Topology and routing optimization for congestion minimization in optical wireless networks

Iradj Ouveysi, Feng Shu, Wei Chen, Gangxiang Shen, Moshe Zukerman

Abstract

Optical wireless networks have appealing features such as very high broadband data rates and cost effectiveness. They represent a potential alternative to the last mile (first mile) wireless access problem. However, they are also highly vulnerable to external disturbances such as adverse weather and building sway. In this paper, we develop robust and efficient methods for outdoor optical wireless networks by jointly considering topology optimization and survivability strategies. We propose linearized congestion minimization schemes with working and protection paths (LCM–WP), in which a mixed integer linear program is formulated to choose the optimal working and protection paths for every OD pair such that the network congestion is minimized. In particular, the objective is to minimize the maximum amount of traffic on the links. To solve realistically sized problems, we consider a restricted version of the LCM–WP, in which only limited sets of candidate working and protection paths are considered. A simple algorithm is developed to find candidate working and protection paths for each origin-destination (OD) pair. Implementation of our LCM–WP schemes demonstrates the efficiency of our approach in terms of the number of constraints and solution time. It also shows that our approach is applicable to realistically sized networks.

1. Introduction

The technology of optical wireless, also known as free-space optics (FSO) communication, has been continuously growing as an attractive solution for the provision of high data rates over short distances in recent years [1–5].

An optical wireless link consists of a transmitter, the propagation channel and a receiver. FSO links typically have a maximum transmission range of a few kilometers and direct line-of-sight (LOS) is imperative in optical wireless communications. Invisible, eye-safe light beams are used between optical transceivers to provide optical bandwidth connections that can send and receive voice, video, and data information [6].

In an optical wireless network, connectivity between end users is achieved using optical transceivers installed on windows, building rooftops or exterior walls. The limitations of imperative LOS and transmission range force FSO transceivers to have direct links only with relatively close-by neighbors. This leads to typically low nodal degrees, and in turn, imposes constraints on the number of disjoint
Fig. 1. An optical wireless transceiver. Source: HCS Technologies [9].

Fig. 2. A fictitious optical wireless network illustration in an urban area of Hong Kong.

paths between origin-destination (OD) pairs which are important for network survivability.

Commercially available optical wireless transceivers (see Fig. 1 for a typical example) can provide high-speed data rates in the range of 100 Mbps to 2.5 Gbps, and a data rate as high as 160 Gbps has been reported in demonstration systems [7]. Potential applications of the optical wireless technology include: last mile network access [8], enterprise connectivity, fiber backup, and disaster recovery. A fictitious urban optical wireless network in Hong Kong is illustrated in Fig. 2.

However, the performance of optical wireless links is highly vulnerable to external disturbances [10], such as:

- atmospheric disturbances (e.g., fog, absorption, scattering, scintillation), which could result in high bit error rate (BER) and transmission delays [11, 12];
- sway of buildings, which leads to a difficulty in maintaining the strict line-of-sight requirement [3, 12].

Consequently, the progress of widely accepting optical wireless technology has been slow. Innovative technologies and architectures are still required to provide solutions for high-speed, resilient and robust optical wireless networks.

Many methods have been proposed to increase the accuracy and robustness of optical wireless links. Kwok et al. [13] proposed a temporal domain diversity reception scheme that reduces BER and improves the reliability of optical wireless transmission links. Such a scheme is especially important under atmospheric turbulence effects. Ho et al. [14] introduced an alignment approach to improve pointing performance using the geometrical dependence between cameras and transceivers in mid- and long-range scenarios. In [15, 16], Muhammad et al. proposed an FSO system with increased resilience to varying weather conditions, by adopting suitable modulation and coding schemes. Yusal et al. [17] analyzed the error performance for coded FSO communication systems operating in atmospheric turbulence channels. O’Brien et al. [18] introduced a solid-state architecture and seven-channel tracking system to design integrated and accurate optical wireless transceivers.

It has been shown that optical wireless links can be expected to be available for more than 99% of the time on links up to 1 km in length, and high performance radio frequency (RF) provides backup connectivity in adverse conditions [19]. Hybrid system architectures that exploit the complementary nature of optical wireless and RF technologies have attracted significant research efforts in recent years [20–25]. In [20], experimental results of an FSO/RF network deployed in the five campuses of Ankara University were reported. Kukshya [21] characterized the link performance of a hybrid FSO and millimeter wave (MMW) system under different weather conditions.

In this paper, we address the issues of robustness and efficiency of optical wireless networks by considering the problems of topology optimization and survivability strategies in a joint manner. It is important to incorporate survivability strategies as well as optimal topology control into the network design to guarantee the high-speed data delivery performance of optical wireless networks. This gives rise to new optimization problems related to FSO networks that are different from existing optimizations of wireline optical fiber networks.

The design of a dynamic topology control which is adaptive to real-time network changes has been extensively studied for wavelength-division multiplexed (WDM) all-optical networks (see e.g., [26–29]). For example, a mesh-based traffic grooming ILP model for sub-wavelength traffic grooming over WDM networks was developed and a lightpath-based virtual topology was designed in [30]. The impact of number of transceivers at each switch node was also evaluated. However, for FSO networks, there are far fewer studies dedicated to topology design. In [31], Desai and Milner extended the optimization problem of topology control to the domain of optical wireless networks. The authors of [31] formulated the congestion minimization problem as a mixed integer program (MIP) to dynamically optimize the network topology in response to link state changes, and proposed various heuristics to find near optimal topologies to solve the problem in an efficient manner. However, they did not consider network survivability in their topology optimization formulation.

Network survivability is of paramount importance for an optical wireless network owing to its many challenges in free space physical transmission [3, 12, 13]. In topology design for an optical wireless network, it is necessary to jointly consider it with survivability. Extensive
research has been performed for the survivability of wired communication networks, which can be categorized into classes of pre-designed protection and online restoration [32–38]. Pre-designed protection schemes are able to immediately change the route to pre-allocated backup paths when network failures on the working paths are detected. By contrast, alternative paths are dynamically established in dynamic restoration schemes when the working path is affected. Therefore, pre-designed protection schemes can react quickly to network failures at the cost of reserving more spare capacity. Dynamic restoration schemes are more efficient than protection schemes in terms of the utilization of network resources, but take longer to recover network connections.

Meanwhile, due to the popularity of GMPLS/MPLS networks, much research has been focussed on GMPLS/MPLS network survivability. In [39], a simple path protection mechanism based on a reverse notification tree structure was proposed for MPLS networks. In [40], a survey on MPLS protection methods and their utilization in combination with online routing methods was presented. In [41], an approach called design-based routing (DBR) was proposed for GMPLS networks to employ optimized paths computed offline to guide online path setups. Also, for virtual topology design in the context of a WDM optical network, a comprehensive ILP model was developed for a dynamic survivable network jointly considering working and protection capacity. For fast online reconfiguration, a heuristic based on LP-relaxation to release the constraint of integer values for the variables of the original ILP model was also proposed [42].

Despite the importance of network survivability and its extensive investigation for a wired communication network, little work has been reported on designing a highly reliable FSO network. In this paper, we focus on this type of design. We employ the 1 + 1 protection technique to support FSO network survivability. Although 1 + 1 protection is well-studied for the traditional wired networks, the present study has several novel aspects. Firstly, in the traditional wired networks, a physical topology (e.g., a fiber network) that provides network capacity is often pre-defined. The 1 + 1 survivable design follows the existing link trail of the physical topology to choose link-disjoint working and protection path pairs for each pair of nodes. In contrast, in the present study, in addition to finding a pair of working and protection paths, we also aim to optimize the topology for the FSO network. Thus, the present problem is more challenging than the wired 1 + 1 survivable network design. Secondly, for the topology design, the traditional approaches often do not consider the physical length limitation of a topology link, which means that each node is not bounded by a nodal degree [30,42]. In contrast, in the current design problem, the nodal degree of each node is constrained by the transmission distance of an FSO link. (Only the neighboring nodes that are within the FSO link transmission distance are allowed to establish direct communication links with a node.)

In this work, we first simplify the optimization problem for topology control by linearizing the formulation of [31]. We then propose an optimal topology control strategy for survivable optical wireless networks by taking network survivability into account based on a path-oriented pre-designed protection approach. We consider only failures associated with a single FSO link. Simultaneous multiple link failures including node failures are beyond the scope of this paper. As we discuss below, an RF network will back-up the FSO network when multiple-link failure occurs, which may happen in the case of heavy fog.

The remainder of the paper is organized as follows. In Section 2, the network model and the assumptions used in this paper are described. The linearized formulation of the traditional problem of topology control based on congestion minimization is discussed in Section 3. In Section 4, we first formulate the linearized congestion minimization problem with working and protection paths for optional wireless networks, and analyze its complexity. We then introduce a restricted version of the problem in Section 5, in which only limited sets of candidate working and protection paths are considered. We include in an Appendix a multi-hour extension of this restricted problem. We also propose an algorithm to find the candidate working and protection paths for each OD pair. Numerical results and related discussions are provided in Section 6. The paper is concluded in Section 7.

2. Network model and assumptions

Consider a network of $N$ nodes as a digraph $G=(V,E)$, where $V$ and $E$ denote the sets of nodes and arcs, respectively. In the context of an FSO network, an arc represents an FSO link which is a directed bit pipe from a transmitting node to a receiving node. Such an FSO link uses a laser diode to transmit and a photodetector to receive multi-channel signals [43]. Therefore an FSO link carries many wavelengths and its total capacity can be used by many OD pairs simultaneously, each of which carries the traffic of many users.

It is assumed that the traffic matrix $R$ of size $N \times N$, where its $[s, d]$th entry $R_{sd}$ denotes the effective bandwidth for OD pair $[s, d]$, is known a priori. In Appendix A, we describe a method to derive effective bandwidth. We further assume that traffic of a given OD pair that uses an arc along the path can also use a subsequent arc if capacity is available there. Such an assumption requires that full wavelength conversion is available at any intermediate node along the path. Availability of full wavelength conversion is justified, as conventional FSO systems use optical-electronic-optical (OEO) conversion in the nodes [44], so wavelength conversion can be used for independent data-format transmission.

3. Linearized formulation for congestion minimization

The problem of congestion minimization has been studied in the context of logical topology design for wavelength-routed optical networks. Desai and Milner [31] extended the research to optical wireless networks, and formulated this problem as a non-linear MIP. To solve the non-linear MIP problem, the authors of [31] considered various heuristics, and sub-optimal solutions were found. However, the non-linear MIP can be in fact reformulated as a linear MIP that finds optimal solutions of the non-linear
MIP in significantly less computation time. We first define the indicator function $I_{ij}$ as:

$$I_{ij} = \begin{cases} 1, & \text{if nodes } i \text{ and } j \text{ can see each other,} \\
0, & \text{otherwise.} \end{cases}$$

We introduce the following notation.

- $x_{ij}^{sd}$: the amount of traffic on arc $e_{ij}$ from origin destination (OD) pair $[s, d]$.
- $b_{ij}$: binary decision variable represents the selection of arc $e_{ij}$.
- $\Delta_i$: the degree constraint on node $i$.

In the following, we reformulate the problem of [31] as a linear MIP.

**Objective function**: minimize $z$

**Subject to**:

1. Flow conservation at node $i$:

$$\sum_j I_{ij} x_{ij}^{sd} - \sum_j I_{ji} x_{ij}^{sd} = \begin{cases} +R_{id}, & \text{if } i = s \\
-R_{id}, & \text{if } i = d \\
0, & \text{otherwise} \end{cases}$$

   \[ \forall i \in V, \forall [s, d]. \]

2. Bound on variable $z$:

$$z \geq \sum_{[s, d]} x_{ij}^{sd}, \quad \forall (i, j) \in E.$$

3. Bound on traffic flow variables:

$$0 \leq x_{ij}^{sd} \leq R_{id} b_{ij}, \quad \forall (i, j) \in E, \forall [s, d].$$

4. Constraint on degree of each node:

$$\sum_j b_{ij} \leq \Delta_i, \quad \forall i \in V.$$

5. Bound on decision variable $b_{ij}$:

$$b_{ij} \in \{0, 1\}, \quad \forall (i, j) \in E.$$

The objective function of the new formulation is linear, and the number of integer and linear variables remains the same as in [31]. In such congestion minimization problems, the optimal topology and optimal routing are searched simultaneously. Throughout this paper, this scheme is referred to as “linearized congestion minimization” (LCM). In LCM the objective is to minimize the maximum amount of traffic on the links.

In practice, one might be more interested in finding the optimal topologies given the set of possible routes for each OD pair. In this work, we propose a simple algorithm to find sets of candidate working and protection paths for each OD pair, and then the congestion minimization is performed to find the optimal working and protection paths such that the congestion of the entire network is minimized. We refer to this scheme as “linearized congestion minimization with working and protection paths” (LCM–WP). The LCM scheme will be used to illustrate the efficiency of our proposed LCM–WP scheme, and the results of the evaluation will be presented later in Section 6.

4. Linearized congestion minimization with working and protection paths

Here we provide a formulation of the linearized congestion minimization problem with working and protection paths. In the proposed scheme, our optimization will select for each OD pair, one working path and one protection path (which is arc-disjoint with the working path) such that the overall selection of paths will globally optimize the objective function which is to minimize network congestion.

In the prespecified protection schemes, the working and protection paths for a given OD pair can operate in either “1 + 1” protection or “1:1” protection [36]. In the “1 + 1” approach, data are sent on the working paths and duplicate data are sent on the protection paths. The receiver can choose the stronger of the two signals. In “1:1” protection, no traffic is transmitted on the backup routes unless network failures are detected. In this work, we assume that our algorithms operate under the “1 + 1” protection scheme. We introduce the following additional notation.

- $e_{ij}$: an arc between two nodes $i$ and $j$.
- $x_{ij}^{sd,W}$: the amount of traffic on arc $e_{ij}$ of the working path.
- $x_{ij}^{sd,P}$: the amount of traffic on arc $e_{ij}$ of the protection path.
- $x_{ij}^T$: total traffic on arc $e_{ij} \in E$.

We now formulate the congestion minimization problem with working and protection paths. For a node-set $D \subseteq V$, we define the cut induced by $D$ in $G$ as

$$\delta_c(D) = \{e_{ij} \in E : i \in D, j \in V \setminus D\}$$

where $\delta_c(D)$ is the set of all arcs with one end-node in $D$ and other end-node in $V \setminus D$. Define:

$$\phi_{ij}^c = \begin{cases} 1, & \text{if arc } e_{ij} \text{ is an arc on the working path} \\
0, & \text{otherwise} \end{cases}$$

and

$$\gamma_{ij}^c = \begin{cases} 1, & \text{if arc } e_{ij} \text{ is an arc on the protection path} \\
0, & \text{otherwise} \end{cases}$$

The following is our MIP problem:

**minimize** $z$

**Subject to**:

For each OD pair $[s, d]$ and for any $[s - d]$ cut induced by $D \subseteq V$ in $G$ we need to have:

$$\chi_c^d(\delta_c(D)) = R_{id} \quad \forall D \subseteq V : s \in D, d \notin D, \emptyset \neq D \neq V$$

where for any $M \subseteq E$,

$$\chi_c^d(M) = \sum_{e_{ij} \in M} x_{ij}^{sd,W}.$$  \hspace{1cm} (1)

To ensure that at each $[s - d]$ cut, only one arc is a member of any working path for OD pair $[s, d]$ traffic, the following set of constraints is introduced:

$$\chi_d^c(\delta_c(D)) = 1 \quad \forall D \subseteq V : s \in D, d \notin D, \emptyset \neq D \neq V.$$  \hspace{1cm} (2)

where for any $M \subseteq E$,

$$\chi_d^c(M) = \sum_{e_{ij} \in M} \phi_{ij}^c.$$
We need to have a protection path as well for OD pair \([s, d]\). Similar to (1), for the protection path, we set the following constraints:

\[
\mathcal{X}^D_{\gamma}(\delta_c(D)) = R_{ad}
\forall D \subseteq V : s \in D, d \notin D, \phi^D \neq D \neq V,
\]

where for any \(M \subseteq E\),

\[
\mathcal{X}^D_{\lambda}(M) = \sum_{e_{ij} \in M} \gamma_{ij}.
\]

As in (2),

\[
\mathcal{X}^D_{\gamma}(\delta_c(D)) = 1
\forall D \subseteq V : s \in D, d \notin D, \phi^D \neq D \neq V,
\]

where for any \(M \subseteq E\),

\[
\mathcal{X}^D_{\gamma}(M) = \sum_{e_{ij} \in M} \gamma_{ij}.
\]

For each arc \(e_{ij}\), we have the following constraints for working paths

\[
\lambda_{ij}^{W, W} \leq \phi_{ij} R_{sd}, \quad \forall [s, d] \in G
\]

and the following set of constraints for protection paths

\[
\lambda_{ij}^{P, P} \leq \gamma_{ij} R_{sd}, \quad \forall [s, d] \in G.
\]

These constraints ensure that unless an arc has been selected to be a member of a path, no traffic flow could be assigned for that arc.

Since the working and protection paths for an OD pair \([s, d]\) need to be arc-disjoint, we introduce the following set of constraints:

\[
\phi_{ij} + \gamma_{ij} \leq 1, \quad \forall [s, d], \text{ and } \forall e_{ij} \in E.
\]

The total capacity required on each arc \(e_{ij}\) is

\[
\lambda_{ij} = \sum_{\forall [s, d] \in G} \lambda_{ij}^{W, W} + \sum_{\forall [s, d] \in G} \lambda_{ij}^{P, P}.
\]

The remaining constraints are:

\[
z \geq \lambda_{ij}, \quad \forall e_{ij} \in E.
\]

\[
\sum_j b_{ij} \leq \Delta_i, \quad \forall i \in V.
\]

\[
\phi_{ij} \leq b_{ij}, \quad \forall [s, d] \in G \text{ and } \forall e_{ij} \in E.
\]

\[
\gamma_{ij} \leq b_{ij}, \quad \forall [s, d] \in G \text{ and } \forall e_{ij} \in E.
\]

\[
\lambda_{ij} \geq 0, \lambda_{ij}^{W, W} \geq 0, \lambda_{ij}^{P, P} \geq 0, \quad \forall e_{ij} \in E \text{ and } \forall [s, d].
\]

\[
b_{ij}, \phi_{ij}, \gamma_{ij} \in \{0, 1\}, \quad \forall e_{ij} \in E, \forall [s, d].
\]

Now we have formulated an MIP problem in which we jointly consider optimal topology design and routing for optical wireless networks. The number of binary variables (\(\phi\) and \(\gamma\)) in this NP-hard optimization problem is of the order of \(N^4\). This is because we have \(O(N^2)\) binary variables for each arc and the total number of arcs is in the order of \(O(N^2)\) as well. Furthermore, the number of constraints presented by (2) and (3) is \(O(2^{N^2}N^2)\). This is explained by considering the fact that the total number of OD node pairs is \(O(N^2)\) and the number of constraints in (2) and (3) for each OD pair is \(O(2^{N^2}N^2)\) — this is because excluding the origin and destination nodes we have \(N - 2\) remaining nodes and total number of combinations of choosing all possible subsets out of a set of \(N - 2\) remaining nodes is \(2^{N^2}N^2\). We can conclude that such an optimization problem is very hard to solve for realistic problems where the value of \(N\) normally exceeds 20.

Therefore, we consider a restricted version of it, which we call the restricted MIP problem. In our restricted MIP problem, we use a limited set of potential working and protection paths, and subsequently decrease the number of binary variables to the order of \(O(N^2)\). This is due to the fact that the binary variables in this path optimization method are defined based on working and protection paths for each OD node pair. This implies that the number of such binary variables in our restricted case is in the order of \(O(N^2)\) and since \(k\) is bounded, then the total number of binary variables in the restricted case is in the order of \(O(N^2)\). Note that in both the arc-based and path-based optimization methods, we ignored the number of binary variables \(b_{ij}\) as the total number of such variables in both cases is dominated by the other binary variables (\(\phi\) and \(\gamma\)).

Also note that our paper addresses a pure FSO network and provides optimal design solutions that can cope with a failure of any single arc. For places and situations where it is expected that many arcs fail at once, e.g. heavy fog, a backup is necessary. One such possible backup can be achieved by directional RF connections. These RF connections are sensitive to different weather conditions and are thus complementary to the FSO network. For example, the FSO network is sensitive to heavy fog but can operate well in rain, while the RF is sensitive to heavy rain and operates well in fog. In case of heavy rain, where the RF connections cannot provide backup, still a single arc obstruction may occur in the FSO network and our proposed optimal design will guarantee survivability.

5. Restricted linearized congestion minimization problem

In this section, we first describe our algorithm that finds the candidate working and protection paths, and present its complexity analysis. We then provide a formulation for the restricted LCMinWP problem which searches the optimal working and protection paths for each OD pair.

Define

- \(W^d\): the set of working paths for OD pair \([s, d]\).
- \(w^{sd}_{ij}\): an arbitrary path in set \(W^d\).
- \(P^d\): the set of protection paths for OD pair \([s, d]\).
- \(p^{sd}_{ij}\): an arbitrary path in set \(P^d\).

5.1. The \(k\)-shortest paths problem

The \(k\)-shortest paths (KSP) problem is generalized from the shortest path problem, and has been well-studied in recent decades. Numerous algorithms have been proposed to find not one but several shortest paths \(45–48\). In this work, we use the approach proposed in \(47\) to find the candidates of working and protection paths for each OD pair. To find the shortest path, the Dijkstra’s algorithm is used as a main subroutine in the KSP approach proposed in \(47\).
5.2. Algorithm for finding candidate working and protection paths

We propose a new method to find a set of candidate working and protection paths for each OD pair in a given network. The purpose of the procedure is to identify a number of potential primary and backup paths for each OD pair, from which our optimization scheme is able to select the optimal working and protection paths such that the total network congestion is minimized.

Assuming it is required to find \( k \) candidate working paths and \( k \) candidate protection paths for OD pair \([s, d]\), we consider the following algorithm.

Step 1: Use Dijkstra’s algorithm to find the shortest path for OD pair \([s, d]\);
Step 2: Disable all the arcs on the shortest path found in Step 1;
Step 3: Use the KSP algorithm to find \( k \) protection paths for OD pair \([s, d]\), and store them to \( P^{sd} \);
Step 4: Disable all the arcs on all the \( k \) paths found in Step 3;
Step 5: Restore the shortest path found in Step 1;
Step 6: Use the KSP algorithm to find \( k \) working paths for OD pair \([s, d]\), and store them to \( W^{sd} \).

In our algorithm, the shortest path is always included as a working path candidate. In order to get high quality protection paths, the next \( k \) shortest paths (which are arc-disjoint with the shortest path) are allocated to the set of protection paths. Finally, the other \( k-1 \) shortest paths are obtained for the candidate working path by disabling all the arcs of the protection paths. In this way, it is guaranteed that any path in \( W^{sd} \) must be arc-disjoint with a path in \( P^{sd} \).

Without loss of generality, our analysis applies to the case where the numbers of candidate working and protection paths are different. For simplicity of notation, we consider here the case where the numbers of candidate working and protection paths are the same and it is denoted \( k \).

Existing work on node disjoint paths (e.g. [49]), may also be, generally speaking, used to obtain arc-disjoint paths. Since in FSO, the electronic/optical nodes are designed for a very high degree of availability, we are interested here in arc-disjoint paths and not in node disjoint paths. Furthermore, given the limitations of imperative LOS and transmission range for FSO links, the nodal degrees in FSO networks are typically not very high, and it may become infeasible to find, for example, two candidate working paths and two candidate protection paths for each OD pair that are node-disjoint.

The procedures of the algorithm are repeated for each OD pair of a given network (digraph) to find the required number of candidate working and protection paths. The networks are randomly generated and we assume that all arcs in a network are assigned the same weights. An example of the algorithm is illustrated in Fig. 3.

5.3. Analysis of algorithm complexity

In the following, we derive the complexities of our algorithm and show that it is polynomial. Let \( m \) denote the number of potential arcs in the network, i.e., \( m = |E| \).

Fig. 3. Algorithm for finding candidate working and protection paths. \( k = 2 \).

Step 1 of the algorithm has the complexity of Dijkstra’s algorithm implemented using the binary heap method, which has complexity of \( O(m + N \log N) \) to obtain the shortest path between any pair of nodes in the network. Steps 2, 4 and 5 have a complexity of \( O(m) \). Steps 3 and 6 are the more greedy parts of the algorithm, which implement repeatedly the shortest path algorithm many times. They have the complexity of \( O(m + N \log N + k) \) [47]. Therefore, the overall complexity of the algorithm is \( O(m + N \log N + k) \).

5.4. Formulating the restricted MIP problem

As stated in Section 4, our original optimization may not be solvable for practically sized problems. We propose here a restricted constraint space, and formulate the restricted MIP problem:

It is beneficial to have a fixed number of practical paths to be chosen as the set of candidate working and protection paths. Consider a network that is currently in operation. Every time when a new customer is added to the network, or an existing customer changes its usage significantly, then we need to solve the problem again. And if we do not dictate the feasible sets of working and protection paths, then according to the new circumstances, the algorithm may choose totally different working and protection paths from the ones before the change. Such dynamic routing implies the pre-emption of all (or most) of the existing paths and establishment of new working and protection paths at each iteration. This will cause service interruption and incur a significant administrative burden which is normally unacceptable to telecom companies. Furthermore, such an approach that gives freedom to the algorithm to select the paths would not provide very significant cost reduction. Notice that in our approach one may increase the sizes of the path sets to include good paths during optimization.

The restricted congestion minimization problem with working and protection paths in our approach can then be formulated as:

**Objective function:**

\[
\text{minimize } z
\]

(15)

The constraints of this optimization problem are described as follows:

5.4.1. Flow conservation constraints

To address the flow conservation constraints, we first define \( \Omega^{W}_{W} \) and \( \Omega^{P}_{P} \) as the set of nodes that appear as the
second node on a path in $W^d$ and $p^d$, respectively:

$$\Omega^d_W = \{v : (s, v) \in W^d_{sd} \}.$$  

Similarly, for the protection paths, we have

$$\Omega^d_p = \{v : (s, v) \in P^d_{sd} \}.$$  

The flow conservation constraints for working paths are as follows:

$$\sum_{v \in \Omega^d_W} \lambda_{sv} = R_{sd}. \quad (16)$$

Similarly, for protection paths,

$$\sum_{v \in \Omega^d_p} \lambda_{sv} = R_{sd}. \quad (17)$$

If a path is selected as a working or protection path for a specific OD pair, the traffic flow on an arbitrary arc (FSO link) of this path is the same as on any other arc of the same path. Therefore, on any working path $w^d = \{s, e_{sv}, v, e_{vg}, g, \ldots, u, e_{ud}, d\}$, where $e_{uv}$ denotes the arc between node $u$ and $v$, we have the following flow conservation constraints for working paths:

$$\lambda_{sv}^W = \lambda_{eg}^W = \cdots = \lambda_{ud}^W. \quad (18)$$

Similarly, on any protection path, $p^d = \{s, e_{sv}, v, e_{vg}, g, \ldots, u, e_{ud}, d\}$, we have,

$$\lambda_{sv}^P = \lambda_{eg}^P = \cdots = \lambda_{ud}^P. \quad (19)$$

5.4.2. Bounds on traffic flow variables

The following binary decision variable is defined:

$$\phi_{w}^{sd} = \begin{cases} 
1, & \text{if path } w^{sd} \text{ is selected as the working path} \\
0, & \text{otherwise.} 
\end{cases}$$

and

$$\gamma_{p}^{sd} = \begin{cases} 
1, & \text{if path } p^{sd} \text{ is selected as the protection path} \\
0, & \text{otherwise.} 
\end{cases}$$

Since we only choose one working path and one protection path for each OD pair, we use the following constraints:

$$\sum_{w} \phi_{w}^{sd} = 1. \quad (20)$$

and

$$\sum_{p} \gamma_{p}^{sd} = 1. \quad (21)$$

If a working path is not selected, all the traffic flow decision variables associated with that are set to zero:

$$\lambda_{sv}^W \leq \phi_{w}^{sd} R_{sd}, \quad \forall \text{ OD pair } [s, d] \quad (22)$$

where $e_{uv} \in w^{sd} \in W^{sd}$. Similarly, for protection paths,

$$\lambda_{sv}^P \leq \gamma_{p}^{sd} R_{sd}, \quad \forall \text{ OD pair } [s, d] \quad (23)$$

where $e_{uv} \in p^{sd} \in P^{sd}$.

5.4.3. Bound on variable $z$

The total traffic amount on arc $e_{uv}$ is given by:

$$\lambda_{uv} = \sum_{[s, d]} \lambda_{uv}^{sd, W} + \sum_{[s, d]} \lambda_{uv}^{sd, P}, \quad \forall e_{uv} \in E.$$  

The bound on variable $z$ is given by:

$$z \geq \lambda_{uv}. \quad (24)$$

5.4.4. Constraint on degree of each node:

$$\sum_{j} b_{ij} \leq \Delta_i, \quad \forall i \in V. \quad (25)$$

5.4.5. Bounds on decision variables

The decision variables are binary, thus we have

$$b_{ij} \in \{0, 1\}, \quad \forall e_{ij} \in E. \quad (26)$$

If a path $w^{sd}$ is chosen as the working path, $\phi_{w}^{sd}$ is set to one and the decision variables of all the arcs on this path will be set to one. On the other hand, if a path $w^{sd}$ is not chosen as the working path, $\phi_{w}^{sd}$ is set to zero, but it is still possible that working paths for other OD pairs include arc $e_{ij}$. Therefore, we have the following constraint:

$$\phi_{w}^{sd} \leq b_{ij}, \quad \forall e_{ij} \in w^{sd}, \quad (27)$$

and similarly for protection paths:

$$\gamma_{p}^{sd} \leq b_{ij}, \quad \forall e_{ij} \in p^{sd}. \quad (28)$$

As a matter of fact, the restricted problem provides a solution that bounds from above the optimal solution of the MIP (since the restricted problem has a smaller feasible region than its MIP equivalent). As we demonstrate in Section 6, the restricted MIP problem is solvable for problems of realistic size.

The level and the distribution of traffic tend to vary in a telecommunications network. The traffic variability may be partly attributable to the stochastic nature of customer demand and some of the variability may be due to change in usage pattern over the course of a day. During working hours, business usage dominates and in the evening residential usage dominates. A telecommunications network should be designed to take into account the multi-hour traffic profile and to be able to carry the traffic over all hours in a cost-effective manner [50–55]. Obtaining a minimum-cost network for multi-hour traffic profile is a large-scale optimization problem. We provide the multi-hour network design modeling for our restricted MIP in Appendix B.

6. Numerical evaluation and discussion

We now evaluate the proposed methods using numerical examples. The results were obtained with CPLEX 6.0 on a 3 GHz Pentium 4 with 1 GB RAM.

Networks up to 40 nodes with potential arcs (FSO links) were randomly generated. The degree constraint of node $\Delta_i$ is set to be six for all nodes in the networks. For a randomly generated network of $N$ nodes, we also randomly generated $N \times N$ traffic matrices with uniformly $[0, 1]$ chosen entries, each of which represents a possible effective bandwidth for OD pair $[s, d]$. The shortest paths are determined based on the number of hops; in other words, all arcs in the networks are assigned the same weights.
6.1. Comparison between LCM and the restricted LCM–WP

Fig. 4 shows the topology of an example network consisting of 15 nodes. The randomly generated traffic matrix of the network is shown in Table 1.

Figs. 5 and 6 demonstrate comparisons between LCM and our restricted LCM–WP schemes. Fig. 5 illustrates a comparison in terms of number of constraints when $N$ varies from five to 25. It shows that the number of constraints in the LCM grows exponentially as the network size $N$ increases. By contrast, the number of constraints in our LCM–WP schemes increases almost linearly. For instance, when $N$ is 20, the number of constraints in LCM–WP is less than one fifth of that in LCM.

The solution time of the LCM–WP schemes is significantly less than that of LCM. This is because the optimal routes are searched only within the path candidates in the restricted LCM–WP schemes. Fig. 6 illustrates a comparison of solution time between LCM–WP and LCM. It can be observed that the solution times are similar when the network size is small (e.g., $N = 5$). When the network size grows, the LCM scheme can take more than ten times longer than the LCM–WP schemes in most cases.

6.2. Validating optimality for problems of realistic size

Evaluation of the optimality of our results is a challenge for realistic size problems because the general MIP is not

---

Table 1

OD traffic matrix of a 15-node network.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0.55</td>
<td>0.88</td>
<td>0.46</td>
<td>0.74</td>
<td>0.94</td>
<td>0.04</td>
<td>0.27</td>
<td>0.57</td>
<td>0.70</td>
<td>0.80</td>
<td>0.96</td>
<td>0.25</td>
<td>0.99</td>
<td>0.40</td>
</tr>
<tr>
<td>2</td>
<td>0.15</td>
<td>0</td>
<td>0.69</td>
<td>0.67</td>
<td>0.35</td>
<td>0.36</td>
<td>0.87</td>
<td>0.46</td>
<td>0.33</td>
<td>0.19</td>
<td>0.17</td>
<td>0.13</td>
<td>0.07</td>
<td>0.12</td>
<td>0.81</td>
</tr>
<tr>
<td>3</td>
<td>0.11</td>
<td>0.41</td>
<td>0</td>
<td>0.04</td>
<td>0.71</td>
<td>0.14</td>
<td>0.84</td>
<td>0.07</td>
<td>0.65</td>
<td>0.61</td>
<td>0.06</td>
<td>0.43</td>
<td>0.90</td>
<td>0.84</td>
<td>0.13</td>
</tr>
<tr>
<td>4</td>
<td>0.80</td>
<td>0.54</td>
<td>0.51</td>
<td>0</td>
<td>0.82</td>
<td>0.61</td>
<td>0.20</td>
<td>0.41</td>
<td>0.71</td>
<td>0.40</td>
<td>0.67</td>
<td>0.82</td>
<td>0.01</td>
<td>0.36</td>
<td>0.59</td>
</tr>
<tr>
<td>5</td>
<td>0.40</td>
<td>0.36</td>
<td>0.25</td>
<td>0.88</td>
<td>0</td>
<td>0.01</td>
<td>0.25</td>
<td>0.43</td>
<td>0.99</td>
<td>0.30</td>
<td>0.64</td>
<td>0.61</td>
<td>0.43</td>
<td>0.51</td>
<td>0.55</td>
</tr>
<tr>
<td>6</td>
<td>0.74</td>
<td>0.89</td>
<td>0.31</td>
<td>0.50</td>
<td>0.70</td>
<td>0</td>
<td>0.83</td>
<td>0.04</td>
<td>0.90</td>
<td>0.91</td>
<td>0.82</td>
<td>0.04</td>
<td>0.07</td>
<td>0.66</td>
<td>0.39</td>
</tr>
<tr>
<td>7</td>
<td>0.93</td>
<td>0.38</td>
<td>0.48</td>
<td>0.63</td>
<td>0.15</td>
<td>0.09</td>
<td>0</td>
<td>0.31</td>
<td>0.68</td>
<td>0.39</td>
<td>0.81</td>
<td>0.92</td>
<td>0.11</td>
<td>0.30</td>
<td>0.02</td>
</tr>
<tr>
<td>8</td>
<td>0.37</td>
<td>0.25</td>
<td>0.67</td>
<td>0.36</td>
<td>0.50</td>
<td>0.42</td>
<td>0.79</td>
<td>0</td>
<td>0.31</td>
<td>0.15</td>
<td>0.14</td>
<td>0.72</td>
<td>0.51</td>
<td>0.99</td>
<td>0.07</td>
</tr>
<tr>
<td>9</td>
<td>0.04</td>
<td>0.57</td>
<td>0.64</td>
<td>0.51</td>
<td>0.26</td>
<td>0.55</td>
<td>0.66</td>
<td>0.64</td>
<td>0.75</td>
<td>0.43</td>
<td>0.89</td>
<td>0.35</td>
<td>0.92</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.39</td>
<td>0.10</td>
<td>0.30</td>
<td>0.74</td>
<td>0.73</td>
<td>0.55</td>
<td>0.65</td>
<td>0.75</td>
<td>0.05</td>
<td>0</td>
<td>0.12</td>
<td>0.54</td>
<td>0.46</td>
<td>0.04</td>
<td>0.67</td>
</tr>
<tr>
<td>11</td>
<td>0.05</td>
<td>0.89</td>
<td>0.72</td>
<td>0.75</td>
<td>0.24</td>
<td>0.32</td>
<td>0.20</td>
<td>0.79</td>
<td>0.23</td>
<td>0.51</td>
<td>0</td>
<td>0.96</td>
<td>0.94</td>
<td>0.96</td>
<td>0.54</td>
</tr>
<tr>
<td>12</td>
<td>0.23</td>
<td>0.31</td>
<td>0.42</td>
<td>0.40</td>
<td>0.49</td>
<td>0.26</td>
<td>0.66</td>
<td>0.31</td>
<td>0.27</td>
<td>0.82</td>
<td>0.93</td>
<td>0</td>
<td>0.10</td>
<td>0.42</td>
<td>0.05</td>
</tr>
<tr>
<td>13</td>
<td>0.65</td>
<td>0.39</td>
<td>0.74</td>
<td>0.46</td>
<td>0.99</td>
<td>0.57</td>
<td>0.71</td>
<td>0.72</td>
<td>0.38</td>
<td>0.77</td>
<td>0.98</td>
<td>0.82</td>
<td>0</td>
<td>0.73</td>
<td>0.62</td>
</tr>
<tr>
<td>14</td>
<td>0.90</td>
<td>0.17</td>
<td>0.34</td>
<td>0.15</td>
<td>0.39</td>
<td>0.67</td>
<td>0.38</td>
<td>0.38</td>
<td>0.22</td>
<td>0.56</td>
<td>0.04</td>
<td>0.14</td>
<td>0.84</td>
<td>0</td>
<td>0.01</td>
</tr>
<tr>
<td>15</td>
<td>0.96</td>
<td>0.13</td>
<td>0.81</td>
<td>0.01</td>
<td>0.60</td>
<td>0.03</td>
<td>0.08</td>
<td>0.25</td>
<td>0.53</td>
<td>0.95</td>
<td>0.73</td>
<td>0.41</td>
<td>0.37</td>
<td>0.22</td>
<td>0</td>
</tr>
</tbody>
</table>
Tables 2–7 demonstrate that in the cases considered, our working and protection paths described in Section 4 of our linearized congestion minimization problem with the optimality of our results, we consider a relaxed version solvable for such problems. Therefore, in order to evaluate the optimality of our results, we consider a relaxed version of our linearized congestion minimization problem with working and protection paths described in Section 4 which we call the relaxed MIP. In our relaxed MIP we consider the MIP of Section 4, but removing a set of constraints (7) that guarantees that the working and protection paths for an OD pair \([s, d]\) are arc-disjoint. We know that a solution of the relaxed MIP provides a lower bound to our optimal solution. To estimate the solution of the relaxed MIP we use a modified version of our algorithm described in Section 5 where we do not aim to achieve that the working and the protection paths for an OD pair \([s, d]\) are arc-disjoint. We call this modified algorithm the relaxed MIP approximation (RMA) algorithm.

This procedure of the RMA algorithm is described as follows.

**Step 1:** Use the KSP algorithm to find \(k\) working paths for OD pair \([s, d]\), and store them to \(W^sd\).

**Step 2:** Disable all the arcs on the shortest path found in Step 1;

**Step 3:** Use the KSP algorithm to find \(k\) protection paths for OD pair \([s, d]\), and store them to \(P^sd\).

In the RMA, the shortest path is always reserved as a candidate working path, and a working path may share the same arc with a protection path. The RMA algorithm should not be used to find optimal working and protection paths for an OD pair, as we do not have protection in a situation where failures happen to a common arc.

The RMA algorithm for a large \(k\) value will find an optimal solution which approximates the optimal solution of the Relaxed MIP problem because for large \(k\) values, almost all paths are allowed and there is no restriction of having arc-disjoint paths. Running the RMA algorithm for a large \(k\) is not computationally difficult. Therefore, we increase the value of \(k\) until the marginal benefit diminishes, or until the running times are excessive. This provides an approximation to the relaxed MIP problem. Then we compare the value of the objective function obtained by our original algorithm with the values obtained by the RMA algorithm. The results are presented in Tables 2–7. In these tables the values of the best objective function for each problem has been provided with the corresponding error bound (EB) provided by the CPLEX optimizer package. EB is the maximum error of the corresponding objective function value. Please note that the length of running time of the optimizer package for each problem was less that 20 min in all cases.

Table 2 demonstrates that in the cases considered, our original algorithm does provide objective function values that are close to those of the RMA algorithm. The results in Tables 2–7 demonstrate that our heuristic approach provides better solutions (lower objective function values) as \(k\) increases. Please note that due to non-scalability of the

### Table 2

The impact of \(k\) on the values of the objective function for \(N = 5\). Here EB represents the maximum error bound on the objective function value.

<table>
<thead>
<tr>
<th>Instance</th>
<th>Results</th>
<th>(k = 2)</th>
<th>(k = 3)</th>
<th>(k = 4)</th>
<th>(k = 5)</th>
<th>Original</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Z</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.753</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>2</td>
<td>Z</td>
<td>1.07</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.756</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>3</td>
<td>Z</td>
<td>1.79</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>4</td>
<td>Z</td>
<td>0.98</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
<td>0.691</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>5</td>
<td>Z</td>
<td>1.53</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.776</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>6</td>
<td>Z</td>
<td>1.25</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.722</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>7</td>
<td>Z</td>
<td>0.77</td>
<td>0.77</td>
<td>0.77</td>
<td>0.77</td>
<td>0.657</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>8</td>
<td>Z</td>
<td>1.47</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.817</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>9</td>
<td>Z</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>10</td>
<td>Z</td>
<td>1.53</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

### Table 3

The impact of \(k\) on the values of the objective function for \(N = 10\).

<table>
<thead>
<tr>
<th>Instance</th>
<th>Results</th>
<th>(k = 2)</th>
<th>(k = 3)</th>
<th>(k = 4)</th>
<th>(k = 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Z</td>
<td>2.87</td>
<td>1.76</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>2</td>
<td>Z</td>
<td>2.43</td>
<td>2.17</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>3</td>
<td>Z</td>
<td>3.02</td>
<td>1.05</td>
<td>1.05</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>10.71%</td>
</tr>
<tr>
<td>4</td>
<td>Z</td>
<td>2.89</td>
<td>2.89</td>
<td>1.57</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>9.02%</td>
</tr>
<tr>
<td>5</td>
<td>Z</td>
<td>3.07</td>
<td>1.19</td>
<td>0.99</td>
<td>9.9</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>6</td>
<td>Z</td>
<td>1.98</td>
<td>1.15</td>
<td>0.95</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>7</td>
<td>Z</td>
<td>2.10</td>
<td>1.08</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>8</td>
<td>Z</td>
<td>3.21</td>
<td>2.12</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>9</td>
<td>Z</td>
<td>2.98</td>
<td>1.01</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>10</td>
<td>Z</td>
<td>4.4</td>
<td>1.93</td>
<td>1.07</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td>EB</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>
original problem we did not perform its implementations for larger $N$ values as the number of constraints in the original modeling is computationally prohibitive.

It is also important to mention that in the case of the 40-node restricted MIP problem, the objective function values corresponding to the feasible integral solutions provided by the CPLEX optimizer package within the starting minutes of running time had an error bound value of no more than 10%. This behavior was not observed for smaller networks. This indicates a potential to develop heuristics for very large networks for which optimal solution is not obtainable. The results presented in Table 7 for the case of a 40 node network ($k = 40$) particularly demonstrate that our method is scalable to provide optimal solutions for realistic networks.

### 7. Conclusion

We have proposed an optimal topology control scheme, called LCM–WP, for minimizing congestion in optical wireless networks, in which topology and routing optimization strategies are jointly considered. We then formulated a linearized congestion minimization problem in the form of MIP that aims to select optimal working and protection routes for each OD pair such that the total network congestion is minimized.

As the MIP is not scalable, we proposed to use a restricted version of the MIP problem where the sets of working and protection paths are limited, and introduced an algorithm that efficiently chooses candidate working and protection paths for each OD pair in the network.

We have evaluated the efficiency of our proposed schemes in comparison to the LCM scheme. As compared
to LCM, in our LCM–WP scheme the number of constraints is significantly lower and hence more tractable for real size problems. The implementation results show that the value of the objective function decreases (improves) as $k$ increases (granting more routing flexibility).

Acknowledgements

The authors wish to thank HCS Technologies for their permission to allow us to publish the image presented in Fig. 1. The authors also extend their thanks to Wendy Zukerman for providing the photograph of Hong Kong used in Fig. 2. The work of I. Ouveysi, G. Shen and M. Zukerman was partly supported by the Australian Research Council and the work of F. Shu was partly supported by the National ICT Australia, Victoria Laboratory. The work described in this paper was also partially supported by a grant from City University of Hong Kong (Project No. 9380044).

Appendix A. Effective bandwidth evaluation

Analysis of measured traffic streams taken from a wide range of sources has shown that traffic streams exhibit Long Range Dependent (LRD) and self-similar characteristics in VBR video [56,57], Ethernet traffic [58,59], and metropolitan area traffic [60]. It has been shown that the Gaussian model accurately represents aggregation of many traffic streams due to the central limit theorem [61]. One way to evaluate effective bandwidth is to use the results in [61] to derive an optimal service rate value independently for any given OD pair. We introduce the following notation.

- $S_{sd}$: a decision variable that represents a service rate for OD pair $[s, d]$ which we aim to optimize.
- $R_{sd}$: the effective bandwidth, which represents the minimal value of the service rate subject to a quality of service (QoS) constraint.
- $D_{sd}$: the mean variable traffic that arrives for OD pair $[s, d]$.
- $\sigma_{sd}$: the standard deviation of the variable traffic that arrives for OD pair $[s, d]$.
- $m_{sd}$: the mean of the net traffic input, i.e., $m_{sd} = D_{sd} - S_{sd}$.
- $t$: a buffer content threshold (expressed in data units, e.g., packets), which is used to define buffer overflow.
- $P_{sf}$: overflow probability, which is defined as the probability that the number of data units in the buffer exceeds a specified threshold $t$.
- $H$: Hurst parameter.

The overflow probability in the LRD case based on the Gaussian model is approximated by [61]:

$$P_{sf}(t) = \frac{\sqrt{2}\pi}{\sigma_{sd}} \psi(-m_{sd}, \sigma_{sd})e^{s^*(t)t}$$

where

$$s^*(t) = -\frac{1}{2\sigma_{sd}^2} |1 - H|^{-2} \left(\frac{H}{|m_{sd}(1 - H)|}\right)^{2H} t^{1-2H}$$

and

$$\psi(x, \sigma_{sd}) = \frac{\sigma_{sd}}{\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_{sd}^2}} - \frac{x}{2\sigma_{sd}\sqrt{2}} \sqrt{\pi} e^{-\frac{x^2}{4\sigma_{sd}^2}}.$$  

It has been validated that (29) is accurate for a stationary LRD Gaussian queue where $0.5 < H < 1$ [62]. We consider only the busiest hour in a day and thus assume the arrival process is stationary. We define the following optimization problem to obtain the effective bandwidth $R_{sd}$.

1. Objective function: $R_{sd} = \min S_{sd}$
2. Subject to: $P_{sf}(t) \leq \epsilon_{qos}$

where $\epsilon_{qos}$ is the preassigned QoS measure, e.g., $\epsilon_{qos}$ can be set to $10^{-5}$. Since $s^*(t) < 0$, $P_{sf}(t)$ is a monotonically decreasing continuous function, the optimal value $R_{sd}$ can be obtained by bisection.

Appendix B. Multi-hour optimization model

While LRD traffic modeling is appropriate for dimensioning links and evaluating effective bandwidth in timescales of micro seconds to a few hours, for longer timescales, metropolitan area networks traffic exhibit multi-hour behavior [50]. This non-stationary periodic traffic behavior is a result of people movements. While intra CBD traffic and suburbs-CBD traffic peak during the business hours, it almost diminishes in the evening when traffic shifts to suburbs. A common methodology to dimension such networks is based on multi-hour traffic profiles. Accordingly, we may consider only two time intervals during a day. Namely, the busiest hour during business hours and the busiest residential traffic activity. We may consider additional time intervals if traffic intensity, for example, in the weekend, justifies their considerations. In the formulation below we assume a general number of time intervals.

We define the following decision variables:

- $x_{ij,t}^{sd,W}$: the amount of traffic on arc $e_{ij}$ of working paths $w_{ij}^{sd}$ at time interval $t$.
- $x_{ij,t}^{sd,P}$: the amount of traffic on arc $e_{ij}$ of protection paths $p_{ij}^{sd}$ at time interval $t$.
- $\lambda_{ij}^t$: total traffic on arc $e_{ij}$ in $E$ at time interval $t$.

Assume that end-to-end effective bandwidth requirements are provided for $N_t$ time intervals. This implies that we have $N_t$ matrices denoted as

$$R^t = (R_{sd}^t) .$$

The multi-hour optimization is formulated as follows.

minimize $z$

Subject to:
Flow conservation constraints

The flow conservation constraints for working paths for time interval $t$ are as follows:

$$\sum_{v \in V^d_{W_{sd,t}}} x_{v_{sd},t}^d = R_{sd}^t. \quad (33)$$

Similarly, for protection paths,

$$\sum_{v \in V^d_{P_{sd,t}}} x_{v_{sd},t}^d = R_{sd}^t. \quad (34)$$

We have the following flow conservation constraints for any working path $w^d = \{s, s_{e_{st}}, v, e_{vg}, g, \ldots, u, e_{ud}, d\}$:

$$\lambda_{v_{sg},t}^d = \lambda_{v_{ug},t}^d = \cdots = \lambda_{v_{ud},t}^d. \quad (35)$$

Similarly we have the following flow conservation constraints for any protection path, $p^d = \{s, v_{e_{st}}, u, e_{ug}, g, \ldots, u, e_{ud}, d\}$:

$$\lambda_{v_{sg},t}^d = \lambda_{v_{ug},t}^d = \cdots = \lambda_{v_{ud},t}^d. \quad (36)$$

Bound on traffic flow variables

The following binary decision variables are defined:

$$\phi_{w_{sd},t}^d = \begin{cases} 1, & \text{if path } w^d \text{ is selected as the working path for OD pair } [s, d] \text{ at time interval } t. \\ 0, & \text{otherwise.} \end{cases} \quad (37)$$

$$\gamma_{p_{sd},t}^d = \begin{cases} 1, & \text{if path } p^d \text{ is selected as the protection path for OD pair } [s, d] \text{ at time interval } t. \\ 0, & \text{otherwise.} \end{cases} \quad (38)$$

Since we only choose one working path and one protection path for each OD pair, we use the following constraints:

$$\sum_w \phi_{w_{sd},t}^d = 1 \quad \sum_p \gamma_{p_{sd},t}^d = 1. \quad (39)$$

If a working path is not selected, all the traffic flow decision variables associated with that are set to zero:

$$\lambda_{w_{sd},t}^d \leq \phi_{w_{sd},t}^d R_{sd}, \quad \forall \text{ OD pair } [s, d]. \quad (40)$$

where $e_{wv} \in w^d \in W^d$. Similarly, for protection paths,

$$\lambda_{w_{sd},t}^d \leq \gamma_{p_{sd},t}^d R_{sd}, \quad \forall \text{ OD pair } [s, d]. \quad (41)$$

where $e_{uv} \in p^d \in P^d$.

Bound on $z$

The total traffic amount on arc $e_{uw}$ is given by:

$$\lambda_{uw}^t = \sum_{[s,d]} \lambda_{w_{sd},t}^d, \quad \forall e_{uv} \in E. \quad (42)$$

The bound on variable $z$ is given by:

$$z \geq \lambda_{uw}^t. \quad (43)$$

Constraint on degree of each node:

$$\sum_j b_j \leq \Delta_i, \quad \forall i \in V. \quad (44)$$

Bounds on decision variables

$$b_j \in \{0, 1\}, \quad \forall e_{ij} \in E. \quad (45)$$

$$\phi_{w_{sd},t}^d \leq \varepsilon_{e_{ij}}^d, \quad \forall e_{ij} \in w^d. \quad (46)$$

$$\gamma_{p_{sd},t}^d \leq b_j, \quad \forall e_{ij} \in p^d. \quad (47)$$

$$z, \lambda_{uw}^t, \Phi_{w_{sd},t}^d, \gamma_{p_{sd},t}^d \geq 0. \quad (48)$$

References


