light-tree decomposition and pruning (N-LT-DP) algorithm for the light-forest scheme. The N-LT-DP algorithm precalculated a single large-size light-tree without considering the quality of transmission (QoT) constraint and modified it to a feasible light-forest by iteratively deleting the destination node that has the longest branch. However, the N-LT-DP algorithm did not incorporate the impact of the destination node number of each sub-tree into sub-tree construction. They assumed that a sub-tree can have as many destination nodes as possible as long as the length of each branch did not exceed the given transmission reach of the adopted modulation level. Fan *et al.* [11], [12] proposed a spectrum-aware link-sharing ratio grouping (SA-LRG) algorithm to construct sub-trees. The SA-LRG algorithm grouped requesters according to their feasible paths. However, the SA-LRG algorithm did not consider the QoT constraint while constructing sub-trees. Therefore, the N-LT-DP algorithm and the SA-LRG algorithm are not suitable for the CSST-OM scheme considering the relationship among the adopted modulation level, the transmission distance of the longest branch, and the number of destination nodes, as described in Eq. (1). In this paper, we use the HMLA-RG algorithm and the LMLA-RG algorithm to group requesters for the CSST-OM scheme. The HMLA-RG-based CSST-OM scheme and the LMLA-RG-based CSST-OM scheme are evaluated in Section VI. For the ST-OM scheme, the distance-adaptive routing, modulation level, and spectrum allocation algorithm are evaluated.

VI. PERFORMANCE EVALUATION

In this section, the performances of the proposed DST-OM, CSST-OM, and ST-OM schemes are measured under static and dynamic scenarios. The ILP formulations are conducted on the n6e8 network presented in Fig. 1 (a). The heuristic approach is conducted on NSFNet and the USB as presented in Fig. 9. All fiber links in these topologies are assumed to be bidirectional. The ILP formulations and the heuristic approach are implemented by using JAVA in ILOG CPLEX V12.5 and Eclipse-JEE-LUNA-SR2, respectively.

A. THE STATIC SCENARIO

In the n6e8 network, each fiber link is assumed to contain 50 FSs. All multicast demands are generated in advance, and each has at least two requesters. In the ILP formulation, a multicast service can be hosted in any data center, but all sub-trees of a multicast demand in the CSST-OM scheme have a common source data center and only one light-tree is constructed for a multicast demand in the ST-OM scheme. For each sub-tree, the modulation levels of BPSK, QPSK, and 8QAM are optional.

Fig. 6, Fig. 7, and Fig. 8 present the spectrum consumption, the number of sub-trees, and the number of the backups of multicast services of the proposed DST-OM scheme, the CSST-OM scheme, the ST-OM scheme, and the heuristic approach. Fig. 6 shows that the proposed DST-OM scheme reduces spectrum consumption by approximately 5.27% more than the CSST-OM scheme and by approximately 30.82% more than the ST scheme. In simulations, we find that the DST-OM scheme exhibits good performance when a multicast demand has four requesters. In other words, the DST-OM scheme is more suitable for the multicast demand with a large number of requesters. This advantage is equally verified in the following dynamic scenario. Fig. 7 shows that the total number of the constructed light-trees of the DST-OM scheme equals that of the CSST-OM scheme and

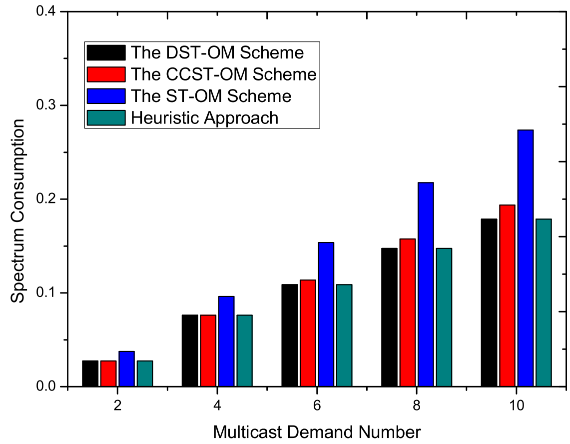


FIGURE 6. Spectrum consumption on n6e8.

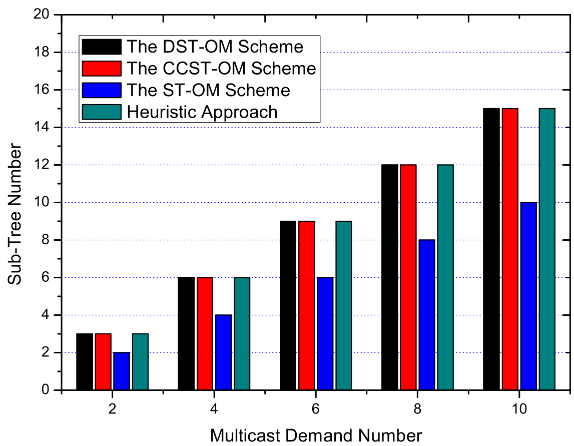


FIGURE 7. The number of sub-trees on n6e8.

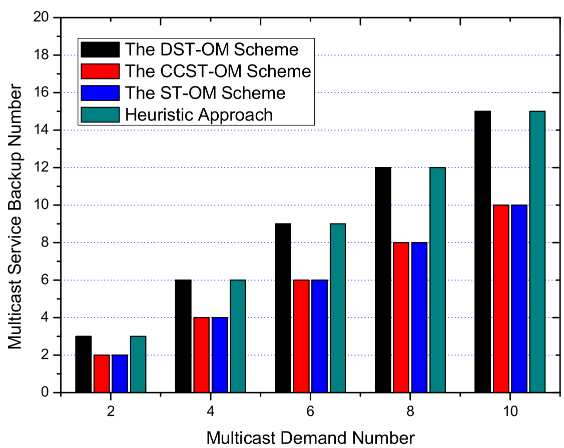


FIGURE 8. The number of the backups of multicast services on n6e8.

is greater than that of the ST scheme. The proposed DST-OM scheme and the CSST-OM scheme both will consume more transmitters. Fig. 8 shows that the DST-OM scheme will consume approximately 50% more storage resources than the CSST-OM scheme and the ST-OM scheme. In other words, approximately half of the multicast services are replicated and maintained in both data centers. The DST-OM scheme reduces the spectrum consumption of multicast demands by consuming more storage resource. For a given network, it is important to conduct the trade-off between spectrum resources and storage resources when deciding which kind of light-tree scheme should be adopted.For current large and ultra-large data centers, storage resources are easily deployed and expanded by adding more hard disk drives and memory banks. It is worth sacrificing storage resources to accommodate more multicast demands. Therefore, the DST-OM scheme provides a promising approach to serve multicast demands in current and future EO-DCNs. Moreover, the performance of the proposed heuristic approach is equivalent to that of the ILP formulation for the DST-OM scheme. In other words, the heuristic approach can also present a good performance in small-scale EO-DCNs. In simulations, CPLEX needs to consume plenty of time to obtain the optimal solution when the number of multicast demands is large. The ILP model of the DST-OM scheme needs to jointly conduct sub-tree construction, modulation-level selection, and spectrum allocation, causing the computational complexity to be very high. In the following sub-section, we verify the performance of the proposed heuristic approach. Using the heuristic approach to realize the DST-OM scheme is efficient and practical for large-scale EO-DCNs with a large number of multicast demands.

B. THE DYNAMIC SCENARIO

In the dynamic scenario, the total number of multicast demands is assumed to be 10000, and the bandwidth requirement of each multicast service is randomly set to {40 Gbps, 60 Gbps, 80 Gbps, 100 Gbps}. For the CSST-OM scheme and the ST-OM scheme, any multicast service is hosted only in one data center. For the DST-OM scheme, any multicast service is hosted in all data centers. For any multicast service, each user has a probability of p. The arrival of each multicast demand follows a Poisson distribution with λ requests per second and the holding time follows the exponential distribution with a mean $1/\mu$. Consequently, the traffic load is measured by λ/μ in Erlang. Since a multicast demand consumes more spectrum resource than the conventional unicast demand, we assume that each link contains 600 FSs.

We measure the blocking probability and the spectrum consumption of the HMLA-RG algorithm-based CSST-OM scheme (HMLA-RG-CSST-OM), the LMLA-RG algorithm-based CSST-OM scheme (LMLA-RG-CSST-OM), the HMLA-RG algorithm-based DST-OM scheme (HMLA-RG-DST-OM), the LMLA-RG algorithm-based DST-OM scheme (LMLA-RG-DST-OM), and the ST-OM scheme when $p = 0.4$ and $p = 0.6$. For NSFNet, Fig. 10 shows that the HMLA-RG-DST-OM scheme can reduce blocking probability by approximately 87.48% better than HMLA-RG-CCST-OM and by approximately 94% better than the ST-OM scheme when $p = 0.4$. Moreover, the blocking probability of the HMLA-RG-CCST-OM scheme is very close to that of the LMLA-RG-CCST-OM scheme, and

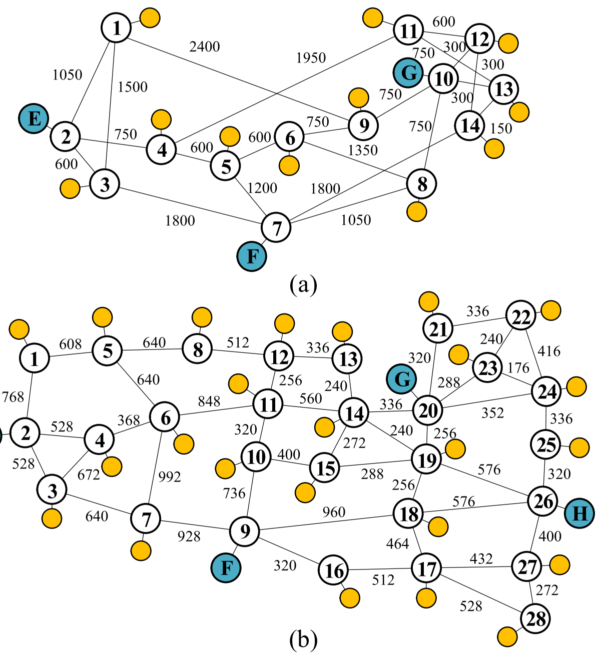


FIGURE 9. Simulation topologies (a) NSFNet, (b) US backbone network (USB).

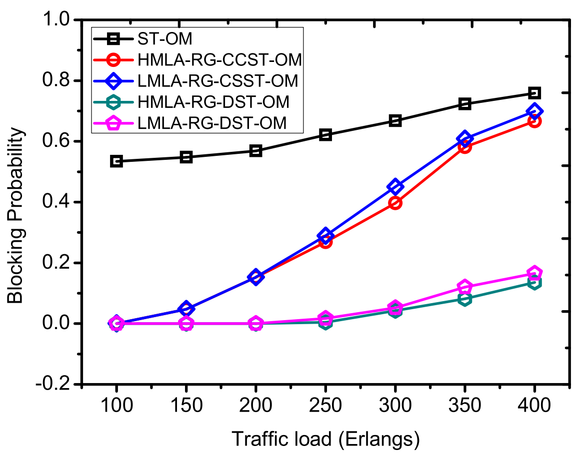


FIGURE 10. Blocking probability on NSFNet with $p = 0.4$.

the blocking probability of the HMLA-RG-DST-OM scheme is also very close to that of the LMLA-RG-DST-OM scheme. In other words, the HMLA-RG algorithm is only slightly better than the LMLA-RG algorithm when serving multicast demands that have relatively few requesters. Fig. 11 shows that the HMLA-RG-DST-OM scheme can reduce blocking probability by approximately 57.8% better than the HMLA-RG-CCST-OM scheme and by approximately 73.9% better than the ST scheme when $p = 0.6$. Moreover, the blocking probability of the HMLA-RG-CCST-OM scheme is very close to that of the LMLA-RG-CCST-OM scheme. However, the HMLA-RG-DST-OM scheme can reduce blocking probability by approximately 26.4% better than the LMLA-RG-DST-OM scheme. In other words, the

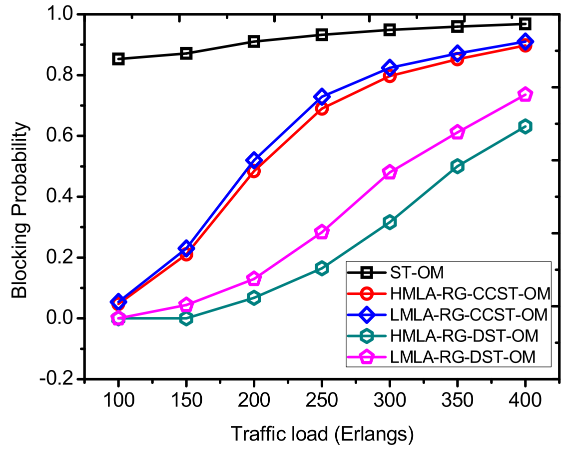


FIGURE 11. Blocking probability on NSFNet with $p = 0.6$.

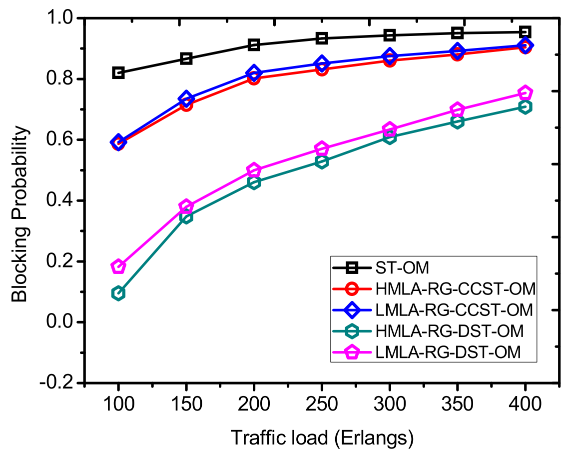


FIGURE 12. Blocking probability on USB with $p = 0.4$.

HMLA-RG algorithm is more suitable for the DST-OM scheme than the CCST-OM scheme when serving multicast demands that have relatively more requesters.

For USB, Fig. 12 shows that the HMLA-RG-DST-OM scheme can reduce blocking rate by approximately 38.9% better than the HMLA-RG-CCST-OM scheme and by approximately 46.5% better than the ST scheme when $p = 0.4$. HMLA-RG-CCST-OM can reduce blocking probability by approximately 1.6% better than the LMLA-RG-CCST-OM scheme and the HMLA-RG-DST-OM scheme can reduce blocking probability by approximately 8.27% better than the LMLA-RG-DST-OM scheme. Fig. 13 shows that the HMLA-RG-DST-OM scheme can reduce blocking probability by approximately 15.9% better than the HMLA-RG-CCST-OM scheme and by approximately 24.9% better than the ST-OM scheme when $p = 0.6$. The HMLA-RG-CCST-OM scheme can reduce blocking probability by approximately 3.68% better than the LMLA-RG-CCST-OM scheme, and the HMLA-RG-DST-OM scheme can reduce blocking probability by approximately 5.37% better than the LMLA-RG-DST-OM scheme. Since the distance between any two nodes in USB is relatively long,the blocking

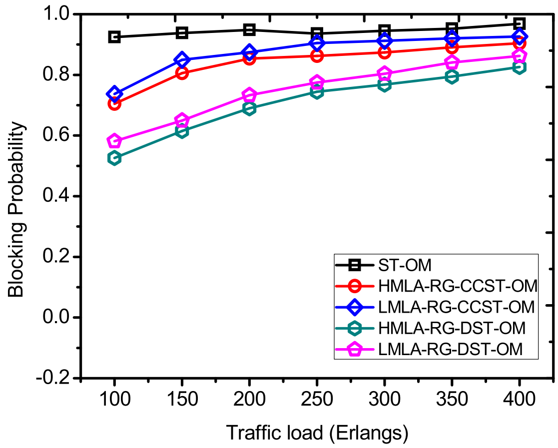


FIGURE 13. Blocking probability on USB with $p = 0.6$.

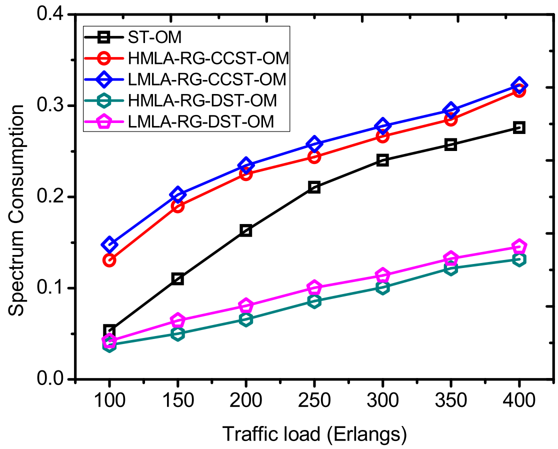


FIGURE 14. Spectrum consumption on NSFNet with $p = 0.4$.

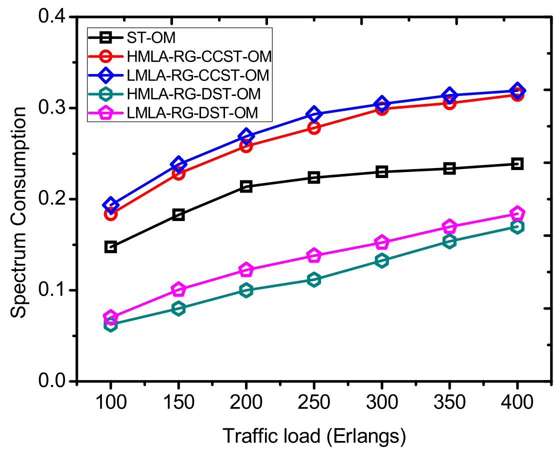


FIGURE 15. Spectrum consumption on NSFNet with $p = 0.6$.

probability is very high for the CCST-OM scheme and the ST-OM scheme. Compared with the HMLA-RG-CCST-OM scheme, the LMLA-RG-DST-OM scheme, and the ST-OM scheme, the HMLA-RG-DST-OM scheme is more suitable for large-scale networks with a large number of requesters.

For NSFNet, Fig. 14 shows that the HMLA-RG-DST-OM scheme can reduce spectrum consumption by approximately

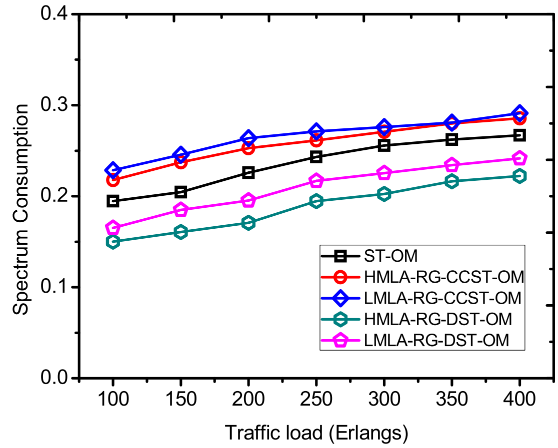


FIGURE 16. Spectrum consumption on USB with $p = 0.4$.

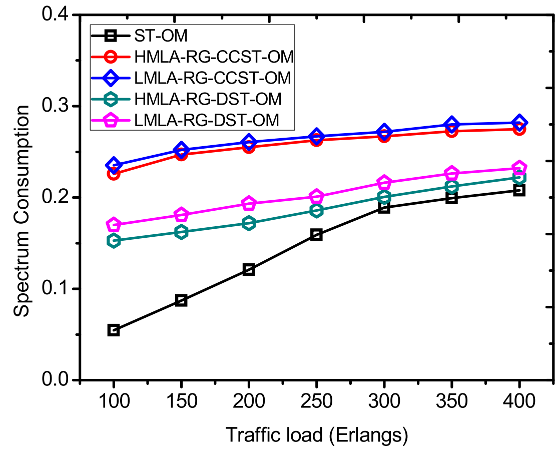


FIGURE 17. Spectrum consumption on USB with $p = 0.6$.

64.2% better than HMLA-RG-CCST-OM when $p = 0.4$. Since the blocking probability of the ST-OM scheme is very low, the spectrum consumption of the ST-OM scheme is lower than that of the DST-OM scheme when $p = 0.4$ and $p = 0.6$. Fig. 15 shows that the HMLA-RG-DST-OM scheme can reduce spectrum consumption by approximately 56.68% better than the HMLA-RG-CCST-OM scheme when $p = 0.6$. Moreover, the spectrum consumption of the HMLA-RG-CCST-OM scheme is lower than that of the LMLA-RG-CCST-OM scheme, and the spectrum consumption of the HMLA-RG-DST-OM scheme is also lower than that of the LMLA-RG-DST-OM scheme. The HMLA-RG algorithm is more suitable for the DST-OM scheme than the CCST-OM scheme when serving multicast demands. For USB, Fig. 16 shows that the HMLA-RG-DST-OM scheme can reduce spectrum consumption by approximately 27.2% better than the HMLA-RG-CCST-OM scheme when $p = 0.4$. Fig. 17 shows that the HMLA-RG-DST-OM scheme can reduce spectrum consumption by approximately 27.8% better than the CCST-OM scheme when $p = 0.6$. Moreover, the HMLA-RG-DST-OM scheme can reduce spectrum consumption by approximately 9.98% better than the

LMLA-RG-DST-OM scheme when $p = 0.4$ and by approximately 7.88% better when $p = 0.6$. The HMLA-RG algorithm presents better performance than the LMLA-RG algorithm for the proposed DST-OM scheme even in a large-scale network. We can find that the proposed DST-OM scheme achieves higher spectrum efficiency and lower blocking probability than the CCST-OM scheme and the ST-OM scheme in small-scale and large-scale networks. The HMLA-RG algorithm presents better performance than the LMLA-RG algorithm both in small-size and large-size networks when $p = 0.4$ and $p = 0.6$.

VII. CONCLUSION

In this paper, we study the problem of multicast service provisioning while leveraging service backup among multiple distributed data centers in EO-DCNs. A DST-OM scheme that achieves low blocking probability and high spectrum efficiency when serving multicast demands is proposed. Moreover, the MSST problem is also defined. The HMLA-RG algorithm presents better performance than the LMLA-RG algorithm while addressing the MSST problem. The HMLA-RG-DST-OM scheme provides a promising approach to serving multicast demands in EO-DCNs.

ACKNOWLEDGMENT

This paper was presented in part at the Optical Fiber Communications Conference, Los Angeles, CA, USA, in 2017.

REFERENCES

[1] Y. Wu, Z. Zhang, C. Wu, C. Guo, Z. Li, and F. C. M. Lau, "Orchestrating bulk data transfers across geo-distributed datacenters," *IEEE Trans. Cloud Comput.*, vol. 5, no. 1, pp. 112–125, Jan./Mar. 2015.

[2] J. Chen, Y. Gong, M. Fiorani, and S. Aleksic, "Optical interconnects at the top of the rack for energy-efficient data centers," *IEEE Commun. Mag.*, vol. 53, no. 8, pp. 140–148, Aug. 2015.

[3] A. Muhammad, M. Fiorani, L. Wosinska, and J. Chen, "Joint optimization of resource allocation for elastic optical intra-datacenter network," *IEEE Commun. Lett.*, vol. 20, no. 9, pp. 1760–1763, Sep. 2015.

[4] W. Xia, P. Zhao, Y. Wen, and H. Xie, "A survey on data center networking (DCN): Infrastructure and operations," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 640–656, 1st Quart., 2016.

[5] S. Mukherjee, F. Bronzino, S. Srinivasan, J. Chen, and D. Raychaudhuri, "Achieving scalable push multicast services using global name resolution," in *Proc. GLOBECOM*, Washington, DC, USA, Dec. 2016, pp. 1–6.

[6] G. N. Rouskas, "Optical layer multicast: Rationale, building blocks, and challenges," *IEEE Netw.*, vol. 17, no. 1, pp. 60–65, Jan. 2003.

[7] R. Lin, M. Zukerman, G. Shen, and W.-D. Zhong, "Design of light-tree based optical inter-datacenter networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 5, no. 12, pp. 1443–1455, Dec. 2013.

[8] T. Rahman, G. Ellinas, and M. M. Ali, "Building light-forest to support group multicast in mesh-based optical grid networks," in *Proc. Opt. Fiber Commun. Conf. Expo. (OFC)*, Anaheim, CA, USA, Mar. 2006, p. 3.

[9] L. Yang, L. Gong, F. Zhou, B. Cousin, M. Molnár, and Z. Zhu, "Leveraging light forest with rateless network coding to design efficient all-optical multicast schemes for elastic optical networks," *J. Lightw. Technol.*, vol. 33, no. 18, pp. 3945–3955, Sep. 15, 2015.

[10] Z. Zhu *et al.*, "Impairment- and splitting-aware cloud-ready multicast provisioning in elastic optical networks," *IEEE/ACM Trans. Netw.*, vol. 25, no. 2, pp. 1220–1234, Apr. 2017.

[11] Z. Fan, Y. Li, G. Shen, and C. C.-K. Chan, "Dynamic resource allocation for all-optical multicast based on sub-tree scheme in elastic optical networks," in *Proc. Opt. Fiber Commun. Conf. Expo. (OFC)*, Anaheim, CA, USA, Mar. 2016, pp. 1–3.

[12] Z. Fan, Y. Li, G. Shen, and C.-K. C. Chan, "Distance-adaptive spectrum resource allocation using subtree scheme for all-optical multicasting in elastic optical networks," *J. Lightw. Technol.*, vol. 35, no. 9, pp. 1460–1468, May 1, 2016.

[13] M. Ruiz and L. Velasco, "Performance evaluation of light-tree schemes in flexgrid optical networks," *IEEE Commun. Lett.*, vol. 18, no. 10, pp. 1731–1734, Oct. 2014.

[14] J. Yao, P. Lu, L. Gong, and Z. Zhu, "On fast and coordinated data backup in geo-distributed optical inter-datacenter networks," *J. Lightw. Technol.*, vol. 33, no. 14, pp. 3005–3015, Jul. 15, 2015.

[15] D. Boru, D. Kliazovich, F. Granelli, P. Bouvry, and A. Y. Zomaya, "Models for efficient data replication in cloud computing datacenters," in *Proc. IEEE Int. Conf. Commun. (ICC)*, London, U.K., Jun. 2015, pp. 6056–6061.

[16] M. F. Habib, M. Tornatore, and B. Mukherjee, "Fault-tolerant virtual network mapping to provide content connectivity in optical networks," in *Proc. Opt. Fiber Commun. Conf. Expo. (OFC)*, Anaheim, CA, USA, Mar. 2013, pp. 1–3.

[17] S. Ferdousi, F. Dikbiyik, M. F. Habib, M. Tornatore, and B. Mukherjee, "Disaster-aware datacenter placement and dynamic content management in cloud networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 7, no. 7, pp. 681–694, Jul. 2015.

[18] X. Li *et al.*, "Design of K-node (edge) content connected optical data center networks," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 466–469, Mar. 2016.

[19] S. S. W. Lee, H.-H. Yen, and A. Chen, "Traffic grooming for IP multicast over WDM networks using light-path and light-tree schemes," in *Proc. 9th Int. Conf. Netw. (ICN)*, Menuires, France, Apr. 2010, pp. 291–293.

[20] X. Huang, F. Farahmand, and J. P. Jue, "Multicast traffic grooming in wavelength-routed WDM mesh networks using dynamically changing light-trees," *J. Lightw. Technol.*, vol. 23, no. 10, pp. 3178–3187, Oct. 2005.

[21] G. M. Fernandez, C. Vazquez, P. C. Lallana, and D. Larrabeiti, "Tap-and-2-split switch design based on integrated optics for light-tree routing in WDM networks," *J. Lightw. Technol.*, vol. 27, no. 13, pp. 2506–2517, Jul. 1, 2009.

[22] D. Andrei, M. Tornatore, C. U. Martel, D. Ghosal, and B. Mukherjee, "Provisioning subwavelength multicast sessions with flexible scheduling over WDM networks," *J. Opt. Commun. Netw.*, vol. 2, no. 5, pp. 241–255, May 2011.

[23] Y. Zhu, X. Gao, W. Wu, and J. P. Jue, "Efficient impairment-constrained 3r regenerator placement for light-trees in optical networks," *J. Opt. Commun. Netw.*, vol. 3, no. 4, pp. 359–371, Apr. 2011.

[24] F. Zhou, M. Molnár, B. Cousin, and C. Qiao, "Cost bounds and approximation ratios of multicast light-trees in WDM networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 3, no. 4, pp. 323–334, Apr. 2011.

[25] Q. Wang and L.-K. Chen, "Performance analysis of multicast traffic over spectrum elastic optical networks," in *Proc. Opt. Fiber Commun. Conf. Expo. (OFC)*, Los Angeles, CA, USA, Mar. 2012, pp. 1–3.

[26] L. Gong, X. Zhou, X. Liu, W. Zhao, W. Lu, and Z. Zhu, "Efficient resource allocation for all-optical multicasting over spectrum-sliced elastic optical networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 5, no. 8, pp. 836–847, Aug. 2013.

[27] K. Walkowiak, R. Goścień, M. Klinkowski, and M. Woźniak, "Optimization of multicast traffic in elastic optical networks with distance-adaptive transmission," *IEEE Commun. Lett.*, vol. 18, no. 12, pp. 2117–2120, Dec. 2014.

[28] S. Li, W. Lu, X. Liu, and Z. Zhu, "Fragmentation-aware service provisioning for advance reservation multicast in SD-EONs," *Opt. Exp.*, vol. 23, no. 20, pp. 25804–25813, Sep. 2015.

[29] A. Cai, J. Guo, R. Lin, G. Shen, and M. Zukerman, "Multicast routing and distance-adaptive spectrum allocation in elastic optical networks with shared protection," *J. Lightw. Technol.*, vol. 34, no. 17, pp. 4076–4088, Sep. 1, 2016.

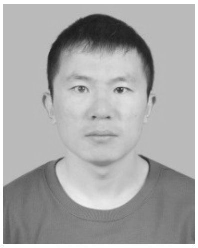
[30] L. Gifre *et al.*, "Experimental assessment of inter-datacenter multicast connectivity for Ethernet services in flexgrid networks," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Cannes, France, Sep. 2014, pp. 1–3.

[31] K. Tokas *et al.*, "A scalable optically-switched datacenter network with multicasting," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Athens, Greece, Jun. 2016, pp. 265–270.

[32] W.-K. Jia and L.-C. Wang, "A unified unicast and multicast routing and forwarding algorithm for software-defined datacenter networks," *IEEE J. Sel. Area Commun.*, vol. 31, no. 12, pp. 2646–2657, Dec. 2013.

[33] M. Zeng, W. Fang, J. J. P. C. Rodrigues, and Z. Zhu, "Orchestrating multicast-oriented NFV trees in inter-DC elastic optical networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.

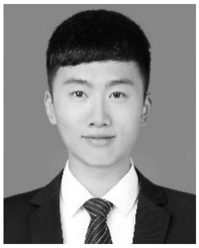
[34] Y. Zhu and J. P. Jue, "Multi-class flow aggregation for IPTV content delivery in IP over optical core networks," *J. Lightw. Technol.*, vol. 27, no. 12, pp. 1891–1903, Jun. 15, 2009.
[35] C. Wang, G. Shen, and S. K. Bose, "Distance adaptive dynamic routing and spectrum allocation in elastic optical networks with shared backup path protection," *J. Lightw. Technol.*, vol. 33, no. 14, pp. 2955–2964, Jul. 15, 2015.
[36] N. Sambo *et al.*, "Demonstration of data and control plane for optical multicast at 100 and 200 Gb/s with and without frequency conversion," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 5, no. 7, pp. 667–676, Jul. 2013.
[37] X. Liu, L. Gong, and Z. Zhu, "Design integrated RSA for multicast in elastic optical networks with a layered approach," in *Proc. GLOBECOM*, Atlanta, GA, USA, Dec. 2013, pp. 2346–2351.
[38] L. Gong and Z. Zhu, "Virtual optical network embedding (VONE) over elastic optical networks," *J. Lightw. Technol.*, vol. 32, no. 3, pp. 450–460, Feb. 1, 2014.



XIN LI was born in Shandong, China, in 1986. He received the B.S. degree in communication engineering from the Shandong University of Science and Technology in 2008 and the Ph.D. degree in communication and information system from the Beijing University of Posts and Telecommunications (BUPT) in 2014. He is currently a Lecturer with BUPT. His research interests include software-defined optical networking, elastic optical networks, datacenter networking, and optical network survivability.

LU ZHANG was born in Shandong, China, in 1994. She received the B.S. degree in communication engineering from the Central South University of Forestry and Technology in 2016. She is currently a Graduate Student majored in electronics and communication engineering with the Beijing University of Posts and Telecommunications. Her research interests include artificial intelligence and photonic firewall techniques.

YING TANG was born in Guangxi, China, in 1993. She received the B.S. degree from Xidian University, Xi'an, China, in 2016. She is currently pursuing the Ph.D. degree with the Beijing University of Post and Telecommunication. Her research interests include routing and spectrum allocating algorithm, optical network virtualization, artificial intelligence, and photonic firewall techniques.



JUNFENG GUO was born in Zhejiang, China. He received the B.S. degree in electronic information engineering from Xiangtan University in 2017. He is currently pursuing the master's degree with the State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications. His research interests include optical networks and artificial intelligence.

SHANGUO HUANG (M'09) received the Ph.D. degree from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2006. He is currently a Professor with the State Key Laboratory of Information Photonics and Optical Communications and the Dean of the School of Science, BUPT. He has been actively undertaking several national projects, published three books and over150 journals and refereed conferences, and authorized 14 patents. His current research interests include networks designing, planning, traffic control and resource allocations, routing algorithms, and performance analysis. He was a recipient of the Beijing Higher Education Young Elite Teacher, the Beijing Nova Program, and the Program for New Century Excellent Talents in University from the Ministry of Education, from 2011 to 2013, respectively.