A Ranked Partial Protection Scheme for Degraded Services in Elastic Optical Networks

Lu Zhang
School of Electronic Engineering
Beijing University of Posts and Telecommunications
Beijing, China
luzhang@bupt.edu.cn

Xin Li
School of Electronic Engineering
Beijing University of Posts and Telecommunications
Beijing, China
xinli@bupt.edu.cn

Jingjie Xin
School of Electronic Engineering
Beijing University of Posts and Telecommunications
Beijing, China
jingjiex@bupt.edu.cn

Chenyu Zhao
School of Electronic Engineering
Beijing University of Posts and Telecommunications
Beijing, China
chenyuzhao@bupt.edu.cn

Daixuan Li
School of Electronic Engineering
Beijing University of Posts and Telecommunications
Beijing, China
lidaixuan@bupt.edu.cn

Shanguo Huang
School of Electronic Engineering
Beijing University of Posts and Telecommunications
Beijing, China
shghuang@bupt.edu.cn

***Abstract*—A data division based ranked partial protection (RPP) scheme is proposed for service differentiated degradation in elastic optical networks. A resource aware RPP algorithm is developed. Results show it improves blocking rate and resource utilization.**

Keywords—Degraded services, differentiated degradation, failure protection, multi-paths routing

I. INTRODUCTION

As of today, optical networks represent the most important communication networks, by virtue of their inherent advantages, such as large capacity, wide bandwidth, low loss, good confidentiality and so on. Especially, the elastic optical network (EON) with the flexibly access ability to the massive bandwidth in optical fibers, is an appealing optical network paradigm. In the face of an explosive growth of network traffic, it has been a hot research topic to relieve the strain of optical network resources. Especially during disasters, network resource insufficiency is more prominent. Thus, degraded service is proposed, which means that services, such as video streaming and file transfers, can still be accepted with reduced bandwidth under a certain tolerance [1]. In addition, optical networks spanning large areas are subject to various failures, such as natural disasters, operation errors, system aging, and malicious attacks [2]. These failures can affect network availability and may cause massive losses of service data. Network survivability is a basic problem for avoiding communication disruption due to failures. Protection is a popular survivability strategy to provide reliability for services [3]. For degraded services, partial protection (PP), providing protection for a part of the requested bandwidth, is more efficient than full protection (FP), i.e., the conventional primary-backup method, to alleviate extensive resource consumption [1], [3], [4].

There are several studies on PP for degraded service. Authors in [3] used multipath routing to support degraded services by ensuring a specified requested bandwidth under failures in Mixed-Line-Rate WDM networks. A content connectivity based multipaths scheme is proposed for the reliable provision of degraded services in [5]. To improve the quality of experience of users and network resource utilization, custom-tailored differentiated degradation for degraded services has been an attractive research direction. There have been many studies on the classification and gradation of services. The authors of [6], [7] provide differentiated degradation for different classes of services. However, exist researches do not consider the inherent difference of the data of the service to be transmitted. With the development of the 6G communication networks and Internet of Things (IoT), content-based services, such as video delivery, augmented reality (AR) and virtual reality (VR) are booming. Content retrieval-based applications contribute the most Internet traffic [8]. With the support of context-awareness, the whole service data can be divided into multiple parts based on their relevance to user request and can be graded accordingly. It may be unreasonable to downgrade the data of a service in a one-size-fits-all manner. If the received service data is of low priority, the small amount of information will result in a poor user experience. Therefore, it is necessary to develop differentiated protection methods based on context-aware data grading for degraded service.

In this paper, a ranked partial protection (RPP) scheme based on service data division is proposed for degraded services in EONs. The data of a service can be divided into several parts, and graded accordingly depending on the user's requirements for reliability. Different protection strategies are adaptively performed on different levels data. A particle swarm optimization (PSO) based RPP (PSO-RPP) algorithm is developed to realize data partition and assignment, aiming at minimizing resource consumption. Results show that compared with the FP scheme, the proposed scheme can make more efficient use of network resources to serve more services.

II. THE RPP SCHEME

This section elaborates our proposed RPP scheme for services with data gradation in EONs. Simply, the degraded service data is split into two levels, i.e., high-level data and low-level data, according to the reliability requirements of receiving end. The high-level data must be correctly received, because it will cause errors in subsequent works once it is lost or destroyed. The low-level data are the part that has little impact on subsequent works, and it just be received as much as possible. In the RPP scheme, we adopt backup strategy for the high-level data to improve reliability. For low-level data, adaptive splitting strategy is used to balance network load according to available network resources.

In the RPP scheme, the data transmitted through each path includes all high-level data and partial low-level data. The high-level data is transmitted through each path, so the degraded service is still acceptable when either path fails. On the other hand, partial low-level data is transmitted through each path, and the sum of low-level data transmitted through all paths is the full low-level data. The constraints of the RPP scheme are described as (1-5). Equation (1) is used to calculate the number of paths k, which is determined by the number of failures f. The data proportion transmitted through each path is determined by (2-5), where h is the proportion of high-level data of the service, l_j is the proportion of low-level data transmitted by corresponding path p_j, x_j is the total data proportion transmitted by path p_j, z_j is the percentage of the low-level data proportion in p_j, and u is the threshold of the percentage of the low-level data proportion in each path. $D = u*(1-h)+h$ is the lower limit of the data proportion that should be retained in each path when the service is degraded. Formulas (2-4) ensure that $x_j \geq D \geq h$. $x_j \geq D$ denotes service provision can meet degradation requirement under f failures. $x_j \geq h$ means the data transmitted through each path includes all high-level data and partial low-level data. $D \geq h$ means the high-level data of service can be reliably received under f failures. Equation (5) ensures the sum of low-level data transmitted through all paths is the full low-level data. Formulas (1-5) denotes that the required total spectrum resources of all paths in the RPP scheme is $\sum_{j=1}^{k} x_j * B = (1+h*f)*B$, where B is the transmission rate requested by the service. Thus, the RPP scheme achieves $\left[(f+1)*1-(1+h*f)\right]*B = (1-h)*f*B$ resources saving than the FP scheme. Given the above constraints, k disjoint paths and the data proportion assigned on each path can be adaptively selected, based on available resources, transmission delay, etc., or their weighted combination, to transmit a degraded service against f failures. On the other hand, as the data proportion transmitted by each path in the RPP scheme is not larger than in the FP scheme, i.e., $x_j \leq 1$, the probability of successful spectrum resource allocation for services in the RPP scheme is higher than in the FP scheme.

$$k = f+1 \tag{1}$$

$$x_j = h + l_j \quad j \in [1,2,\cdots,k] \tag{2}$$

$$l_j = z_j(1-h) \quad 0 \leq z_j \leq 1 \tag{3}$$

$$z_j \geq u \quad 0 \leq u \leq (1/k) \tag{4}$$

$$\sum_{j=1}^{k} z_j = 1 \tag{5}$$

Figure 1 illustrates the comparison of the RPP scheme and the FP scheme in an EON, where the notations on the link denotes the link length and frequency slots (FSs) occupancy. The EON is modeled as $G(V, E)$, where V is the set of optical nodes, and E is the set of fiber links. The number of FSs in each link is Λ=8. As shown in Fig.1 (a), there are two service requests r_1 (s=1, d=3, B=100Gbps, h=0.5, f=1) and r_2 (s=4, d=2, B=60Gbps, h=0.5, f=1) in Fig.1, where s is the source node, and d is the destination node. For the RPP scheme, according to formula (1), two disjoint paths should be calculated for r_1 and r_2. We set $u = 50\%$ for r_1, thus the lower limit of transmission rate in each path is $BD = B*D = B*\left[u*(1-h)+h\right] = 100\text{Gbps}*\left[50\%*(1-0.5)+0.5\right] = 75\text{Gbps}$. Similarly, the lower limit of transmission rate in each path for r_2 is 36Gbps when $u = 20\%$. As shown in Fig.1 (b), two disjoint paths p_{11}($1 \rightarrow 0 \rightarrow 3$) and p_{12}($1 \rightarrow 2 \rightarrow 3$)are calculated for r_1 , and two other disjoint paths p_{21}($4 \rightarrow 3 \rightarrow 2$) and p_{22}($4 \rightarrow 5 \rightarrow 2$) are calculated for r_2, using k-shortest paths method. According to (1-5), the sum of required transmission rate of all paths for r_1 is $(1+h*f)*B = (1+0.5*1)*100\text{Gbps} = 150\text{Gbps}$. Therefore, the transmission rate of p_{11} or p_{12} is 75Gbps. However, the sum of required transmission rate of all paths for r_2 is $(1+h*f)*B = (1+0.5*1)*60\text{Gbps} = 90\text{Gbps}$. Therefore, the transmission rate of one path ranges from 36Gbps to 45Gbps, and the transmission rate of the other path ranges from 45Gbps to 54Gbps. We set the transmission rates of $p_{21}$and p_{22} are 40Gbps and 50Gbps, respectively. In this paper, we adopt three kinds of modulation formats, i.e., BPSK, QPSK, and 8QAM. According to the transmission rate per FS and maximum transmission distance corresponding to each modulation format [9], the modulation formats adopted by the paths p_{11} and p_{12} are 8QAM, the modulation formats adopted by the paths $p_{21}$and p_{22} are QPSK. Therefore, the number of required FSs for r_1 in each link of p_{11} or p_{12} is $\lceil 75/(3*12.5) \rceil = 2$. Similarly, the number of FSs required by r_2 in each link of p_{21}($4 \rightarrow 3 \rightarrow 2$) and p_{22}($4 \rightarrow 5 \rightarrow 2$) is $\lceil 40/(2*12.5) \rceil = 2$ and $\lceil 50/(2*12.5) \rceil = 2$, respectively.. While for the FP scheme, the number of required FSs for r_1 in each link of p_{11} or p_{21} is $\lceil 100/(3*12.5) \rceil = 3$, while the number of required FSs for r_2 in each link of p_{21} or p_{22} is $\lceil 60/(2*12.5) \rceil = 3$. As shown in Fig.1 (c), after spectrum is allocated to r_1 on the link $3 \rightarrow 2$, there are no longer 3 consecutive available FSs. Thus the path p_{21} cannot be established for r_2. In this paper, if any path for a service request fails to be established, that is, failures cannot be prevented, the service request should be blocked and resources are not allocated to all paths. Therefore, compared with the RPP scheme, the FP scheme only serves the request r_1, and 4 more FSs are consumed.

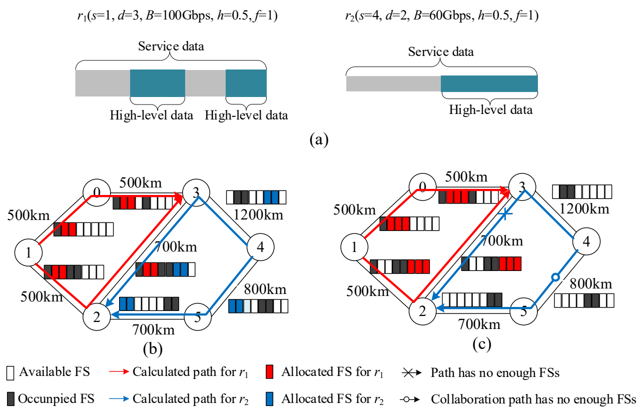

Fig. 1. (a) Service requests (b) RPP scheme (c) FP scheme.

III. Resource Aware RPP Algorithm

In this section, a resource aware RPP algorithm based on PSO is developed. In the RPP algorithm, the PSO algorithm is used to adaptively determine the optimal data proportion assignment according to available spectrum resource of calculated paths, with the aim of load balance. The main procedure of the RPP algorithm is stated in Algorithm 1. The main content of the algorithm is to firstly calculate whether there are enough link-disjoint paths against f failures. We adopt the dijkstra algorithm based on link pruning to calculate link-disjoint paths. The maximum number of link-disjoint paths is set to k. If k link-disjoint paths can be calculated, then the algorithm jumps to the sub-algorithm 1, i.e., a PSO algorithm, to obtain final data proportion assignment solution; Otherwise, the request is blocked. Finally, spectral resources is allocated to establish the transmission paths.

Algorithm 1: PSO-RPP Algorithm

Input: G (V, E), a set of service requests R
Output: routing and resource allocation

for each $r(s,d,B,h,f) \in R$ **do**
  Determine the value of k according to formulas (1);
  Search k shortest link-disjoint paths between s and d in G (V, E), the paths set is denoted as P;
  if $k > |P|$ **do**
    r is blocked;
  else
    Decide the data transmitted by each path using **Sub-Algorithm 1**;
    Perform resource allocation for r.
  end if
end for

In the PSO algorithm, each particle can be represented as a potential solution to the problem. In this paper, a particle is made up of k elements, and each element corresponds to a path. The value of each element is the proportion of data transmitted through each path. For any particle i, it can be denoted as $X^i = \left(x_1^i, x_2^i, \cdots, x_k^i\right)$. Each particle has a velocity $V^i = \left(v_1^i, v_2^i, \cdots, v_k^i\right)$ that determines the search direction to get closer to the optimal solution. In the Sub-Algorithm 1, we firstly initialize the particle population and the velocity of each particle. According to (1-5), each particle must have k non-zero elements, the value range of an element is $x_j^i \in \left[h + u(1-h), h + (1-u)(1-h)\right]$, and the sum of all element values of a particle is $1 + h * f$. In addition, (6) describes the updating of the values of each element of a particle. Hence, to make x_j^i variable in its range, the range of v_j^i is $v_j^i \in \left[h + u(1-h) - x_j^i, h + (1-u)(1-h) - x_j^i\right]$. While the updating of the velocity of a particle is calculated as (7), which is related to the optimal solution found by the particle itself called the individual extremum $Y_{best}^i = \left(y_1^i, y_2^i, \cdots, y_k^i\right)$, and the optimal solution found from the whole population called the global extremum $Q_{best} = \left(q_1, q_2, \cdots, q_k\right)$. Equation (7) consists of three parts, $\lambda_0 * v_j^i$ is the inertial part, $\lambda_1 \beta_1 \left(y_j^i - x_j^i\right)$ is the individual learning part, and $\lambda_2 \beta_2 \left(g_j - x_j^i\right)$ is the population learning part, where β_1 and β_2 are learning factors, λ_0, λ_1 and λ_2 are weight factors between [0, 1].

$$x_j^i = x_j^i + v_j^i \quad j \in (1, 2, \cdots, k) \tag{6}$$

$$v_j^i = \lambda_0 * v_j^i + \lambda_1 \beta_1 \left(y_j^i - x_j^i\right) + \lambda_2 \beta_2 \left(q_j - x_j^i\right) \quad j \in (1, 2, \cdots, k) \tag{7}$$

$$Fitness^i = \xi_0 * \sum_{j=1}^{k} C_j^i + \xi_1 * e_1^i + \xi_2 * e_2^i \tag{8}$$

$$C_j^i = (x_j^i * B / m_j * 12.5) * N_j \tag{9}$$

$$e_1^i = \begin{cases} 0, & \text{when meet (1-5)} \\ MAX, & \text{when not meet (1-5)} \end{cases} \tag{10}$$

$$e_2^i = \begin{cases} -A, & w = 1 \\ MAX, & w = 0 \end{cases} \tag{11}$$

Sub-Algorithm 1: PSO algorithm

Input: G (V, E), P, population size S, and maximum iteration period T
Output: Data proportion assignment

1: Initialize the population, and set current iteration period $t = 0$;
2: Calculate the fitness value of each particle in the population; Obtain { Y_{best}^i }and Q_{best};
3: Update the element values and velocity of each particle according to formula (6) and (7), respectively;
4: If $t == T$ or the fitness value of the two generations of optimal particle is less than ε, algorithm is terminated and the output result is Q_{best}; Otherwise, $t = t + 1$, and jump to Step 2.

The selection of the individual extremum and the global extremum are based on the fitness function. The fitness function reflects the difference between the particle and the optimal solution. The fitness function of each particle in the PSO algorithm is expressed by (8), which is composed of two parts, spectrum resource consumption and penalty, where ξ_0, ξ and ξ_2 are the weight factors. Formula (9) defines spectrum resource consumption C_j^i of each path, where m_j is modulation format, and N_j is the number of hops of the path. The penalty contains two parts, routing penalty e_1^i and resource penalty e_2^i, which are expressed by (10) and (11), respectively. The penalty prevents the particle from entering the impossible region. e_1^i constrains the number of used paths and the data proportion, where MAX is a large constant. e_2^i is the network resource constraint, w equals 1 if there are enough resources for the paths, and 0 otherwise, where A is the sum of available resources on k paths. In this algorithm, particles with smaller fitness values are more likely to be selected. Y_{best}^i is updated by comparing the fitness value of the current particle and the fitness value of the historical individual extremum. If the fitness value of the current particle is smaller, Y_{best}^i is replaced by the element values of the current particle. The update of Q_{best} is based on the comparison of the fitness values of all particles in the current population with the fitness value of the historical global extremum.

IV. RESULTS ANALYSIS

This Section presents the simulation setup and some numerical results based on extensive simulations. In the RPP algorithm, we set S=20, T=30, $\lambda_0 = 0.5$, $\lambda_1 = 0.3$, $\lambda_2 = 0.2$, $\beta_1 = 1$, $\beta_2 = 1$, $\xi_0 = 100$, $\xi_1 = 1$ and $\xi_2 = 1$. The 24-node U.S. network topology with Λ=100 is used as the simulation topology. The performance of the RPP scheme is evaluated from three aspects. First, the performance of the RPP scheme and the FP scheme under different proportion of high-level data is compared. Second, the influence of different limitation of low-level data proportion is evaluated. Third, we compare the RPP scheme with the FP scheme in terms of blocking probability and resource consumption, where *h* of each service request is randomly selected from {0.2, 0.4, 0.6, 0.8} with equal probability. The reason of blocking is twofold. One is the number of disjoint paths needed is not enough (path shortage), due to the connectivity limitations between nodes or the longest distance limitation that a fiber path can reach in different modulation formats. Another is the shortage of available FSs (resource shortage).

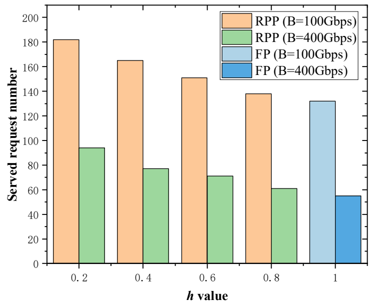

Fig. 2. Comparison of the number of served services under different *h* values.

Figure 2 shows the comparison of the number of services that can be served by the RPP scheme and the FP scheme against single failure under different proportions of high-level data, i.e., *h*. When *h*=1, the RPP scheme is equivalent to the FP scheme. In terms of the overall trend, with the increase in the proportion of high-level data, the number of served services decrease. When *h* is 0.2, 0.4, 0.6 and 0.8, the RPP scheme can carry 37.9%, 25%, 14.4% and 4.5% more requests than the FP scheme at *B*=100Gbps, respectively. When *h* is 0.2, 0.4, 0.6 and 0.8, the RPP scheme can carry 70.9%, 40%, 29.1% and 10.9% more requests than the FP scheme at *B*=400Gbps, respectively. Thus the higher the transmission rate of the requests, the more obvious the advantage of the RPP scheme. The number of served requests at *B*=100Gbps is slightly more than double that at *B*=400Gbps.

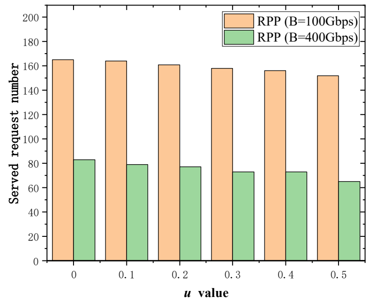

Fig. 3. Comparison of the number of served services under different u values.

As shown in Fig. 3, we evaluate the performance of the RPP scheme under different constraints of low-level data proportion of each path under *h*=0.4, *f*=1. With the increase in u value, the number of served services slightly decrease. The results may be caused by the same optimal allocation of low-level data when u takes different values in the RPP scheme.

Figure 4 shows the comparison of blocking probability and spectrum resources consumption between the RPP scheme and FP scheme under single failure. The source node and destination node of request are randomly generated, and the transmission rate is 100Gbps. From Fig. 4 (a), the blocking caused by path shortage of both the RPP and FP schemes is about 2.5%. In addition, when the service number reaches 100, blocking occurs due to resource shortage in the RPP scheme. While when service number is 80, some service requests are already blocked due to resource shortage in the FP scheme. From a general perspective, the blocking probability caused by resource shortage of both schemes increases with the increase of the number of services, but the RPP scheme increases more slowly. Specifically, the blocking probability of the proposed scheme is lower than the FP scheme with same number of services. The blocking probability caused by resource shortage in our scheme is about 11% less than in the FP scheme. As shown in Fig. 4 (b), as the number of requests increases, resource consumption increases for both schemes. In terms of the same number of requests, the RPP scheme consumes fewer resources but serves more requests the FP scheme.

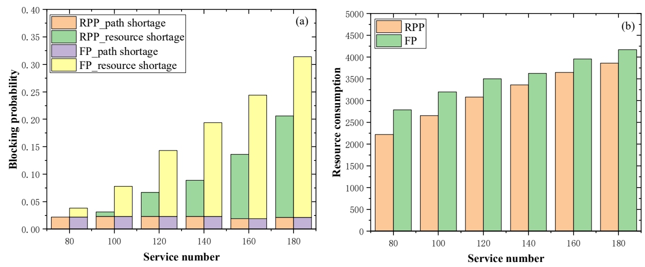

Fig. 4. Performance comparison against single failure.

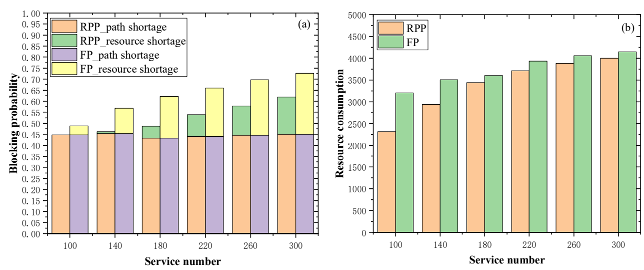

Fig. 5. Performance comparison against double failures

The comparison of blocking probability and spectrum resources consumption between the RPP scheme and FP scheme under double failures is shown in Fig. 5. Similar as in Fig. 4 (a), the blocking probability caused by resource shortage of RPP scheme increases more slowly than the FP scheme as the number of services increases in Fig. 5 (a). However, when blocking caused by resource shortage happens, the number of requests is higher than under single failure. The reason for this result is that the blocking probability caused by path shortage is high, about 40.4%, because three disjoint paths cannot be calculated for a large number of requests. When service number is less than 140, there is no blocking due to resource shortage in the RPP scheme. While when service number reaches 100, blocking occurs due to resource shortage in the FP scheme. When service number is same, the blocking probability caused by resource shortage in our scheme is about 12% less than the FP scheme. As shown in Fig. 5 (b), with the increase in the number of services, the resource consumption of both schemes continuously increases. And when the number of services is greater than 180, the difference of spectrum consumption between the two schemes decreases, which is because the RPP scheme serves more services than the FP scheme due to its lower blocking probability.

V. CONCLUSION

We propose an RPP scheme for degraded services to accommodate as many services as possible in EONs. It adaptively decides the number of disjoint paths against failures and low-level data proportion according to the available spectrum resource while ensuring the reliable transmission of high-level data. The simulation results confirm that the RPP scheme serves more services than the FP scheme in same network environment, which reflects the proposed scheme can utilize network resources more efficiently than the FP scheme.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China (62171050, 62125103), the Fund of State Key Laboratory of Computer Architecture (CARCH201906), and Open Fund of State Key Laboratory of Information Photonics and Optical Communications of BUPT (IPOC2021ZT15).

REFERENCES

[1] S. S. Savas, et.al, “Network adaptability to disaster disruptions by exploiting degraded-service tolerance,” IEEE Communications Magazine, vol. 52, no. 12, pp. 58-65, December 2014.

[2] S. Huang, M. Xia, C. U. Martel and B. Mukherjee, "A Multistate Multipath Provisioning Scheme for Differentiated Failures in Telecom Mesh Networks," Journal of Lightwave Technology, vol. 28, no. 11, pp. 1585-1596, June 2010.

[3] C. S. K. Vadrevu, et al., “Degraded service provisioning in mixed-line-rate WDM backbone networks using multipath routing,” IEEE/ACM Trans. Netw., vol. 22, no. 3, pp. 840–849, Jun. 2014.

[4] L. Zhang, et al., “End-to-End Multi-Connection Partial Protection Scheme Ensuring Tolerable QoS in Optical and Wireless Converged Networks,” in Proc. ACP, Shanghai, China, Oct. 2021.

[5] G. Le et al., “Reliable Provisioning with Degraded Service Using Multipath Routing from Multiple Data Centers in Optical Metro Networks,” IEEE Transactions on Network and Service Management, 2023, DOI 10.1109/TNSM.2023.3235354.

[6] A. Santos, et al., “An Online Strategy for Service Degradation with Proportional QoS in Elastic Optical Networks,” in Proc. ICC, Kansas City, MO, USA, May 2018.

[7] Z. Zhong, et al., “On QoS-assured degraded provisioning in service differentiated multi-layer elastic optical networks,” in Proc. GLOBECOM, Washington, DC, USA, Dec. 2016.

[8] J. Li, K. Xue, J. Liu, Y. Zhang and Y. Fang, “An ICN/SDN-Based Network Architecture and Efficient Content Retrieval for Future Satellite-Terrestrial Integrated Networks,” IEEE Network, vol. 34, no. 1, pp. 188-195, January/February 2020.

[9] Z. Zhu, et al., “Impairment- and Splitting-Aware Cloud-Ready Multicast Provisioning in Elastic Optical Networks,” IEEE/ACM Transactions on Networking, vol. 25, no. 2, pp. 1220-1234, 2017.