Historical Data Learning Based Dynamic LSP Routing for Overlay IP/MPLS over WDM Networks[☆]

Xiaojun Yu, Gaoxi Xiao*, Tee Hiang Cheng

School of Electric and Electronic Engineering, Nanyang Technological University, Singapore, 639798

Abstract

Overlay IP/MPLS over WDM network is a promising network architecture starting to gain wide deployments recently. A desirable feature of such a network is to achieve efficient routing with limited information exchanges between the IP/MPLS and the WDM layers. This paper studies dynamic label switched path (LSP) routing in the overlay IP/MPLS over WDM networks. To enhance network performance while maintaining its simplicity, we propose to learn from the historical data of lightpath setup costs maintained by the IP-layer integrated service provider (ISP) when making routing decisions. Using a novel historical data learning scheme for logical link cost estimation, we develop a new dynamic LSP routing method named Existing Link First (ELF) algorithm. Simulation results show that the proposed algorithm significantly outperforms the existing ones under different traffic loads, with either limited or unlimited numbers of optical ports. Effects of the number of candidate routes, add/drop ratio and the amount of historical data are also evaluated.

Keywords: Optical networks, IP over WDM, overlay model, label switched path, dynamic routing, historical data learning.

[☆] This research was supported in part by A*STAR SERC Grant (Award No. 0721010019).

[*] Corresponding author

Email addresses: `E070035@e.ntu.edu.sg` (Xiaojun Yu), `EGXXiao@ntu.edu.sg` (Gaoxi Xiao), `ETHCHENG@ntu.edu.sg` (Tee Hiang Cheng)

1. INTRODUCTION

Over the past few decades, Internet traffic has been growing exponentially, stimulating widespread deployments of wavelength-division-multiplexing (WDM) based optical networks [1, 2]. To handle the huge amount of traffic, it is widely believed that network control has to be simplified in the next-generation optical Internet [3]. Among the various emerging networking technologies, the overlay IP/MPLS over WDM network architecture is regarded as a promising candidate [4]. The recent developments and standardizations of some new technologies, such as Multi-Protocol Label Switching (MPLS) [5], Generalized MPLS (GMPLS) [6, 7], User Network Interface (UNI) [8] and path computation element (PCE) [9], etc., also help make such network architecture feasible to be deployed in the near future.

In IP/MPLS over WDM networks, IP routers with MPLS functions are called label switched routers (LSRs). They are interconnected by the optical core network and are used to groom or switch finer-grained label switched paths (LSPs). The optical layer network, which consists of photonic cross-connects (PXCs) and optical fiber links, provides dynamic point-to-point connectivity service to the IP/MPLS-layer network in the form of lightpaths. A lightpath may span several optical links, and if all PXCs have no wavelength conversion capability, it has to be established with the same wavelength along its route, known as the wavelength continuity constraint [10]. Since the granularity of an LSP is typically much smaller than that of a lightpath, several LSPs are usually groomed into a single lightpath. Once a lightpath is no longer used by any LSPs, it will be torn down and the wavelength used is released.

According to the different interconnection methods between IP/MPLS-layer and WDM-layer networks, three architectural alternatives have been proposed, namely overlay, augmented and peer models respectively [11]. The peer model adopts a unified control plane, and shares all the network information, e.g., the topology, routing and link state information, among all network nodes across its two layers; the overlay model has its IP and WDM layers controlled and managed independently, with limited information exchanges between the two layers for handling service requests. The augmented model aims to make a compromise between the first two models by sharing certain selected information between the two layers; however, there is still no consensus on what kind of information should be shared. Since IP-layer network and the optical transport network are usually owned by different operators, the overlay model is widely regarded as the most realistic one for near-term deployments. Extensive work has been done on this model [12, 13], and the recently demonstrated service oriented optical networks [14, 15] is a practical application of such an overlay architecture.

In this paper, we focus on the dynamic LSP routing problem in the overlay IP/MPLS over WDM networks. Although a number of algorithms have been proposed, none of them has considered the fact that a logical-layer ISP typically has the historical records of its own service requests that have been supported by the WDM-layer network. Such records, if properly used, can help an ISP make better routing decisions.

We propose a novel dynamic LSP routing algorithm for overlay IP/MPLS over WDM networks which utilizes the historical records. By developing a historical data learning scheme for logical link cost estimation and utilizing the K-loopless shortest path (KSP) algorithm [16] for the LSP routing, an algorithm named Existing-Link-First (ELF) is proposed. Simulation results show that the ELF algorithm outperforms the existing ones under different traffic loads, with either limited or unlimited number of optical ports in each node. Studies on the effects of the number of candidate routes, add/drop ratio, and the amount of historical data are also carried out.

The remainder of this paper is organized as follows. Section 2 presents a brief review of the existing algorithms for the overlay network model. Section 3 describes the system model and provides necessary definitions. Section 4 describes the ELF algorithm. Simulation results and discussions are presented in Section 5. Finally, Section 6 concludes the paper.

2. PREVIOUS WORK

The multi-layer routing and traffic grooming problems have received extensive research interests in the past few years, and an efficient generic graph model was proposed in [17, 18]. Such graph models, however, cannot be extended to the overlay networks: in overlay networks, the two layers are independent of each other, owned by two different owners, making the complete information on the two different layers unavailable to any user. In this paper, we focus on overlay IP/MPLS over WDM networks. Specifically, in this section, we shall present a brief review of the existing algorithms for dynamic LSP routing in such networks, and then choose the best ones among them for performance comparisons. We classify these existing algorithms into two categories, i.e., sequential routing and resource based routing, according to the routing decision strategies adopted by them.

2.1. Sequential routing algorithms

The Logical-Layer-First (LGF) and Optical-Layer-First (OPF) are two representative sequential routing algorithms. In the LGF algorithm, the following two steps are carried out upon the arrival of each transmission request: (1) try to route the request over the residual bandwidth on the existing logical links; (2) if Step 1 fails, then try to set up a new lightpath directly between the ingress and egress LSRs on the WDM-layer network. For the OPF algorithm, the above two steps are reversed. In both algorithms, if both of the two steps fail, the request is blocked.

In [19], Ye *et al.* used the LGF sequential routing algorithm to set up the primary path in their integrated routing/protection strategy. Niu *et al.* presented both the LGF and OPF algorithms and made some comparisons between them when the fixed-path routing is applied to set up new lightpaths [20]. Zhong *et al.* improved the OPF algorithm by adopting dynamic least congested shortest path routing in the WDM-layer network [21]. Although both Niu *et al.* and Zhong *et al.* have studied the influence of add/drop ratio on network performance, they did not take any measures to improve the optical port utilization when setting up new lightpaths. Therefore wavelength utilization tends to be low when lightpaths are long.

To improve both the port and the wavelength resources utilizations, Ye *et al.* proposed an algorithm named Short Lightpath Establishment Approach (SLEA) [22]. Through dynamically assigning link costs in the auxiliary graph considering the optical hop constraint, i.e., assigning a high cost once the optical hop length reaches a certain threshold value and a low cost otherwise, SLEA tries to eliminate the inefficient long lightpaths. Simulation results show that SLEA significantly improves the network blocking performances, and it even outperforms the integrated routing algorithm described in [23] if an appropriate hop constraint is found.

In this paper, we use the OPF and SLEA algorithms for performance comparisons.

2.2. Resource based routing algorithms

Resource based routing algorithms aim at efficiently utilizing the network resources. Two typical examples of such algorithms are the Existing Capacity First (ECF_OVLY) and Minimum Logical Hop (MLH_OVLY) methods proposed by Koo *et al.* [24]. The main idea of ECF_OVLY is to firstly try to use the residual bandwidths of the existing logical links to serve the arriving requests. If that fails, it then tries to set up some new lightpaths, not necessarily from source to destination, for the request. By encouraging setting up shorter lightpaths, ECF_OVLY lowers the chance that lightpaths are under-utilized for a long time.

MLH_OVLY aims to minimize the number of logical hops traversed by each incoming request. For each arriving request, MLH_OVLY first tries to serve it using a single-hop logical link, by either using an existing logical link or setting up a new lightpath between the source and destination LSRs; if that fails, it then tries to find a route with the minimum number of logical hops, where new lightpaths are set up when necessary.

Simulation results in [24] show that MLH_OVLY outperforms ECF_OVLY. In this paper, we use the MLH_OVLY algorithm for comparisons.

3. SYSTEM MODEL AND DEFINITIONS

A typical overlay IP/MPLS over WDM network consists of two layers, with the IP/MPLS layer residing over the WDM layer in a client-server fashion. In such architecture, the IP/MPLS-layer networks are clients while the WDM-layer networks are bandwidth servers, and the two layers' network control and management are independent of each other. Figure 1 shows a sample overlay IP/MPLS-over-WDM network architecture.

With centralized management systems, each of the IP/MPLS and WDM network layers is controlled by its own network operators. Such operators keep all the information of their own network layers, and distribute necessary information and commands to their network elements through the control channels. For each arriving request, the controllers on the two layers can work cooperatively according to their service contracts. Specifically, when a request arrives, the IP-layer controller will firstly try to find a logical-layer route. If this step fails, the request is then transferred to the WDM layer through the well-defined network interfaces, e.g., the UNI. Note that whether to transfer a request to WDM layer is decided by the logical-layer controller; when the request is transferred to the WDM layer, whether the request can be served or not is decided by the WDM-layer controller. Also note that the information exchanges between the two layers can be through any UNI, not necessarily through a direct connection between the two layers' central controllers (if such a direct connection exists at all). It is for the ease

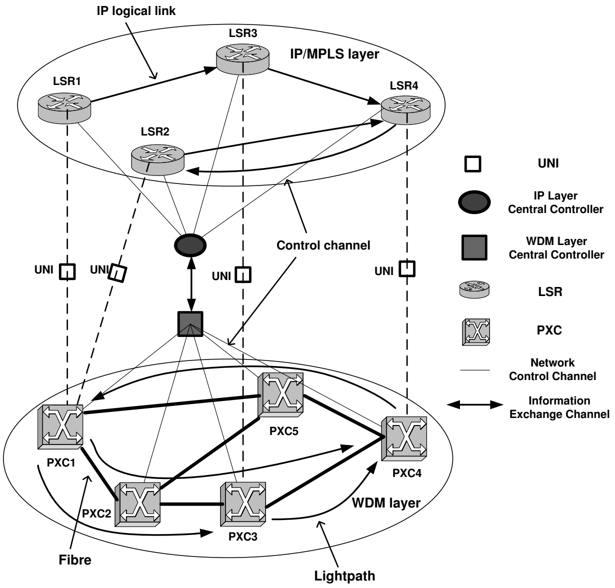


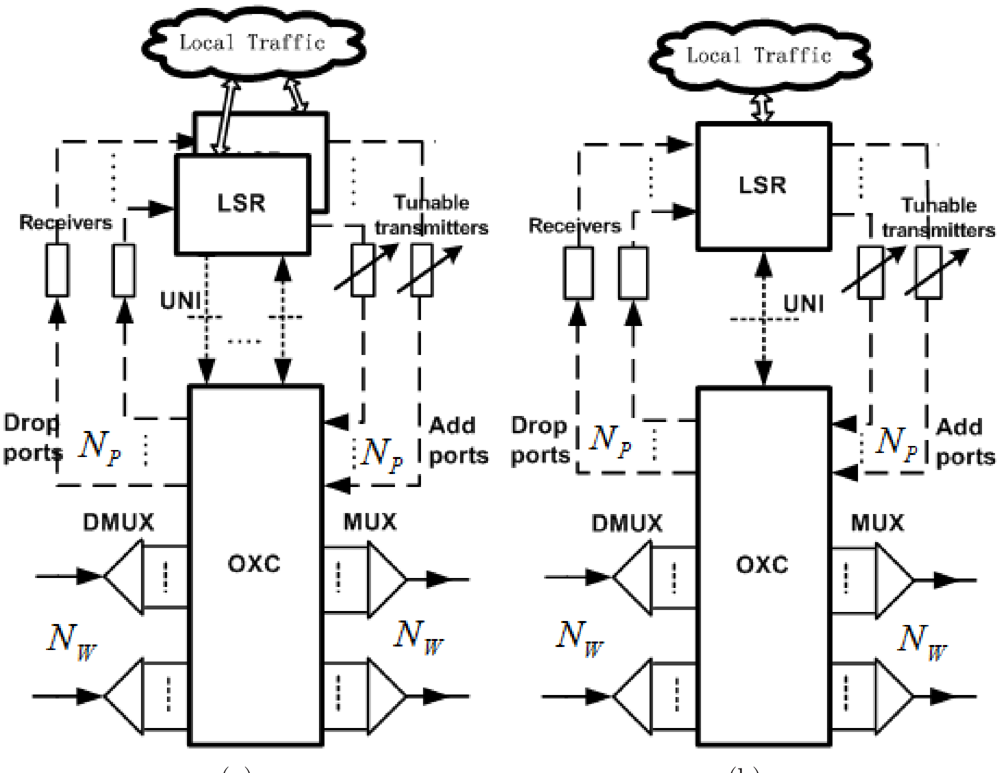
Figure 1: A sample IP/MPLS over WDM network architecture. The WDM layer network consists of five PXCs and six fiber links, and four lightpaths have been set up on it. The IP/MPLS-layer network comprises four LSRs interconnected by the corresponding WDM lightpaths. Each layer is controlled by its own controller, with limited information exchanges between the two controllers.

of illustration and discussion that a direct information exchange channel between the two layers' controllers is shown in Fig. 1.

In the overlay network architecture, each network node is a PXC interconnected with zero, one or more LSRs through the UNI. When there is no LSR connected to it, a PXC is only responsible for switching the bypass traffic, from its input ports to its output ports transparently [25, 26]. When a PXC is connected with one or more LSRs, it has additional functions. Fig. 2(a) illustrates a typical PXC architecture. With several LSRs connected to it, the PXC can receive traffic terminated at the LSR to the local network or transmit traffic originated at the LSR from the local network. The LSRs are used to multiplex local traffic streams into a higher capacity request that PXC can support and also to generate/terminate traffic to/from a lightpath. Note that the number of add/drop ports of a PXC equals the number of transmitters/receivers the node has. Such input/output ports are typically of high costs due to the high-speed electronic processing units they have. Therefore, to save the network cost without sacrificing network performance, a favorable solution is to let each PXC to be equipped with a limited number of add/drop ports shared by all the wavelength channels going through it [27]. In this paper, we call such a PXC architecture as *port-limited*; while in a *port-unlimited* PXC, each wavelength channel is assigned a dedicated pair of add/drop ports. Define *add/drop ratio* r [21] of a PXC as $r = N_p/N_W$ $(0 < r \leq 1)$, where N_p denotes the number of add/drop port pairs and N_W is the number of incoming/outgoing wavelength channel pairs of the PXC. Apparently, $r = 1$ for a port-unlimited PXC.

In this paper, we consider the general case of an overlay IP/MPLS over WDM network with N network nodes and L bi-directional optical fiber links where each link carries W wavelengths. Without loss of generality, each network node is assumed to be a PXC interconnected with a single LSR [22], and all PXCs have no wavelength conversion capability. Note that the work can be easily extended to networks with full or partial wavelength conversion. The node architecture adopted in this paper is shown in Fig. 2(b).

For the network considered, it is assumed that there is a single ISP on top of the WDM network, with the exact link-state information of the logical-layer network. Each time when a new request arrives, the IP-layer ISP will try to find an appropriate route for it and decide whether to use the existing logical links or to set up new lightpaths between LSRs. If new lightpaths need to be established, the ISP sends the request to the WDM-layer bandwidth provider, consulting on the costs of purchasing these lightpaths. Based on



(a) (b)

Figure 2: Typical node architectures for the overlay IP/MPLS over WDM networks. (a) An PXC with multiple LSRs connected to it, where each LSR connects to the PXC through a fixed number of add/drop ports; (b) an PXC with a single LSR connected in it.

the service contract and network resource availability, the WDM-layer bandwidth provider either feedbacks with the lightpath setup costs or rejects the lightpath establishment request if it has run out of wavelength channel and/or add/drop port resources. Such decisions made by the WDM layer operator are independent of those by the logical-layer ISP, as it has no knowledge of either the network topology or the available resources on the logical layer; and vice versa. Once an incoming request is provisioned on the logical layer network using either existing links or new lightpaths, or both, its routing will be kept unchanged by ISP so that the end users' services will not be interrupted.

The information exchanges between the IP and WDM layers basically include only the cost enquiries and feedbacks for the candidate lightpaths to be set up. Since theoretically speaking the number of candidate lightpaths may easily increase exponentially with network size, the IP-layer ISP has to smartly select a small set of candidate lightpaths without the knowledge of WDM-layer topology or resources availability. To make good decisions, it makes sense for the ISP to make use of historical records of lightpath costs. In this paper, we make the reasonable assumption that the ISP can keep record of the lightpath setup costs during a past period of time as well as the time at which such costs are reported by the WDM-layer bandwidth provider. Such historical records can be used to estimate the cost for setting up each candidate lightpath and consequently decide on the candidate route(s). The cost estimation and candidate routes selection methods will be discussed in the next section.

Table 1 presents a summary of the notations used in this paper.

4. PROPOSED ALGORITHM

Section 4.1 describes the graph generation and cost assignment process for making routing decisions. Section 4.2 discusses the historical data learning and cost updating strategies; Section 4.3 presents the complete ELF algorithm; and finally Section 4.4 analyzes the complexity of the algorithm.

4.1. Graph Generation and Cost Assignment

For each incoming request, the ELF algorithm runs the K-shortest path (KSP) algorithm on top of a generated graph to find a desired number of candidate routes. The generated directed graph represents the IP/MPLS-layer network with its nodes being LSRs, and its edges either existing logical

Table 1: NOTATIONS USED IN THIS PAPER

Symbol	**Means**
K	Number of candidate routes for each incoming request
n	The ID number of each incoming connection request
$\bar{H}$	The average number of optical hops traversed by each lightpath
$L(i, j)$	Logical link between LSR i and j
$P(i, j)$	New lightpath between LSR i and j
ω_{ij}	Number of idle wavelengths along $P(i, j)$
C_{ij}	Cost of $P(i, j)$
L_{ij}	Cost of $L(i, j)$
H_{ij}	Minimum number of optical hops between OXC i and j
p_i	Number of idle optical ports available on LSR i
U	The maximum number of historical records of cost for each logical link kept by the IP-layer ISP
T^n	The arrival time of the n-th LSP request
C_{ij}^n	Cost of $L(i, j)$ reported by WDM layer operator at T^n
C_{ij}^{est}	Estimated cost for $L(i, j)$ after the expiration time T_{ij}
T_{ij}	Estimated expiration time for L_{ij}
T_{ij}^{cal}	Calculated expiration time for C_{ij}^{est}
r	Add/drop ratio
$\bar{R}_{ij}$	Estimated average changing rate of L_{ij} during the time before T^n

links with sufficient residual bandwidth for the incoming request or potential new lightpaths to be set up on the WDM layer. We call those edges corresponding to the existing logical links with sufficient residual bandwidths as *existing links*, and those corresponding to the potential new lightpaths as *candidate new lightpaths* (CNLs). After running the KSP algorithm on the generated graph, each of the resulted candidate routes may consist of only existing links, only CNLs, or both.

For each CNL involved in the candidate routes, the logical-layer ISP may signal to the WDM-layer bandwidth provider to enquire its cost. If this CNL is finally chosen to serve the request, a new lightpath will be set up on the WDM layer to support it. However, since an IP-layer ISP does not have the link-state information of the WDM-layer network, CNLs on the candidate routes may turn out to be infeasible due to exhausted wavelength channel and/or input/output port resources. If all the candidate routes are infeasible, the request is blocked.

Upon receiving a lightpath establishment request from the IP layer, theoretically speaking, the WDM-layer bandwidth provider can use any routing and wavelength assignment (RWA) strategy to decide whether and how to set up the required lightpaths. In this paper, since the main focus is to study utilizing historical data on the IP layer, we adopt the shortest hop-count path routing and first-fit wavelength assignment strategy on the WDM layer. Other more sophisticated RWA strategies certainly can also be used. As shall be seen in Section 5, by using the simplest RWA strategy, the proposed algorithm nevertheless outperforms the best existing ones, in many cases by one or two orders of magnitude.

For each lightpath enquired by ISP, the WDM-layer bandwidth provider feedbacks with a market price (say, measured in dollar) and a virtual cost for setting it up. The virtual cost is agreed in service contract. The ISP will keep record of the virtual costs and carefully utilize such records in deciding candidate routes and candidate lightpaths.

The virtual cost shall reflect the resource consumption for setting up the enquired lightpath without revealing detailed WDM-layer information. Also it should discourage over-utilizing a certain link or PXC to avoid emergence of hot spots. The virtual cost therefore should reflect resource consumption as well as resource redundancy/scarcity of the enquired lightpath. To give ISP strong incentives to minimize the virtual cost for setting up a connection, the market price and the virtual cost have to have strongly positive correlation (e.g., the market price may increase faster than being linearly proportional to the virtual cost). This may be a reasonable assumption in most cases since setting up a required connection at a lower virtual cost, when the virtual cost is properly defined, is also of the WDM bandwidth provider's benefits. The bandwidth provider therefore should be willing to reward the cooperative ISP with a lower market price. In this paper, we assume that the ISP always try to lower the virtual cost and always select among the candidate routes the one with the minimum virtual cost. The more complicated cases where the market price may not be positively correlated to the virtual cost have to be discussed in a separate report.

In a port-limited PXC, the limited number of optical ports plays an important role in governing network performance. To improve the performance of a network with limited numbers of wavelengths and optical ports, these two types of resources should be consumed in a balanced manner. Specifically, if both wavelengths and optical ports are abundant, the costs of consuming them should be low and not so different from each other; while if any one of them becomes scarce, the cost of consuming the scarce resource should become higher to impose a penalty to utilizing it. Therefore, we define the cost C_{ij} of a new lightpath $P(i, j)$ by taking into account the costs of consuming optical ports and wavelength resources as follows:

$$C_{ij} = \begin{cases} \left(\frac{\alpha}{p} - H_{ij} \ln\left(1 - \frac{1}{\omega_{ij}+1}\right)\right) \times amp & \text{if } \omega_{ij} > 0 \text{ and } p > 0 \\ \infty & \text{if } \omega_{ij} = 0 \text{ or } p = 0 \end{cases} \tag{1}$$

This cost function is intended for both port-limited and unlimited cases. Specifically, it consists of two parts. The first part reflects the cost of consuming a pair of optical ports at LSR i and and LSR j: $p = \min(p_i, p_j)$ denotes the minimum number of optical ports available at the two end nodes of the candidate lightpath. The parameter $\alpha = \frac{(1-r)}{r(\bar{H}+1)}$ regulates the relative weights of the costs of a wavelength and an optical port: a smaller add/drop ratio leads to a higher cost of consuming a pair of optical ports. Note that the cost for consuming optical ports reduces to zero in port-unlimited case where $r = 1$. The second part calculates the cost of consuming a wavelength along each hop of the lightpath. The negative symbol is to ensure the second part a positive quantity, and $\omega_{ij} + 1$ is used to avoid generating an infinity value when $\omega_{ij} = 1$. Finally, *amp* is an amplification factor regulating the ratio of the cost of using existing logical links and that of setting up new lightpaths. Such a definition of lightpath cost helps avoid selecting a route with too few idle optical ports or too few idle wavelengths, or too many optical hops. Traffic loads on the WDM layer therefore may be better balanced.

We now discuss the logical-layer link cost assignment. For simplicity, we classify the CNLs into *cost enquired* and *cost unknown ones.* If the cost of a link has been enquired before, it is a cost enquired link; otherwise, it is a cost unknown one. The costs of different types of links are defined as follows.

$$L_{ij} = \begin{cases} 1 & \text{an existing logical link} \\ M_{ij} & \text{a cost unknown virtual link} \\ C_{ij}^{est} & \text{a cost enquired virtual link} \\ 2M_{ij} & \text{a failed lightpath for } L(i,j) \end{cases} \tag{2}$$

where M_{ij} is a default value of the cost of $L(i,j)$. The default value can be suggested by the WDM-layer network operator to the IP-layer ISP, e.g., as an average from some past experiences, or it can be calculated by using some typical values of the relevant parameters. In our experiences, a simple M_{ij} calculation method as below can steadily lead to satisfactory performance.

For a given WDM-layer network with an average nodal degree δ, denote the average hop length of each lightpath when adopting the fixed minimum-hop routing method as $\bar{H}$, and the average number of idle wavelengths on each optical link at a certain network status as ω. Since the average number of add/drop ports on each node approximately equals $W \times r \times \delta$, the average number of idle optical ports on each node at such status can be estimated as

$$p = W \times r \times \delta - \frac{1}{\bar{H}+1} \times \delta \times (W - \omega) \; . \tag{3}$$

The second part on the right side of the above equation comes from the fact that each lightpath only uses add/drop ports on its two end nodes. To avoid having a zero or negative value for p under heavy traffic loads, we let

$$p = \max\left(W \times r \times \delta - \frac{1}{\bar{H}+1} \times \delta \times (W - \omega), 1\right) \; . \tag{4}$$

The default value of M_{ij} for $L(i,j)$ can be calculated as

$$M_{ij} = \left(\frac{\alpha}{p} - H_{ij} \ln\left(1 - \frac{1}{\omega+1}\right)\right) \times amp \; , \tag{5}$$

where the value of ω can be anything between 0 and W, depending on network status. Our experiences show that the performance of the proposed algorithm is not very sensitive to the value of ω. A convenient option with satisfactory performance is to let $\omega = \frac{W}{2}$, which is adopted in all the simulations reported in this paper.

4.2. Data Learning and Cost Expiration Process

As described in Sec. 3, each time a new request arrives at the network, the WDM-layer bandwidth provider may report the costs of some CNLs to the IP-layer ISP upon request, while the IP-layer ISP keeps record of such information. To avoid keeping excessive records of historical data, only a limited number of latest records are kept for each logical link. As shall be seen later, keeping a large number of dated records may not help improve network performance.

Utilizing the historical data records, the IP-layer ISP is able to estimate the cost of each logical link. However, as the WDM-layer network operations are independent of those on the IP/MPLS layer, the cost of building a new lightpath between two end nodes may change significantly over time. To avoid outdated information leading to bad routing decisions, we introduce a cost expiration strategy: a link cost record which was updated long time ago (e.g., longer than a pre-defined threshold) is deemed as outdated and thus should be adjusted, e.g., towards a certain default value. The threshold time for a link-cost record to be adjusted is termed as its *expiration time*. Specifically, the cost expiration process for a logical link works as follows: whenever the cost of a logical link is reported by the WDM layer, its logical-layer record is updated accordingly. Meanwhile its expiration time is calculated. Upon the arrival of a new transmission request at a certain time T^n, the expiration time of the logical link is compared to T^n. If the expiration time has not been reached yet, the link cost record is regarded as valid, and thus can be used directly in calculating the candidate routes. Whereas if the expiration time has already been reached, it means that the link cost information has been kept for too long and thus should be adjusted towards its default value. After the adjustment, a new expiration time is calculated if needed. The cost expiration process is repeated until the estimated link cost equals its default value, or it gets updated by the latest information reported from the WDM-layer network.

In this paper, we propose to let the estimated link cost to be adjusted towards its default value in a few steps upon expiration. The corresponding expiration time is calculated by utilizing the average cost changing rate from the historical records. Specifically, upon receiving the accurate information of the cost of a logical link $L(i,j)$, denoted as C_{ij}^n, at time T^n, let

$$d_{ij} = \operatorname{sgn}(C_{ij}^n - M_{ij}) \ . \tag{6}$$

and define its average cost changing rate as

$$\bar{R}_{ij} = \begin{cases} \frac{1}{m}\sum\limits_{t=1}^{m}\left|\frac{C_{ij}^t - C_{ij}^{t-1}}{T^t - T^{t-1}}\right| & m < U \\ \frac{1}{U}\sum\limits_{t=m-U+1}^{m}\left|\frac{C_{ij}^t - C_{ij}^{t-1}}{T^t - T^{t-1}}\right| & m \geq U \end{cases} , \tag{7}$$

where U denotes the maximum number of historical data records kept for each link. The expiration time of the link cost can be calculated as

$$\delta t = \begin{cases} \infty & C_{ij}^n = M_{ij} \\ T^n - T^{n-1} & C_{ij}^n \neq M_{ij} \text{ and } \overline{R}_{ij} = 0 \\ \frac{\left|C_{ij}^n - M_{ij}\right|}{\overline{R}_{ij}} & C_{ij}^n \neq M_{ij} \text{ and } \overline{R}_{ij} \neq 0 \end{cases} , \tag{8}$$

Then for any request arriving at a certain time T^m , the estimated cost of $L(i,j)$ can be calculated as follows,

$$C_{ij}^{est} = \begin{cases} C_{ij}^{n} - d_{ij}\Delta \left| C_{ij}^{n} - M_{ij} \right| \times \min\left(\frac{1}{\Delta}, \left\lfloor \frac{T^m - T^n}{\delta t} \right\rfloor\right) & \delta t \neq \infty \\ M_{ij} & \delta t = \infty \end{cases} \tag{9}$$

where parameter Δ ($0 < \Delta \leq 1$) is a constant for controlling how much the link cost should be adjusted towards its default value once it is deemed to be expired.

The next expiration time of the newly adjusted link cost estimation C_{ij}^{est} can be calculated as

$$T_{ij}^{cal} = \begin{cases} T^n + \left\lfloor 1 + \frac{T^m - T^n}{\delta t} \right\rfloor \delta t & C_{ij}^{est} \neq M_{ij} \\ \infty & C_{ij}^{est} = M_{ij} \end{cases} . \tag{10}$$

Once the cost record of a logical link is deemed as expired, L_{ij} and T_{ij} are updated by C_{ij}^{est} and T_{ij}^{cal} respectively, and such updating process is repeated until C_{ij}^{est} equals to M_{ij} or until it is updated by the new cost reported from the WDM layer. Simulation results show that the BBR performance is not very sensitive to the value of Δ. In our simulations, we let $\Delta = 0.2$, which steadily achieves slightly better performance than $\Delta = 1.0$ (where link cost is adjusted to be equal to its default value once it is deemed as expired).

Note that in such data expiration process, different logical links may have different cost expiration time, which makes sense since different links may be under different traffic loads, leading to different frequencies of link-cost changes. Our experiences show that having different expiration time for different links steadily leads to better performance than updating all links' costs with the same interval. Further, note that the data learning and cost expiration process does not introduce any significant additional computational complexity to the routing process.

With the graph generation method and the data learning and cost expiration strategies as described above, we now present the ELF algorithm.

4.3. Existing Link First (ELF) Algorithm

The algorithm begins by assigning a cost to each link in the logical-layer generated graph, and then running the KSP algorithm on it to find a desired number of candidate routes for each incoming request. As discussed in Sec. IV-A, some candidate routes may contain only existing links while the others contain some CNLs. The algorithm gives using existing links a higher priority since such a strategy generally leads to better performance [17], [28].

Specifically, it checks through all the candidate routes. If there exist routes with only existing links, the one with the minimum cost is selected to serve the request; while if no such route exists, the IP-layer ISP shall then query the WDM layer for the costs of all CNLs along the candidate routes. Based on the CNL costs feedback from the WDM layer, the feasible route with the minimum overall cost, if any, is selected to serve the incoming request. Note that historical records of CNL costs are used in calculating the estimated link costs when finding the candidate routes. And such records are (partially) updated each time when there is feedback of link costs from WDM-layer network.

Since the algorithm gives a higher priority to utilizing the existing logical links while enquiring WDM layer for CNL costs only when necessary, we term it the Existing Link First (ELF) algorithm. Algorithm 1 shows its main steps.

4.4. Algorithm Complexity

As described in Sec. 2, since all the existing algorithms find their candidate routes for each arriving request using Dijkstra's shortest path algorithm, the computational complexity of these algorithms is $O\left(K\left(L + N\log N\right)\right)$.

Compared with that of the existing algorithms, differences in complexity of ELF mainly come from three aspects:

(1) The K-loopless shortest path (KSP) algorithm used to find K logical-layer candidate routes for each incoming request, with complexity of
$O\left(KN\left(L + N\log N\right)\right)$;

(2) The data learning and cost expiration process, which introduces hardly any additional computational complexity, as discussed;

(3) The storage space required for keeping record of historical costs and their corresponding reported time, with a complexity of $O\left(UN^2\right)$;

For typical optical networks with a moderate number of nodes, the additional computational complexity and storage space needed by the ELF algorithm should be acceptable considering the performance improvement.

5. PERFORMANCE EVALUATION

As in previous work [20–24], a dynamic traffic model is utilized to study the blocking performance of the proposed algorithm. Without loss of generality, only unidirectional LSP requests are considered. Assume that the

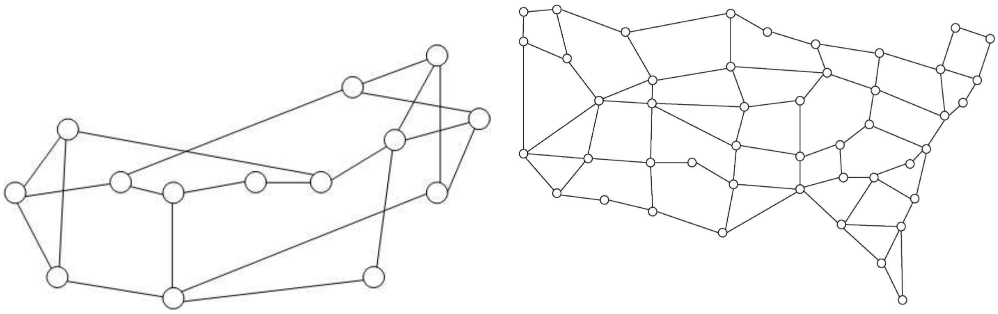
Algorithm 1: EXISTING LINK FIRST (ELF) ALGORITHM

```
   input  : Network $G(V, E)$, Request $R(s \to d, b)$, $K$
   output : A path route for request $R(s \to d, b)$
1  Initialization.  foreach PXC pair do find a minimum hop optical
   layer route;  foreach Logical link do $L_{ij} = M_{ij}$, $T_{ij} = \infty$ ;
2  for Each arriving request do
3      if It is a connection request then go to Line 6;
4      else go to Line 20;
5  end
6  for n-th LSP request arriving at the network at time $T^n$ do
7      update the estimated cost and expiration time for those links
       whose costs become expired, i.e., if $T_{ij} < T^n$ then $L_{ij} = C_{ij}^{est}$,
       $T_{ij} = T_{ij}^{cal}$;
8      IP-layer graph generation and cost assignment;
9      Run the KSP algorithm to find $K$ candidate routes;
10     Check the link property of all links along the $K$ candidate routes;
11     if there exist routes containing only existing links, then
12         choose the route with the minimum cost;
13     else
14         Enquire $C_{ij}^n$ for all CNLs along the $K$ candidate routes;
15         Update cost and expiration time for all cost enquired CNLs:
           $L_{ij} = C_{ij}^n$, $T_{ij} = T^n + \delta t$;
16         Choose the route with the minimum overall cost, if applicable;
17     end
18     Serve the connection request using the selected route if there exists
       at least one feasible route; otherwise, block the connection request;
19 end
20 Update both WDM and IP network status;
```

LSP requests arrive at the network independently following a Poisson process with a mean arrival rate of λ, and the LSP holding time is exponentially distributed with a unit mean, i.e., $\frac{1}{\mu} = 1$. The source and destination node pair of each LSP request is randomly chosen among all network nodes. The bandwidth of each wavelength is divided into 16 units, and the number of bandwidth units requested by each LSP is an integer uniformly distributed between 1 and 16. Each LSP request has to be handled along a single route without splitting.

We evaluate the proposed algorithm mainly by measuring the *bandwidth blocking ratio* (BBR) [20, 21] of the network. Extensive simulations are carried out on two typical network topologies. As shown in Fig. 3, they include the 14-node NSFnet and 46-node USNET. The average number of optical hops traversed by each lightpath is $\overline{H} = 2.18$ for NSFnet and $\overline{H} = 4.4$ for USNET respectively. Since in practical networks, the cost for setting up a new lightpath is generally much higher than that of using an existing logical link, we set the amplification factor *amp* to be 10 and 5 for NSFnet and USNET respectively, such that the cost of using a new lightpath is roughly about 5 times [29] as much as that of using an existing logical link in a port-unlimited network where averagely half of all the wavelength channels are still idle and the corresponding average number of idle optical ports is calculated by Eq. 3. Some other assumptions adopted in the simulations include: 1) each fiber link carries $W = 8$ wavelengths; 2) for the SLEA algorithm, the optical hop constraint is set to 2 for both NSFnet and USNET; 3) the number of optical ports at a network node is set to be $W \times r \times \delta_i$ with δ_i being the node's fanout degree. Results shown in each of the following figures are the average of more than 30 independent simulations, each of which simulating 10^5 connection requests. We observe that the simulation results turn out to be highly consistent, with variance smaller than 4% when there is a single logical-layer candidate route for each communication request (or in other words when $K = 1$), and smaller than 7% for $K = 2$ and $K = 3$.

The following subsections compare the BBR performances of the ELF algorithm against those of the existing ones in port-unlimited and port-limited networks respectively. We also evaluate the effects of the number of optical ports and the amount of recorded historical information. Since all conclusions hold for both topologies, unless otherwise specified, we present only the results on NSFnet for comparisons and discussions.



(a) (b)

Figure 3: Two network topologies used for simulations. (a) 14-node NSFnet. (b) 46-node USNET.

5.1. Performance Comparison in Port-Unlimited Networks under Different Traffic Loads

We first consider the case with port-unlimited PXCs. Figure 4 shows the BBR performances against network traffic loads measured in Erlangs. As can be seen, ELF significantly outperforms the three existing algorithms, namely SLEA, MLH_OVLY and OPF respectively. Specifically, when $K = 1$, ELF outperforms MLH_OVLY and SLEA by about an order of magnitude, and even more over OPF. When $K = 3$, the improvements increase to more than two orders of magnitude.

Note that among the three existing methods, OPF performs the worst, which is not a surprise since it tries only a single candidate route at a single layer for each arriving request. MLH_OVLY outperforms OPF by 70% when traffic load is about 70 Erlangs and 40% when traffic load is around 100 Erlangs, as it tests multiple routes for each request. By imposing an optical hop constraint on new lightpaths, the effect of which is the same as that of using a limited number of wavelength converters on certain network nodes, SLEA outperforms MLH_OVLY by about 20% within the whole range of traffic loads.

Comparison between ELF and MLH_OVLY convincingly demonstrates the effectiveness of utilizing historical data to improve the BBR performance of overlay IP-over-WDM networks: both methods find more than one candidate route for each incoming request, yet ELF performs much better, thanks to its careful utilizations of historic data. Comparison between ELF and SLEA shows that proper utilization of historical information helps the IP-

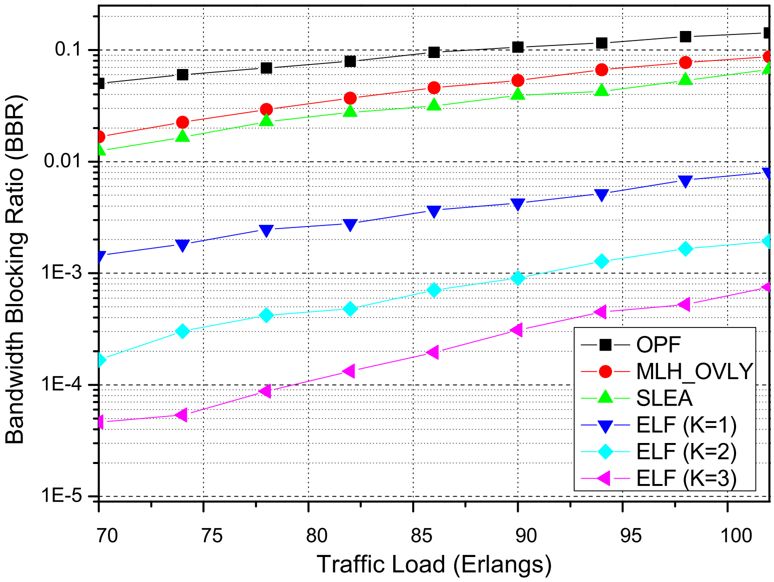


Figure 4: BBR comparisons between the ELF algorithm and the existing ones under different traffic loads without optical port limitation ($r = 1.0$, $\Delta = 0.2$, $U = 5$).

layer ISP make better routing decisions, leading to better performance than that of using high-cost wavelength converters on certain network nodes.

Another interesting observation from Fig. 4 is that increasing the number of candidate routes K for each incoming request improves the BBR performance of ELF. Specifically, when increasing the number of candidate routes from $K = 1$ to $K = 2$, the average BBR improvement is 80% under different traffic loads; further increasing from $K = 2$ to $K = 3$, the average further improvement is 71%. Such improvements mainly come from two aspects: (i) a larger value of K gives ELF a higher chance to find a more appropriate candidate route for each incoming request, and thus lowers the blocking probability; and (ii) a larger K allows more CNLs to be involved in the candidate routes. Their costs are therefore updated more frequently, enabling more accurate selections of candidate routes.

5.2. Performance Comparisons in Port-Limited Networks under Different Traffic Loads

We now study the BBR performance in port-limited networks under different traffic loads. In this subsection, we In this subsection, we let $r = 0.6$ for all LSRs, while the effects of different values of r will be evaluated in the next subsection. Figure 5 shows the BBR performance against network traffic loads. Results again demonstrate that ELF significantly outperforms the existing algorithms: in average, when $K = 1$, ELF performs about 77%, 57% and 37% better than OPF, MLH_OVLY and SLEA respectively; while when $K = 3$, the improvements over the three algorithms increase to about 89%, 79% and 69% respectively.

A noteworthy observation in Fig. 5 is that when the number of candidate routes for each incoming request increases, the performance of ELF also improves, though the improvements are not as significant as those in port-unlimited networks. Results show that when increasing from to , the average improvement is about 46%, further increasing from $K = 2$ to $K = 3$, the average further improvement is at a much lower value of 11%. The improvements are mainly due to the fact that increasing K makes more CNLs to be involved in the candidate routes and thus increases information exchanges between the two layers. The effects of having more candidate routes to increase the chance of finding a feasible route, i.e., the contribution (i) discussed in the last subsection, meanwhile become less significant in port-limited networks. This can be understood: the limited port resources make

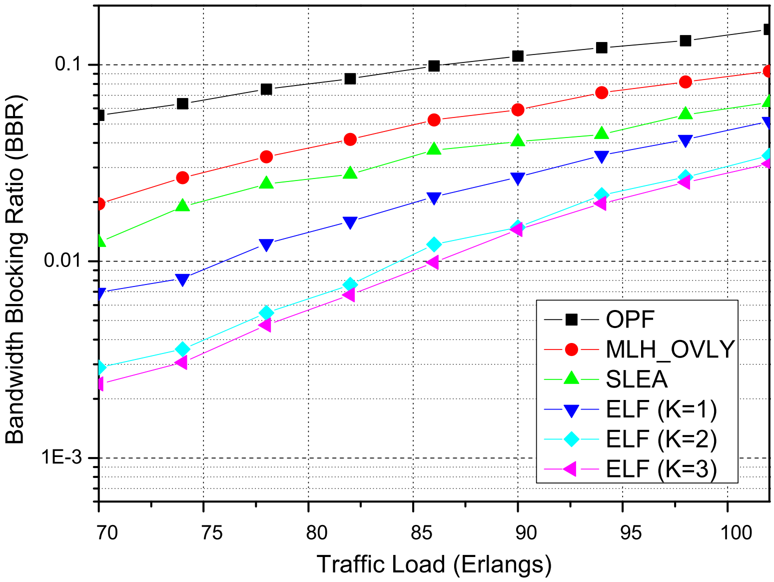


Figure 5: BBR comparison between ELF and the existing ones under different traffic loads when considering optical port limitation ($r = 0.6$, $\Delta = 0.2$, $U = 5$).

longer routes going through more existing links (and consequently consuming more optical ports) less favorable.

Overall, the results shown in Fig. 5 demonstrate the effectiveness of the ELF algorithm in port-limited networks.

5.3. Influence of the Limited Number of Optical Ports

Figure 6 compares the performance of the ELF algorithm with those of the three existing algorithms with different add/drop ratio r. The traffic load is fixed at 80 Erlangs. Results show that ELF significantly outperforms the existing algorithms within a wide range of add/drop ratio. For $K = 1$, ELF outperforms the existing algorithms once $r > 0.45$; when $K \geq 2$, it outperforms them once $r \geq 0.4$. Closer observations reveal that, when the add/drop ratio is large enough, e.g., $r > 0.55$, ELF steadily outperforms any existing algorithms with a wide margin; whereas when r is of a small value, all the algorithms perform nearly the same. This can be explained as follows: when limited optical ports become the bottleneck resource dominating network performance, different algorithms do not make significant differences. However, once the bottleneck constraint is relaxed to a certain extent, i.e., when r is large enough, the ELF algorithm, with its capability of more efficiently utilizing network resources, easily stands out.

Another interesting observation from Fig. 6 is that the performance of ELF improves steadily with an increasing value of r, which is different from that of the existing algorithms of which the performances stay largely unchanged once r is larger than a certain threshold value. This is because when r is large enough, wavelength resources, instead of optical port resources, become the bottleneck. Once the existing algorithms such as OPF, MLH_OVLY and SLEA reach their respective best utilizations of wavelength resources, they will not benefit from the redundant optical port resources. On the contrary, the ELF algorithm, with its logical-layer dynamic routing process, enjoys better flexibility in utilizing network resources. Specifically, by carefully utilizing historical records and the enquired information of CNL costs, ELF efficiently avoids those CNLs with limited wavelength resources, sometimes at the cost of using more optical ports. Better utilization of redundant optical port resources therefore steadily leads to better performance with an increasing value of r.

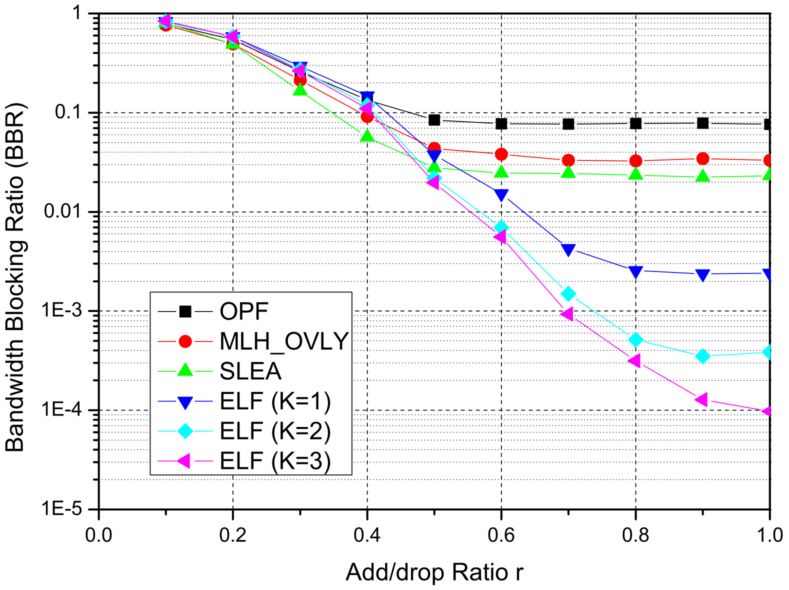


Figure 6: Performance comparison between ELF and the existing algorithms with different values of the add/drop ratio r. The traffic load is fixed at 80 Erlangs ($\Delta = 0.2$, $U = 5$).

5.4. Influence of the Amount of Historical Data

As mentioned in Section 4.2, for each logical link, ELF only keeps record of a small number of latest link costs reported from the WDM layer. We now evaluate how the amount of historical link cost information affects the network performance.

As shown in Eq. (7), in the ELF algorithm, the amount of historical data kept for each logical link is decided by the parameter U. A larger value of U lets the average cost changing rate to be calculated over a relatively longer time, and vice versa. Figure 7 shows the BBR performance with different values of U. Simulation results show that a smaller value of U basically leads to better performance, yet when U is too small (e.g., $U = 1$ or 2), there exist bigger fluctuations in BBR performance. Such observations can be explained as follows: the link cost changing rate may not be very stable, especially considering the fact that we are using the dynamic routing method on the IP/MPLS layer. . If we use a long-term average to estimate the link cost changing rate within a short period of time in future, over- or under-estimation may happen. That is why a larger value of U does not lead to better performance. When U is too small, however, the average changing rate is only estimated by how link cost is changed with the latest one or two connection requests, which may easily cause fluctuations. In this paper, we set $U = 5$ in our simulations as it leads to more stable BBR performance, convenient for comparison against those of the existing methods. If only the BBR performance is concerned, however, setting $U = 1$ may be a better option: a small value of U helps achieve better BBR performance, and meanwhile saving the space for historical data storage.

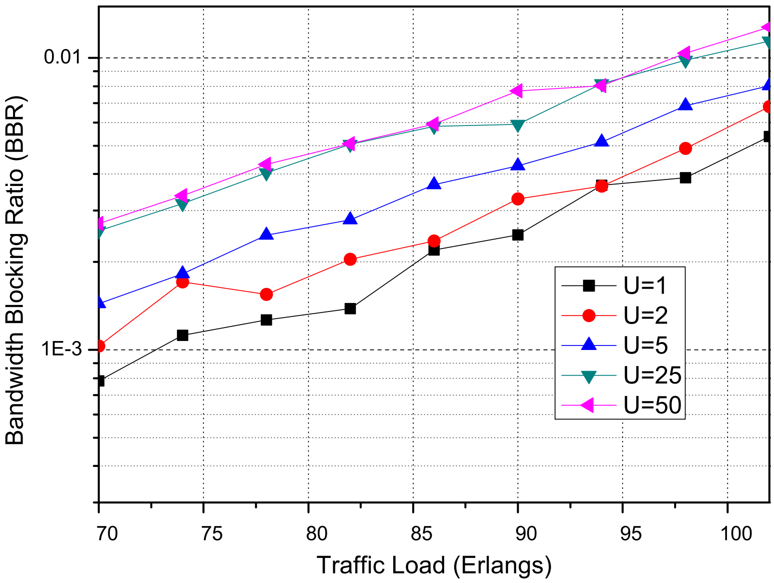


Figure 7: ELF performance with different number of recorded costs without optical port limitation ($r = 1.0$, $\Delta = 0.2$, $K = 1$).

6. CONCLUSIONS AND FUTURE WORK

In this paper, we reported a study on the dynamic LSP routing problem in overlay IP/MPLS over WDM networks. To improve the overlay network performance, we proposed to learn from the historical information maintained by the IP/MPLS-layer ISP. By carefully utilizing a data learning and cost expiration scheme for logical link cost estimation, and adopting the KSP algorithm for logical-layer routing, a novel algorithm named Existing-Link-First (ELF) was proposed. Extensive simulation results show that the proposed algorithm significantly outperforms all the existing ones under different traffic loads, with either limited or unlimited resources of optical ports as long as such resources are not too restrictive. The very significant improvements in BBR performances come at a cost of a negligible additional computational complexity and a small amount of historical data storage on the IP/MPLS layer.

Since the main focus of this study is on utilizing historical data for dynamic LSP routing, detailed discussions on information exchanges between the IP/MPLS and the WDM layers have been largely omitted. As IP/MPLS over WDM technologies are maturing quickly, appropriate protocols almost certainly can be developed in the near future for such information exchanges.

Note that the three algorithms we adopted for comparisons belong to two different classes. OPF and SLEA are sequential routing methods, which try to provision a request on a single network layer. Information exchanges between different layers are not needed and thus, keeping historical data does not help improve network performances. In contrary, MLH_OVLY belongs to the resource based methods, which allow setting up a new connection across two different layers to achieve more efficient utilization of network resources. Algorithms of this class may be revised to make use of the historical data to help facilitate logical layer routing. It would be of our future research interest to investigate how much historical data learning could help improve the performances of these algorithms. Another interesting topic is to extend the ELF algorithm to support multicast communications.

References

[1] A. M. Odlyzko, Internet traffic growth: Sources and implications, in: Proc. SPIE, Vol. 5247, 2003, pp. 1–15.

[2] R. Ramaswami, Optical networking technologies: What worked and what didn't, IEEE Commun. Mag. 44 (9) (2006) 132–139.

[3] M. A. Ali, A. A. Shami, C. M. Assi, Y. Ye, R. Kurtz, Architectural options for the next-generation networking paradigm: Is optical internet the answer?, Photon Netw. Commun. 3 (1/2) (2001) 7–21.

[4] N. Ghani, S. Dixit, T.-S. Wang, On IP-over-WDM integration, IEEE Commun. Mag. 38 (3) (2000) 72–84.

[5] E. Rosen, A. Viswanathan, R. Callon, Multiprotocol label switching architecture, IETF RFC 3031.

[6] G. Swallow, J. Drake, Boeing, H. Ishimatsu, Y. Rekhter, Generalized multiprotocol label switching (GMPLS) use (Oct. 2005).

[7] D. Papadimitriou, M. Vigoureux, K. Shiomoto, D. Brungard, J. L. Roux, Generalized MPLS (GMPLS) protocol extensions for multi-layer and multi-region networks (MLN/MRN) (Oct. 2010).

[8] User network interface (UNI) 1.0 signaling specification, release 2 (Feb. 2004).
URL `www.oiforum.com/public/documents/OIF-UNI-01.0-R2-RSVP.pdf`

[9] J. P. Vasseur, J. L. Roux, Y. Ikejiri, A set of monitoring tools for path computation element (PCE)-based architecture (Jun. 2010).

[10] G. N. Rouskas, H. G. Perros, A tutorial on optical networks, NETWORKING 2002 Tutorials, LNCS 2497 (2002) 155–193.

[11] J. Y. Wei, Advances in the management and control of optical internet, IEEE J. Sel. Areas Commun. 20 (4) (2002) 768–785.

[12] K. H. Liu, C. Liu, J. L. Pastor, A. Roy, J. Y. Wei, Performance and testbed study of topology reconfiguration in IP over optical networks, IEEE Trans. Commun. 50 (10) (2002) 1662–79.

[13] L. Ruan, F. Tang, Survivable IP network realization in IP-over-WDM networks under overlay model, Comput. Commun. 29 (10) (2006) 1772–1779.

[14] B. Martini, F. Baroncelli, P. Castoldi, A. Muchanga, L. Wosinska, The service oriented optical network (SOON) project, in: Proc. 2005 7th International Conference on Transparent Optical Networks, 2005, IEEE ICTON '2005, 2005, pp. 429–432.

[15] G. S. Z. amd Valerio Martini, Y. Qin, E. Escalona, R. Nejabati, D. Simeonidou, F. Baroncelli, B. Martini, K. Torkmen, P. Castoldi, Service-oriented multigranular optical network architecture for clouds, J. Opt. Commun. Netw. 2 (10) (2010) 883–891.

[16] E. Q. V. Martins, M. M. B. Pascoal, A new implementation of Yen's ranking loopless paths algorithm (2000).
URL `http://citeseer.ist.psu.edu/martins00new.html`

[17] H. Zhu, H. Zang, K. Zhu, B. Mukherjee, A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks, IEEE/ACM Trans. Netw. 11 (2) (2003) 285–299.

[18] B. Chen, G. N. Rouskas, R. Dutta, On hierarchical traffic grooming in WDM networks, IEEE/ACM Trans. Netw. 16 (5) (2008) 1226–1238.

[19] Y. Ye, C. Assi, S. Dixit, M. A. Ali, A simple dynamic integrated provisioning/protection scheme in IP over WDM networks, IEEE Commun. Mag. 39 (11) (2001) 174–182.

[20] X. Niu, W. D. Zhong, G. Shen, T. H. Cheng, Connection establishment of label switched paths in IP/MPLS over optical networks, Photon Netw. Commun. 6 (1) (2003) 33–41.

[21] W. Zhong, X. Niu, B. Chen, S. K. Bose, Y. Jin, Performance comparison of overlay and peer models in IP/MPLS over optical networks, Photon Netw. Commun. 9 (1) (2005) 121–131.

[22] T. Ye, Y. Su, K. Zhu, L. Leng, Q. Zeng, Y. Jin, SLEA: A novel scheme for routing in overlay IP/WDM networks, IEEE/OSA J. Lightwave Technol. 23 (10) (2005) 2934–2944.

[23] Q. Zheng, G. Mohan, An efficient dynamic protection scheme in integrated IP/WDM networks, in: Proc. IEEE ICC'2003, Vol. 2, 2003, pp. 1494–1498.

[24] S. Koo, G. Sahin, S. Subramaniam, Dynamic LSP routing in IP/MPLS over WDM networks, IEEE J. Sel. Areas Commun. 24 (12) (2006) 45–55.

[25] S. Melle, D. Perkins, C. Villamizar, Network cost savings from router bypass in IP over WDM core networks, in: Proc.OFC/NFOEC '2008, 2008.

[26] X. Cao, V. Anand, Y. Xiong, C. Qiao, A study of waveband switching with multilayer multigranular optical cross-connects, IEEE J. Sel. Areas Commun. 21 (7) (2003) 1081–1095.

[27] G. Shen, W. D. Grover, Dynamic path-protected service provisioning in optical transport networks with a limited number of add/drop ports and transmitter tunability, IEEE J. Sel. Areas. Commun. 25 (6) (2007) 121–134.

[28] A. Giorgetti, L. Valcarenghi, P. Castoldi, Combining multi-granularity dynamic provisioning and restoration schemes, in: Proc. ONDM, Ghent, Belgium, 2004.

[29] E. A. Doumith, M. Gagnaire, Impact of traffic predictability on WDM EXC/OXC network performance, IEEE J. Sel. Areas Commun. 25 (5) (2007) 895–904.