Minimize Sub-carrier Reallocation in Elastic Optical Path Networks using Traffic Prediction

Sunny Shakya, Yang Wang, Xiaojun Cao
Department of Computer Science
Georgia State University
Atlanta, Georgia 30303, USA
Email: sshakya1@student.gsu.edu, ywang, cao@cs.gsu.edu

Zilong Ye, Chunming Qiao
Department of Science and Engineering
SUNY-Buffalo
Buffalo, NY 14260, USA
Email: zilongye@buffalo.edu, qiao@computer.org

Abstract—Spectrum-sLICed Elastic optical path (SLICE) networks enable elastic and flexible allocation of spectral resources. SLICE networks distribute data on a number of sub-carriers overlapped in frequency domain to provide efficient sub-wavelength and super-wavelength traffic accommodation. In SLICE networks, a routing and spectrum allocation algorithm assigns a spectrum path to any demand with just enough contiguous sub-carriers while following the sub-carrier consecutiveness and spectrum-continuity constraints. In this paper, we propose novel sub-carrier allocation algorithms that employ the proposed Interference Graph technique to assign sub-carriers to a spectrum path based on the historic traffic profile. These algorithms try to achieve minimal disruptions to the live connections while minimizing the blocking probability. Simulation results show that the proposed schemes can effectively accommodate the dynamic traffic while minimizing the network reconfiguration cost in SLICE networks, with or without the traffic prediction.

I. INTRODUCTION

Wavelength Division multiplexing (WDM) networks divide the fiber frequencies into channels with a fixed-size spectral bandwidth. This rigid and coarse resource allocation leads to a poor utilization of the spectrum resources when the traffic demand between any two optical nodes vary significantly. Besides this, the guardband between two wavelengths can be in the order of wavelengths, which results in a wastage of the expensive spectral resources [1], [2]. To make more efficient and elastic use of the spectrum, Spectrum-sLICed Elastic optical path (SLICE) networks are proposed [2]. SLICE networks employ the orthogonal frequency division multiplexing (OFDM) technology such that the adjacent spectrum bands can overlap each other for better spectrum utilization [2]–[5].

Unlike the wavelength routed networks (WRN), the SLICE networks carry data on a bunch of sub-carriers. Each sub-carrier has a much lower granularity than a single wavelength, which makes it one of the best candidates to optically accommodate sub-wavelength traffic [3]. The super-wavelength traffic can be accommodated by combining several sub-carriers together without placing guard frequencies (or carriers) in between. However, in SLICE networks, a routing and spectrum assignment (RSA) algorithm has to consider the spectrum continuity and sub-carrier consecutiveness constraints in the process of assigning a spectrum path (SP) to any incoming connection [4], [5]. The spectrum continuity constraint, similar to the wavelength continuity constraint in WRN, requires continuous availability of sub-carriers along a routing path (if no frequency converter is provided). The sub-carrier consecutiveness constraint requires that the sub-carriers assigned to any SP should be consecutive in spectrum domain to take full advantages of the OFDM technology. For the case with static traffic, the RSA problem is shown to be NP-Complete [5]. The RSA problem in the dynamic traffic can also be very challenging due to the random traffic arrival/departure and the fluctuation of the traffic demands over time. When the traffic demand fluctuates over time, a straightforward solution is to reconfigure the spectrum assignment periodically (or on demand) by doing the RSA process from scratch based on the current traffic demands. This approach, however, may require shifting of many (if not all) existing sub-carriers allocation, reconfiguration of some network elements (e.g. transponders and receivers), and interruptions to the live connections, which can be very expensive for the network providers.

In the literature, some work has been done to assign sub-carriers to connections in networks with dynamic traffic demands. A policy to allocate sub-carriers to requests whose rate fluctuates with time was proposed in [6] while an adaptive spectrum allocation algorithm for time-dependent traffic was investigated to share neighboring sub-carriers in [7]. The work in [8] studied a Multi-Hour Routing and Spectrum Allocation (MHRSA) optimization problem and the authors in [9] explored the periodic behavior of internet traffic in WDM networks with traffic grooming. However, network reconfiguration and traffic prediction in SLICE or FlexGrid networks has not been considered in the existing work. In this paper, we study how to effectively accommodate time-varying traffic with minimum network reconfiguration in SLICE networks.

The remainder of this paper is organized as follows. Section II presents the time-varying spectrum allocation problem in SLICE networks while Section III describes three spectrum allocation strategies for time-varying traffic. In Section IV, we propose the minimum reconfiguration spectrum allocation algorithm. Section V shows the performance evaluation of the proposed schemes and Section IV concludes this work.

II. TIME-VARYING TRAFFIC IN SLICE NETWORKS

It has been observed that although the real-world data traffic may fluctuate randomly in a very short period of time (e.g.
within one hour), such a change follows a similar (or periodic) pattern over a long period of time (e.g., a day or a month) [10], [11]. More specifically, the Internet traffic exhibits a similar pattern every 24 hours [12]. This is largely due to the habitual usage of the network users. The similar or periodic pattern over a long period has been referred to as time-varying traffic [9]. For time-varying traffic, we can assume that the traffic pattern repeats after a certain time period (P). For example in Fig. 1, the traffic pattern repeats in P time period and we divide P into n fixed length time slots (e.g. one hour time slots). To handle time-varying traffic, we assume a slotted control plane in SLICE networks where the traffic demand between each node-pair is almost constant within each time slot and the highest intensity of the flow is used as the traffic demand. The significant traffic change only happens at the boundary of the time slot. We further assume that the traffic requests for the next time unit say \( t_{i+1} \) are given before the end of the current time slot \( t_i \). In addition, at time slot \( t_i \), the traffic requests for the time slot \( t_{i+2} \) are not precisely known.

Consider a network as \( G(V, E, N) \) where \( V \) represents the set of nodes, \( E \) represents the set of fiber links between nodes in \( V \) and \( N \) is the maximum number of sub-carriers in each fiber. Then, we can define the Time-varying Spectrum Allocation (TSA) problem in SLICE networks as follows:

**Time-varying Spectrum Allocation (TSA) problem:** Given a set of spectrum path assignments \( D_{t-1} = \{ f_{t-1} \} \) on a network \( G(V, E, N) \), and a set of demands for the next time slot \( D_t = \{ f_t \} \), is it possible to accommodate all the flows in demand \( D_t \) with minimum interruptions to existing traffic flows?

**Theorem 1.** The time-varying spectrum allocation problem (TSA) in SLICE networks is NP-complete.

**Proof:** The TSA problem consists of an instance of static spectrum allocation problem in the SLICE networks. The static spectrum allocation problem has been shown to be NP-Complete [5] via its connection with the static lightpath establishment (SLE) problem in WRN. Since SLE problem in WRN is known to be NP-Complete [13], the TSA problem in SLICE networks is also NP-Complete.

Fig. 2 shows a linear network with the existing spectrum assignments between node-pairs at time slot \( t_0 \). The expected traffic demands between the node-pairs at time slots \( t_1 \) and the predicted traffic demands at time slot \( t_2 \) are given in Table I. If we denote the traffic demand of a given node-pair at time slot \( t_i \) by \( D_{t_i} \), then we can use the following equation to calculate the traffic change between a node-pair from time slot \( t_i \) to time slot \( t_{i+1} \)

\[
\Delta_{i+1} = D_{t_{i+1}} - D_{t_i}
\]

The value of \( \Delta_{i+1} \) can be either negative, positive or zero. The negative value of \( \Delta_{i+1} \) indicates the decrease in traffic demand. This may create vacant or unused sub-carriers but does not require reconfiguration of network elements such as transponders and receivers. The positive value indicates the increase in traffic demand. If the sub-carriers contiguous to this demand are occupied, sub-carrier reallocation and network reconfiguration have to be carried out resulting in interruptions to some live connections. Moreover, if there are not enough empty sub-carriers to relocate this demand, such demand could be dropped or blocked. However, if there are vacant sub-carriers adjacent to this demand, then we may reduce the reconfiguration overhead by simply stretching the original SP instead of reallocating this SP to a new set of sub-carriers. For example, in Fig. 2, node-pair (A, D) uses sub-carriers \( S_5 \) and \( S_4 \) at time slot \( t_0 \). At time slot \( t_1 \), if the sub-carrier \( S_5 \) is available to expand for node-pair (A, D) to 3 sub-carriers, then node-pair (A, D) does not have to be assigned with a different set of sub-carriers. In Fig. 2, the sub-carrier \( S_5 \) is occupied by the SP between node-pair (A, E). Hence, in order to free up the sub-carrier \( S_5 \) for node-pair (A, D), the SP for node-pair (A, E) has to be moved from sub-carriers \( S_5, S_6 \) to the different set of sub-carriers. If there are not enough sub-carriers to relocate SP for node-pair (A, E), then the connection between node-pair (A, D) would have to be blocked.

Reallocation of connections from one set of sub-carriers to a different set of sub-carriers may be very costly to network providers as it involves re-tuning optical lasers and may also require additional resources such as transponders or bandwidth-variable cross connects. Compared to the case of...
adapting totally disjoint sub-carriers, the reconfiguration cost for stretching the sub-carriers to vacant adjacent sub-carriers may be negligible since the connection of the traffic demand is not fully interrupted and it may not require additional resources. Moreover, when the traffic demand of a node-pair decreases, the reconfiguration involves turning-off operations of network elements, whose cost is also negligible. Hence, our objective in this paper is to assign sub-carriers to time-varying traffic while minimizing the network reconfiguration.

III. SPECTRUM ALLOCATION FOR TIME-VARYING TRAFFIC

As discussed in Section II, TSA problem is an NP-complete problem. Hence, it is difficult to obtain an optimal spectrum assignment. In this section, we present three heuristic algorithms to assign sub-carriers to time-varying traffic, namely, Spectrum Allocation from Scratch (SAS), Planning-ahead Spectrum allocation (PAS) and Predicative Planning-ahead Spectrum allocation (PPAS). SAS solely considers the traffic demands for the spectrum assignment while PAS and PPAS also consider current state of the spectrum assignments in addition to the traffic demands.

A. Spectrum Allocation from Scratch (SAS)

Spectrum Allocation from Scratch (SAS) adopts a straightforward method to allocate sub-carriers to each SP based on the current traffic demands. At the end of time slot $t_i$, the traffic demands for next time slot $t_{i+1}$ are given in a set $TD(t_{i+1}, s, d)$ where $TD(t_{i+1}, s, d)$ is the number of sub-carriers requested by node-pair $(s, d)$ at time slot $t_{i+1}$. SAS only considers the traffic demands at time slot $t_{i+1}$ in the process of spectrum allocation as shown in Algorithm 1. SAS is a naive approach and may result in frequent setup and tear-down of connections due to the traffic fluctuations.

B. Planning-ahead Spectrum allocation (PAS)

Planning-ahead Spectrum allocation (PAS) considers two factors for the spectrum assignment between node-pairs $(s, d)$: $SA(t_i, s, d)$ and $TD(t_{i+1}, s, d)$ where the former is the spectrum assignment at time slot $t_i$ and the latter is the traffic demand at time slot $t_{i+1}$. For example, in Fig. 2, the node-pairs (A, D), (D, E) and (A, E) are assigned with 2 sub-carriers at time slot $t_0$. At time slot $t_1$, the traffic demands between node pair (A, D) and (D, E) is 3 sub-carriers (i.e. $\Delta_1 = 1$). Hence, to minimize the network reconfiguration cost, PAS may assign node pair (A, E) with sub-carriers $S_6, S_7$ as shown in Fig. 3. This allows the $SP$ between node pairs (A, D) and (D, E) to utilize the consecutive sub-carrier $S_5$ during time slot $t_1$. Hence, in this scheme, only $SP$ assigned to node-pair (A, E) has to be reallocated while $SP$s assigned to node-pair (A, D) and (D, E) can operate without network interruptions. Since PAS needs a method to determine the minimum number of $SP$s that need to be reallocated, the detailed steps for PAS are given in Section IV.

Algorithm 1 Spectrum Allocation from Scratch (SAS)

**Input:** $\{TD(t_{i+1}, s, d)\}$

**Output:** $\{SA(t_{i+1}, s, d)\}$

1: Sort the traffic demands in the descending order of sub-carriers requested
2: while there are requests in $TD(t_{i+1}, s, d)$ do
3: Find the shortest route for $(s, d)$
4: Use First-fit scheme to find $TD(t_{i+1}, s, d)$ sub-carriers along the shortest route
5: if found available sub-carriers for $(s, d)$ then
6: Allocate those sub-carriers to $(s, d)$
7: else
8: Block the request for $(s, d)$
9: end if
10: Update $SA(t_{i+1}, s, d)$ and $TD(t_{i+1}, s, d)$
11: end while

C. Predicative Planning-ahead Spectrum allocation (PPAS)

Predictive Planning-ahead spectrum allocation (PPAS) takes advantage of the periodic behavior of the internet traffic...
demands. PPAS considers three factors during the spectrum assignment process: the spectrum assignments at time slot $t_i$ given by $SA(t_i, s, d)$, the traffic demands at time slot $t_{i+1}$ given by $TD(t_{i+1}, s, d)$ and the predicted traffic demands at time slot $t_{i+2}$ given by $TP(t_{i+2}, s, d)$. For example in Fig. 4(a), the spectrum assignments for time slot $t_1$ happen at the end of time slot $t_0$ where $SA(t_0, s, d)$ and $TD(t_1, s, d)$ are known while $TP(t_2, s, d)$ can be predicted using the internet traffic history. In PAS as shown in Fig. 3(a), one needs to further reconfigure the network at the end of time slot $t_1$ to accommodate 4 sub-carriers for node-pair (A, D). In contrast, if the projected traffic demand at time slot $t_2$ is also considered, the proposed PPAS can shift the SP between node-pair (A, E) to sub-carriers $S_8, S_9$ during time slot $t_1$ while reserving sub-carrier $S_3$ for node-pairs (A, C) and (C, E). Similarly, sub-carrier $S_7$ can be virtually reserved for node-pair (A, D) to accommodate increase in traffic demands at time slot $t_2$, which can further reduce the overall network reconfiguration cost. The detailed steps for PPAS are given in next section.

IV. MINIMUM RECONFIGURATION SPECTRUM ALLOCATION ALGORITHM

We now describe the Minimum Reconfiguration Spectrum Allocation algorithm (MRA) to minimize the total reconfiguration cost over time, which will be adopted in PAS and PPAS. MRA consists of two steps. In Step 1, we transform the spectrum allocation information for current time slot $t_i$ into an auxiliary Interference Graph ($IG$). To construct the $IG$, we create a vertex for each $SP$ at the current time slot $t_i$ and connect two vertices (say $i$ and $j$) if sub-carriers of $j$ has to be reallocated for the accommodation of vertex $i$. In Step 2, we will find the Maximum Independent Set (MIS) in the $IG$. Finding the MIS in a graph is a well known NP-Complete problem. There is a parallel version of algorithm [14] which can achieve $O((\log n)^3)$ with a space complexity of $O(n/\log n)^3$. For the simplicity of the implementation, we use the algorithm in [15] to find the MIS in an $IG$.

Algorithm 2 Planning-ahead Spectrum (PAS)

Input: $\{SA(t_i, s, d)\}, TD\{(t_{i+1}, s, d)\}$
Output: $\{SA(t_{i+1}, s, d)\}$

1: Sort the traffic demands in the descending order of sub-carriers requested
2: while there are requests in $TD(t_{i+1}, s, d)$ do
3: Find the shortest route for $(s, d)$
4: Record the $SP$s to be reallocated using MRA in Set $R$
5: $relocation \leftarrow true$
6: while Set $R$ is not empty do
7: Relocate $SP$s in $R$
8: if $SP$ relocation fails then
9: Revert the spectrum assignment for $SP$ in $R$
10: $relocation \leftarrow false$
11: break;
12: end if
13: end while
14: if $relocation == true$ then
15: Stretch the sub-carriers of $(s, d)$ to $TD(t_{i+1}, s, d)$ sub-carriers
16: end if
17: if $relocation == false$ then
18: Allocate $TD(t_{i+1}, s, d)$ sub-carriers to $(s, d)$ using First-fit scheme
19: Block the request of $(s, d)$ if not enough sub-carriers
20: end if
21: Update $SA(t_{i+1}, s, d)$ and $TD(t_{i+1}, s, d)$
22: end while

A. Interference graph for PAS

For the spectrum assignment given in Fig. 3, the $IG$ for time slots $t_1$ and $t_2$ are shown in Fig. 5(a) and (b), respectively. At time slot $t_0$, there are 5 $SP$s. Hence, in the $IG$ for time slot $t_1$, there are 5 vertices - $AC, AD, DE, CE$ and $AE$ and there is an edge between vertices $AD-AE$ since $SP$ for node-pair (A, E) has to be moved to provide sub-carrier $S_5$ for node-pairs (A, D) and (D, E). With the $IG$, we can then calculate the reconfiguration cost by finding the MIS among the vertexes. The MIS of the $IG$ for time slot $t_1$ consists of shaded nodes $AC, AD, DE$ and $CE$ as shown in Fig. 5(a). Similarly, for time slot $t_2$, the size of the MIS is 3. Hence, the reconfiguration cost for time slot $t_1$ is $5 - 4 = 1$ since only one $SP$ needs to be interrupted and relocated (i.e. $SP$ for node-pair (A, E)). Therefore, the total reconfiguration cost for PAS for time slots $t_1$ and $t_2$ is $(5 - 4) + (5 - 3) = 3$. The detailed steps for PAS are given in Algorithm 2.

B. Interference graph for PPAS

The $IG$ of PPAS for time slot $t_{i+1}$ considers the predicted traffic demands at time slot $t_{i+2}$ in addition to the known traffic demands at time slot $t_{i+1}$ and the spectrum assignments for time slot $t_i$. For example, the $IG$ for time slot $t_1$ and $t_2$ in Fig. 3 are shown in Fig. 6(a) and (b), respectively when PPAS is employed. The detailed steps for PPAS are given

Fig. 5. PAS interference graph for (a) time slot $t_1$ (b) time slot $t_2$

Fig. 6. PPAS interference graph for (a) time slot $t_1$ (b) time slot $t_2$
in Algorithm 3. PPAS may involve reallocation of some of the SPs as given in Step 7 of Algorithm 3. Reallocation of SPs can cause network disruption in SLICE networks. Hence, to minimize disruption, we use a make-before-break scheme. Using this scheme, SPs are first established on different set of sub-carriers before releasing the original SP which can minimize the disruption time [16].

Algorithm 3 Predictive Planning-ahead Spectrum (PPAS)

Input: \{SA(t_i,s,d), TD\{(t_{i+1},s,d)\} and \{TP(t_{i+2},s,d)\}
Output: \{SA(t_{i+1},s,d)\}

1: Sort the traffic demands in the descending order of sub-carriers requested
2: \textbf{while} there are requests in \textit{TD}(t_{i+1}, s, d) \textbf{do}
3: \hspace{1em} Find the shortest route for \textit{(s, d)}
4: \hspace{1em} Record the SPs to be reallocated using MRA in Set \textit{R}
5: \hspace{1em} relocation \leftarrow true
6: \hspace{1em} \textbf{while} Set \textit{R} is not empty \textbf{do}
7: \hspace{2em} Relocate SPs in \textit{R}
8: \hspace{2em} \textbf{if} SP relocation fails \textbf{then}
9: \hspace{3em} Revert the spectrum assignment for SP in \textit{R}
10: \hspace{3em} relocation \leftarrow false
11: \hspace{3em} break;
12: \hspace{2em} \textbf{end if}
13: \hspace{1em} \textbf{end while}
14: \hspace{1em} \textbf{if} relocation \textit{==} true \textbf{then}
15: \hspace{2em} Stretch the sub-carriers of \textit{(s, d)} to \textit{max}\{(TP(t_{i+2}, s, d), TD(t_{i+1}, s, d))\} sub-carriers
16: \hspace{2em} \textbf{end if}
17: \hspace{1em} \textbf{if} relocation \textit{==} false \textbf{then}
18: \hspace{2em} Allocate \textit{TD}(t_{i+1}, s, d) sub-carriers to \textit{(s, d)} using First-fit scheme
19: \hspace{2em} Block the request of \textit{(s, d)} if not enough sub-carriers
20: \hspace{2em} \textbf{end if}
21: \hspace{1em} Update \textit{SA}(t_{i+1}, s, d) and \textit{TD}(t_{i+1}, s, d)
22: \textbf{end while}

V. PERFORMANCE EVALUATION

We evaluate the performance of SAS, PAS and PPAS in a 14-node NSFNET. We assume that the period \textit{(P)} is divided into 12 time slots. There are 50 sub-carriers in each fiber. The simulation is done in two phases. The connection requests arrive according to the Poisson process at a rate of \textit{λ} and the holding time of each connection conforms to a negative exponential distribution with parameter \textit{μ}. The demand for each node-pair is randomly generated between \textit{(1, 10)} in terms of the number of sub-carriers.

Fig. 7(a) shows the average number of network reconfiguration when increasing \textit{λ}. From the figure, it is clear that both PAS and PPAS outperform the SAS by a considerable margin. This is because SAS assigns sub-carriers to the traffic demands without considering the existing spectrum assignments. Hence, in order to accommodate increasing traffic demands, SAS may frequently tear down existing connections due to the lack of enough sub-carriers to expand and move these connections to a different set of sub-carriers. This may require the reconfiguration of the whole spectrum assignment. Unlike SAS, PAS and PPAS consider the current spectrum assignments in addition to the traffic demands. In addition, PPAS also considers future traffic demands prediction to assign sub-carriers which results in even lower network reconfiguration cost.

In Fig. 7(b), we compare the blocking probability when increasing \textit{λ}. From the figure, one can see that SAS slightly outperforms PPAS and both are better than PAS in terms of the blocking probability. This is because SAS assigns sub-carriers to \textit{SPs} from scratch without considering the existing spectrum assignments. Hence, SAS makes better use of the spectrum. However, SAS does not consider the current spectrum assignments and \textit{SPs} may be frequently torn down and set up, increasing the number of disruptions to the live connections. To reduce the reconfiguration/disruption cost, PAS and PPAS may reserve some sub-carriers for \textit{SPs} and thus, results in (or causes) a higher blocking probability. When comparing with PPAS, PAS causes a higher number of reallocations, which indicates that more connections are moved to a different set of sub-carriers. After this moving/relocation, the previously assigned sub-carriers are now vacant and may or may not be sufficient to establish new connections. Hence, PAS leads to
more blocking than PPAS.

To simulate the real network traffic, we introduce noise by allowing the actual future traffic demands to fluctuate from the predicted traffic demands. In particular, the percentage of fluctuation from the predicted traffic demands considered in this paper are 0%, +/-10%, +/-25% and +/-50% which are denoted by PPAS-0%, PPAS-10%, PPAS-25% and PPAS-50%, respectively. In Fig.8, as the prediction error increases, the amount of network reconfiguration also increases, which indicates the performance degradation of PPAS. Clearly, the better traffic prediction indicates fewer network reconfiguration for PPAS.

VI. CONCLUSION

Efficient spectrum allocation is a key to obtain better utilization of network resources and to maintain a lower network cost in SLICE networks. In this paper, we have proposed novel spectrum assignment schemes based on Interference graph to assign sub-carriers to time-varying traffic, namely Planning-ahead Spectrum allocation (PAS) and Predicative Planning-ahead Spectrum allocation (PPAS). PPAS adopts a plan-ahead mechanism whereby the spectrum allocation process not only considers the traffic demands in the current time slot, but it also considers the spectrum assignment in previous time slot as well as the traffic prediction in future time slots. Using this approach, PPAS can prevent unnecessary network reconfiguration in future time slots. Simulation results have shown that PPAS can achieve much less network reconfiguration while maintaining a low blocking probability for the time-varying traffic.

REFERENCES


