Availability-Guaranteed Differentiated Provisioning in Integrated Satellite-Terrestrial Optical Networks

Lu Zhang,[1, 2] Xin Li,[1, *] Massimo Tornatore,[2] Jingjie Xin,[1] and Shanguo Huang[1]
[1] School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, 10086, China
[2] Department of Electronics Information and Bioengineering, Politecnico di Milano, 20133, Milan, Italy
** xinli@bupt.edu.cn*

Abstract: This paper investigates differentiated provisioning in integrated satellite-terrestrial optical networks. Two connection availability models are developed considering network dynamic nature. Two availability-guaranteed differentiated provisioning algorithms are proposed. Their effectiveness is verified by numerical results.

1. Introduction

Integrated satellite-terrestrial optical networks (ISTONs) have shown great potential to provide global coverage and seamless connectivity [1]. An ISTON is composed of a satellite optical network (SON) and a terrestrial optical network (TON). The availability of a terrestrial fiber link (TFL) in the TON can be affected by failures, but it is generally considered static. While in the SON, due to the movement of satellites, the visibility and distance between satellites and between satellite and ground station (GS) change over time. Hence, the availability of an inter-satellites link (ISLs) or a satellite-terrestrial link (STL) in an ISTON is dynamic, and it can be influenced by atmospheric attenuation, atmospheric turbulence [2], accuracy of acquisition, tracking and pointing (ATP), interference from stars or other satellites, space debris, etc. To guarantee Quality of Service (QoS), it is necessary to establish connections that meet given availability requirements, typically specified in service level agreements (SLAs) stipulated by service providers and customers [3], and which can be different for different services. Sometimes serving a connection using a single path cannot meet availability requirement, and one or more backup paths are needed to improve a connection's availability [4].

Considering ISTON's dynamic nature, we propose two new protection strategies, a Multi-PathSets-based Protection (MPP) based on full path protection, and a Time-Slice-based Protection (TSP) based on partial path protection. For these two strategies, we devise two connection availability models and two Availability-guaranteed Differentiated Provisioning (ADP) algorithms, in which a connection can be protected adaptively based on availability requirement. Numerical results show that our proposed algorithms outperform 1+1 dedicated protection. To the best of our knowledge, this is the first study to investigate ADP in ISTONs.

2. Connection Availability Models for Multi-PathSets-based Protection and Time-Slice-based Protection

An ISTON can be modeled as a dynamic graph over time $G(t) = \langle V, E(t)\rangle$, where V is the node set, including satellite nodes and terrestrial nodes, and $E(t)$ is the dynamic link set at time t, including ISLs, STLs and TFLs. The satellite movement period can be divided into a series of discrete time slices. As the duration of each time slice is very short (e.g., time-slice duration in Iridium system is 2.5 minutes), link availability is considered to be constant during a time slice. The availability of a (unprotected or protected) connection depends on the availabilities of the paths that constitute the connection, and a path is considered available only when all links in the path are available [4]. A service request can be modeled as $r(s, d, t_a, T_h, Ar)$, where, s is the source node, d is the destination node, t_a is the arrival time, T_h is the holding time, Ar is the required availability. When the availability of a connection is not less than Ar, the service can be successfully provisioned by the connection; otherwise, the service is blocked.

To guarantee availability requirements, we propose two protection strategies, MPP and TSP. Since the ISTON topology changes in different time slices, the paths calculated in different time slices may be different (or, even if the paths calculated in different time slices are same, their availabilities may be different as link availability changes in different time slices). In MPP, we denote a set of N paths (one path for each time slice during the holding time of a service) as a *PathSet* $P=\{p^1,\cdots,p^n,\cdots,p^N\}$, where N is the total number of time slices during the holding time, and p^n is a path in n^{th} time slice. MPP is based on full path protection, hence a service is provisioned by setting up a connection, where a primary *PathSet* is protected by one or more backup *PathSets*, and the number of backup paths in all time slices are same. Unlike MPP, TSP is based on partial path protection. In TSP, we consider that a connection is split into N sub-connections, i.e., one independent connection for each time slice. Each sub-connection in each time slice can be protected by a different number of backup paths (in particular, some sub-connections may not have a backup path).

Connection availability model for MPP: In MPP, a connection is available when at least one *PathSet* is available. A *PathSet* is available only when all paths that constitute the *PathSet* are available. Hence, the availability of a *PathSet* can be formalized as Eq. (1), while the overall connection availability can be formalized as Eq. (2), where P_0 is the primary *PathSet*, and P_i is i^{th} backup *PathSet* ($1 \le i \le K$), K is the number of backup *PathSets*.

Connection availability model for TSP: In TSP, a connection is available only when all sub-connections are available. The set of the availabilities of all sub-connections is denoted as $A_{SC} = \{A^1, \cdots, A^n, \cdots A^N\}$, where A^n is the availability of the n^{th} sub-connection. Hence, the availability A_C of a connection using TSP can be described by Eq. (3). The availability of a sub-connection is given by the availabilities of the paths calculated in the corresponding time slice and can be calculated by Eq. (4) [4], where $A_{p_i^n}$ is the availability of i^{th} path in n^{th} time slice ($0 \le i \le k^n$), p_0^n is the primary path in n^{th} sub-connection, k^n is the number of backup paths in n^{th} time slice.

$$A_P = \prod_{p^n \in P} A_{p^n} \quad (1) \qquad A_C = \sum_{i=0}^{K}\left(A_{P_i} \times \prod_{j=0}^{i-1}(1 - A_{P_j}) \right) (2) \qquad A_C = \prod_{A^n \in A_{SC}} A^n \quad (3) \qquad A^n = \sum_{i=0}^{k}\left(A_{p_i^n} \times \prod_{j=0}^{i-1}(1 - A_{p_j^n}) \right) (4)$$

Using the connection shown in Fig.1 as an example, we calculate the availability of the connection according to the above two models. In the case of MPP, the connection can be considered to consist of two *PathSets* P_0 and P_1, where the primary path p_0^1 in time slice ts_1 and the primary path p_0^2 in time slice ts_2 form the primary *PathSet* P_0, the backup path p_1^1 in time slice ts_1 and the backup path p_1^2 in time slice ts_2 form the backup *PathSet* P_1, and the availability of the connection is calculated by Eq. (5). In the case of TSP, the connection can be considered to consist of two sub-connections, where p_0^1 and p_1^1 form the first sub-connection sc_1, p_0^2 and p_1^2 form the second sub-connection sc_2, the connection availability is calculated by Eq. (6). Note that Fig.1 is just an example of application of Eq. (1-4), while, typically, MPP and TSP lead to different backup allocations.

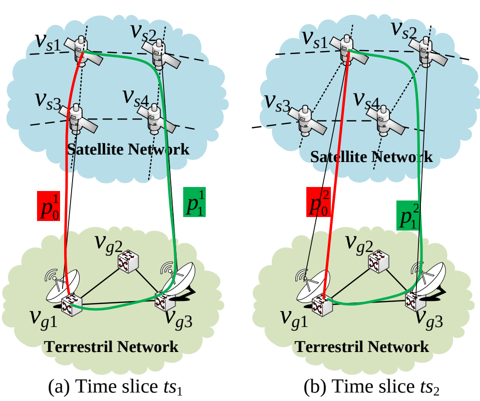

Fig. 1. A connection for a service with two time slices.

$$A_C = A_{P_0} + (1 - A_{P_0}) \times A_{P_1} = A_{p_0^1} A_{p_0^2} + (1 - A_{p_0^1} A_{p_0^2}) \times A_{p_1^1} A_{p_1^2} \quad (5)$$

$$A_C = A^1 A^2 = \left[A_{p_0^1} + (1 - A_{p_0^1}) A_{p_1^1} \right] \times \left[A_{p_0^2} + (1 - A_{p_0^2}) A_{p_1^2} \right] \quad (6)$$

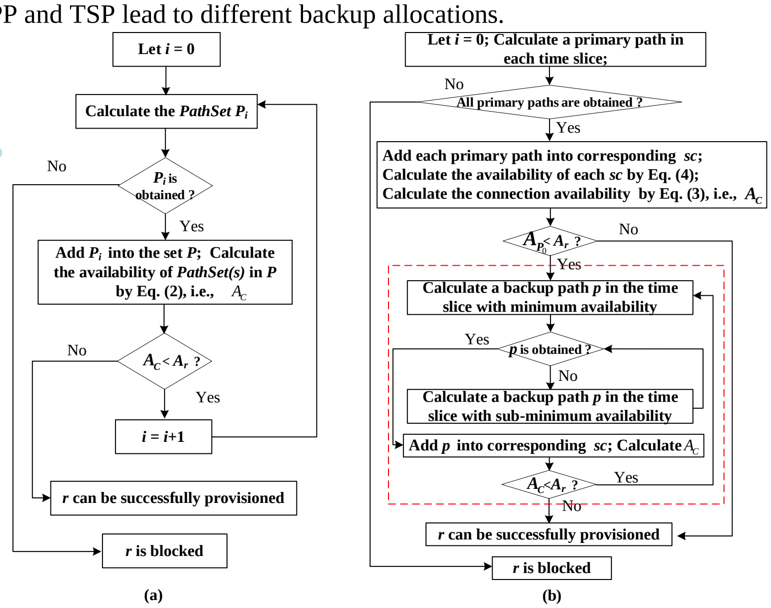

Fig. 2. (a) MP-ADP algorithm; (b) TS-ADP algorithm.

3. Availability-Guaranteed Differentiated Provisioning Algorithms in ISTONs

Our goal is establishing connections for as many services as possible while meeting their availability requirements. To meet this goal, we propose two algorithms, a MPP-based ADP (MP-ADP) algorithm and a TSP-based ADP (TS-ADP) algorithm. The general idea of the two algorithms is that we first calculate a primary path with greatest availability in each time slice, then, if the primary path does not meet the availability requirement, we add backup path(s) until availability requirement is guaranteed or no new backup path can be calculated due to network capacity. The main difference between these two algorithms resides in the way the backup path(s) are calculated. To consider path priority in terms of availability, we set the link weight as a function of link availability, $C_l = -\log(A_l)$ [4], so that the shortest-path is the path with the greatest availability. The details of the two algorithms are as follows (see Fig. 2 for their flow charts).

MP-ADP algorithm: Based on the N time-slice topologies, the primary *PathSet* P_0 can be firstly obtained by calculating a primary path with greatest availability in each time slice. If the availability of the primary *PathSet* does not meet the availability requirement, another backup *PathSet(s)* is calculated. Backup *PathSets* are added one by one, until the overall connection availability is not lower than A_r. Finally, resources are allocated to all *PathSet(s)*.

TS-ADP algorithm: Same as in MP-ADP algorithm, the primary path with greatest availability is calculated in each time slice. Then, since $0 \le A^n \le 1$ and A_c is the product of the availabilities of sub-connections, A_c is smaller than the minimum in A_{sc}, so, to use as few backup paths as possible to meet availability requirements, the backup paths are calculated in the time slice with smallest availability iteratively. If the algorithm fails to calculate a new backup path in the time slice with the smallest availability of sub-connections, it then calculates a backup path in the time slice with next smallest availability of sub-connections. Finally, if the connection meets the availability requirement, then we allocate resources to all paths for service provisioning. Otherwise, the service is blocked.

4. Simulation Results

In our simulation, the ISTON is composed by the NeLS constellation [5] and by a terrestrial network of 24 nodes. A GS is located at each terrestrial node. The System Tool Kit simulator is used to obtain visibility between satellites and between satellites and GSs. Each satellite or GS establishes links with up to 4 satellites. The number of wavelengths for TFL, ISL and STL is set to 80, 16, and 8 respectively. Wavelength allocation is based on first-fit algorithm. Source and destination nodes are chosen randomly. The bandwidth requested by each service is set as one wavelength. The arrival of service requests in the simulation obeys the Poisson distribution. The holding time is subject to the exponential distribution with a mean of 100 seconds. The number of service requests is set as 2000. All results are plotted by statistically averaging 10 simulation runs.

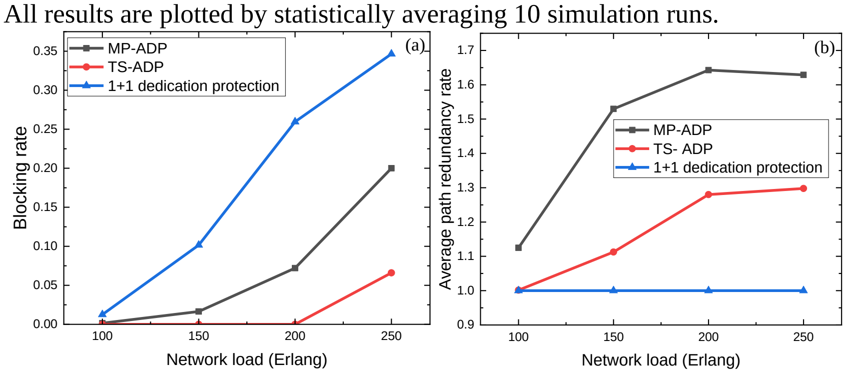

Fig. 4. Comparison for availability requirement 0.9999, (a) blocking rate, (b) APR.

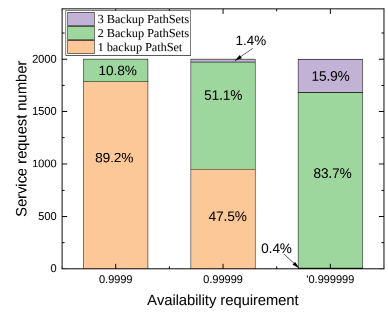

Fig. 5. Distribution of backup *PathSets* number required by connections for differentiated services.

In Fig.4, we compare the two proposed algorithms with traditional 1+1 dedicated protection (i.e., one backup *PathSet* is always provided for protection) for services with availability requirement 0.9999, in terms of blocking rate (blocking can be due to inability of guaranteeing the availability requirement as well as to lack of capacity) and average path redundancy ratio of connections (APR), which is the ratio of the number of all backup paths to the number of primary paths. For increasing load, blocking rate and APR increase. 1+1 dedicated protection has a higher blocking rate than ADP algorithms as it cannot guarantee the availability requirements. Since TS-ADP consumes less wavelength resources (thanks to partial path protection), its blocking rate and APR are significantly lower than MP-ADP. Main blocking reason in MP-ADP and TS-ADP is that it is not possible to find enough paths due to limited network capacity. The simulation results for services with other availability requirements are similar to 0.9999, and are not presented due to space limitations. In particular, Fig. 5 shows, for the MP-ADP algorithm, the amount of backup *PathSets* required by connections with three availability classes {0.9999, 0.99999, 0.999999}, when the network load is 50 Erlang (without blocking). It can be seen that, for higher the availability requirement, the proportion of services with large number of backup *PathSets* becomes higher.

5. Conclusion

Based respectively on full path protection and partial path protection, we devise two connection availability models and two algorithms, MP-ADP and TS-ADP, in which a connection can be protected adaptively based on availability requirement. Numerical results show ADP algorithms performs better on blocking, compared to traditional 1+1 dedicated protection. Especially, TS-ADP has lower blocking rate and redundancy ratio than MP-ADP.

Acknowledgement

This work was supported in part by National Natural Science Foundation of China (Nos. 62171050, 62125103, 61821001), Fundamental Research Funds for the Central Universities (2023PY08), and Open Fund of State Key Laboratory of Information Photonics and Optical Communications of BUPT (IPOC2021ZT15).

References

[1] J. Liu, et.al., IEEE Commun. Surveys & Tutorials **20(4)**, 2714–2741 (2018).
[2] H. Kotake, et.al., Opt. Express **31,** 21351-21366 (2023).
[3] B. Kantarci, et.al., J. Lightw. Technol. **27(20)**, 4595-4602 (2009).
[4] L. Song, et.al., IEEE J. Selected Areas Commun., **26(6)**, 84-91 (2008).
[5] X. Sun, et.al., J. Lightw. Technol. **38(8)**, 2106-2113 (2020).