Multicast routing and wavelength assignment with delay constraints in WDM networks with heterogeneous capabilities ☆

Ming-Tsung Chen[a], B.M.T. Lin[b,c],*, Shian-Shyong Tseng[a]

[a]Department of Computer and Information Science, National Chiao Tung University, Hsinchu 300, Taiwan, ROC
[b]Department of Information and Finance Management, National Chiao Tung University,
Hsinchu 300, Taiwan, ROC
[c]Institute of Information Management, National Chiao Tung University, Hsinchu 300, Taiwan, ROC

Abstract

Because optical wavelength division multiplexing (WDM) networks are expected to be realized for building up backbone in the near future, multicasting in WDM networks needs to be addressed for various network applications. This paper studies an extended multicast routing and wavelength assignment (RWA) problem called multicast routing and wavelength assignment with delay constraint (*MRWA-DC*) that incorporates delay constraints in WDM networks having heterogeneous light splitting capabilities. The objective is to find a light-forest whose multicast cost, defined as a weighted combination of communication cost and wavelength consumption, is minimum. An integer linear programming (ILP) model is proposed to formulate and solve the problem. Experimental results show that using *CPLEX* to solve the ILP formulation can optimally deal with small-scale networks. Therefore, we develop a heuristic, *near-k-shortest-path heuristic* (*NKSPH*), to solve the problem in large-scale networks. Numerical results indicate that the proposed heuristic algorithm can produce approximate solutions of good quality in an acceptable time.
© 2006 Elsevier Ltd. All rights reserved.

Keywords: WDM network; Multicasting; Wavelength assignment; Delay bound; ILP; Heuristics

1. Introduction

Based on optical technology in optical networks (Green, 1992), a high-capacity telecommunication network can be constructed to provide routing, grooming, and restoration at wavelength level. The technology of wavelength division multiplexing (WDM) network (Lowe, 1998), based on optical wavelength-division multiplexing on optical fibers of optical network to form multicommunication channels at different wavelengths with electronic processing speed, provides connectivity among optical components to let optical communication meet the increasing demands for high channel bandwidth and low transmission delay. The utilization of wavelength for routing data is referred as *wavelength routing*, and an optical switch employing the technique is called a *wavelength-routing switch*. Therefore, in a wavelength-routing WDM network constructed using optical fiber links to connect the input ports and the output ports of wavelength-routing switches, data can be routed to other optical switches based on wavelengths of optical fibers. If the transmission between the input port and the output port of a switch involves two different wavelengths, the switch must be able to perform wavelength conversion.

In the (wavelength-routing) WDM network, a *light-path* (Chlamtac et al., 1992), a connection based on wavelength to carry data without optical-to-electrical conversion, would be set up in a way similar as circuit-switched networks to transmit data among (wavelength-routing) switches. The collection of light-paths is referred as a logical topology of a WDM network for transmitting optical signals. The cost of utilized wavelengths and the delay time of transmitting optical signal to a destination by a light-path are referred as *communication cost* and *transmission delay* of the light-path, respectively. The communication cost may depend on the numbers or the costs of fibers and switches used for establishing the connection.

Many network applications, such as videoconferencing, video on demand system, real-time control, on-line shopping, gaming, stock exchanging, and so on, have inspired new communication models. *Multicasting* is one of the most important models used to send data (messages) from a single source to multiple destinations. So far, two schemes (Zhang et al., 2000), *multiple-unicast* and *multicast*, have been employed to route data. The multiple-unicast scheme is a virtual topology consisting of a set of light-paths from the source to all destinations, where the number of light-paths equals the number of destinations. If there exists one link shared by more than one light-path, each light-path would need one different wavelength for routing data. As shown in Fig. 1(a), two light-paths, v_1–v_2–v_3 and v_1–v_2–v_4, would need two different wavelengths λ_1 and λ_2 because the
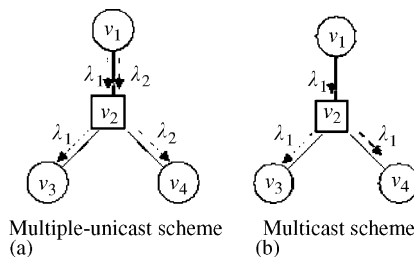

Fig. 1. Two multicast communication schemes: (a) multiple-unicast scheme and (b) multicast scheme.

link between v_1 and v_2 is shared. If each light-path requires one specific wavelength, the wavelength consumption may become unaffordable. The multicast scheme is thus proposed to reduce wavelength consumption.

As for providing the multicast scheme, a switch with or without *light splitting capability*, referred as *MC* (*multicast capable*) node or *MI* (*multicast incapable*) node (Zhang et al., 2000), can or cannot split a (optical) signal of input port to multiple signals of output port without optical-to-electrical conversions. The split signals can be transmitted by links to other switches concurrently. The light splitting capacity of a switch is used to describe the maximum number of split signals in an output port; that is, the light splitting capacity of an MC node (respectively, MI node) is greater than (respectively, equal to) 1. Therefore, locating an MC node for routing data to several destinations would have significant wavelength saving over the multiple-unicast scheme. As shown in Fig. 1(b), since v_2 is an MC node, only the wavelength λ_1 is required for routing data to v_3 and v_4 and the wavelength λ_2 can be saved. As defined above, the trail of routing data may be a light-tree (Sahasrabuddhe and Mukherjee, 1999) in which each internal node is an MC node. Similar to a light-path which is a logical topology to route data, a light-tree is a tree without infeasible node whose number of outbound edges is greater than its light splitting capacity. If all nodes in a network are MC nodes, one light-tree may be sufficient for routing data to all destinations; otherwise, a set of light-trees, referred as a *light-forest*, may be required for the network with sparse light splitting in which some of the nodes are MC nodes. The problems of finding light-paths in the multiple-unicast scheme and finding a light-tree or a light-forest in the multicast scheme are termed as a Routing Problem and a Multicast Routing Problem, respectively.

For solving the two problems, several algorithms (Kadaba and Jaffe, 1983; Ballardie, 1996; Rouskas and Baldine, 1997) and one protocol (Kosiur, 1998) have been proposed in the literature for traditional networks. Several heuristics (Zhang et al., 2000; Sreenath et al., 2001; Jia et al., 2001; Chen and Wang, 2002; Chen and Tseng, 2002; Yang and Liao, 2003) and integer linear programming (ILP) formulations (Krishnaswamy and Sivarajan, 2001; Kumar and Kumar, 2002; Yan et al., 2001) have also been proposed for WDM networks. Communication cost and wavelength consumption are usually discussed so as to evaluate the efficiency of the routes for providing high quality of service (QoS). Moreover, to guarantee that video and audio signals can be efficiently transmitted in interactive multimedia applications, transmission delay from the source to all destinations for routing the type of data will be limited under a given delay bound. The delay bound may be decided according to the emergence degree or priority of the data. Therefore, transmitted data with delay bounds reflects the realistic demand in the future. For the multicast routing problem in WDM networks with sparse light splitting and without wavelength conversion, Yan et al. (2001) proposed an ILP formulation and a tabu search algorithm for minimizing the number of fibers used. Sreenath et al. (2001) studied the multicast routing problem in WDM networks with sparse light splitting and wavelength conversion, but the cost of wavelength conversion was ignored. The Multicast Routing under Delay Constraint Problem (*MRDCP*) in WDM networks, proposed in (Chen and Tseng, 2002), takes into account MC nodes with different light splitting capacities and delay bounds so as to minimize a linear combination of communication cost and wavelength consumption. The problem of determining a light-path to connect a source and a destination and allocating the same wavelength on all links in the light-path to transmit data is referred as the Routing and Wavelength Assignment (*RWA*) problem in WDM networks. In

Krishnaswamy and Sivarajan (2001) and Kumar and Kumar (2002) ILP models were used to solve the RWA problem.

Another problem of determining one or more light-trees to transmit data from a source to all destinations and allocating the same wavelength on all links in the light-tree is referred as the multicast routing and wavelength assignment (*MRWA*) problem in WDM networks. When all nodes provide a light splitting capability, one light-tree is sufficient to transmit the request (Jia et al., 2001; Yan et al., 2001; Chen and Wang, 2002); otherwise, a set of light-trees is required (Zhang et al., 2000; Yang and Liao, 2003). To reduce the complexity, the problem is separated into multicast routing and wavelength assignment. In Jia et al. (2001) and Chen and Wang (2002), approximation algorithms were developed. Chen and Wang (2002) considered both the wavelength cost and the conversion cost and proposed an integrated approximation algorithm to solve the *MRWA* problem in WDM networks with wavelength conversion. It had also been proved that wavelength assignment on a light-tree can be solved in $O(n\gamma)$, where *n* is the number of nodes in the tree and γ is the number of wavelengths provided on each link. Therefore, this problem is not NP-hard. The *MRWA* problem for routing a request with delay bound was solved by Jia et al. (2001) under the assumption that every node in a network has a light splitting capability. Two integrated algorithms were proposed to minimize the sum of wavelength cost and communication cost. For the *MRWA* problem in WDM networks with sparse light splitting and without wavelength conversion, Zhang et al. (2000) proposed multicast routing heuristics to find a light-forest for routing a request without delay bounds. They also described how wavelengths were assigned for these light-trees in the light-forest. For routing on a network with power splitters having full range wavelength conversion and with wavelength converters having an unlimited splitting capacity, a mixed ILP (MILP) was proposed by Yang and Liao (2003) to solve the *RWA* of light-trees with delay bound. In the paper, the object was not only to minimize used fibers and to obtain the optimal placement of power splitters but also to design the logical topology based on light-trees for multiple connection demands. In Yang and Liao (2003), based on the assumption that a multicast request is routed only by a light-tree, it is possible that no light-tree can be found to satisfy the delay bound constraint and to cover all destinations in the network without enough power splitters or enough wavelength converters.

In Zhang et al. (2000), Sreenath et al. (2001), Jia et al. (2001), Yan et al. (2001), Chen and Wang (2002) and Yang and Liao (2003), the capacity of splitting the input signal of MC nodes was not discussed. They assumed that the MC nodes have the capability of splitting an input signal into multiple output signals, where the number of output signals is equal to the number of connected links. Due to the complicated architectures (Green, 1992) of MC nodes, using MC nodes with a superior light splitting capacity to build a WDM network is usually more expensive than those with inferior light splitting capacity or MI nodes. Therefore, nodes providing different light splitting capacities seem more realistic in network deployment. In this paper, we are concerned with the *WDM network with heterogeneous capabilities* (*WDM-He* network), in which the light splitting capacities of all nodes can be different.

From the above survey, we know that few studies have been done on the *MRWA* problem for routing a request with delay bounds in a WDM-He network with or without wavelength conversion for minimizing the total cost incorporating communication cost and wavelength consumption. To better provide a realistic objective function to reflect the cost for routing a request, we consider a linear combination of communication cost and

wavelength consumption, $\alpha \times$ communication_cost) $+ \beta \times$ wavelength_consumption. This objective is called the *multicast cost* hereafter in this paper. Notice that communication cost ratio α and wavelength consumption ratio β can be appropriately chosen based upon the topology and load of the network.

The multicast routing and wavelength assignment with delay constraint (MRWA-DC) problem is to find an optimal light-forest with minimum multicast cost and assigning wavelengths to the light-trees in the light-forest for routing a request under a given delay bound in a WDM-He network. In this paper, the *MRWA-DC* problem will be formulated in an ILP model. For large-scale networks, the *MRWA-DC* problem cannot be solved in a reasonable execution time. Therefore, an efficient *near-k-shortest-path heuristic* (*NKSPH*) is proposed to produce approximate solutions in polynomial time. Two sets of experiments, comparisons among different wavelength consumption ratios and performance assessment of the ILP, are implemented by *CPLEX* to study the performance of the ILP formulation. The other two experiment sets, comparisons among different values of k in *NKSPH* and comparisons between *NKSPH* and ILP, are tested using the same input data. The experimental results show that *NKSPH* provides quality approximate solutions to the *MRWA-DC* problem in large-scale networks.

The remainder of this paper is organized as follows. In Section 2, the *MRWA-DC* problem is formally defined. In Section 3, an ILP formation is proposed and proved to solve the problem. A simple heuristic *NKSPH* is proposed in Section 4 for developing approximate solutions. Section 5 presents the experiments for the ILP formation and *NKSPH*, and Section 6 gives some concluding remarks.

2. Problem definition

The following assumptions are given for the problem studied in this paper:

(1) The WDM network is an arbitrary connected graph.
(2) All links in the network are directed and provide the same set of wavelengths.
(3) Some of the nodes are MC nodes with probably different light splitting capacities.
(4) No node provides wavelength conversion.
(5) All connected demands are serialized such that a request is processed at a time.

A weighted graph $G = (V, E, M, \theta, c, d, w)$ is used to represent a WDM-He network with switch set $V = \{v_1, v_2, \ldots, v_n\}$, directed optical link set $E = \{e_1, e_2, \ldots, e_m\}$, and wavelength set $M = \{\lambda_1, \lambda_2, \ldots, \lambda_\gamma\}$. Function $\theta: V \rightarrow \mathbb{N}$ defines the light splitting capacity of switches, function $c: E \rightarrow \mathbb{R}^+$ defines the communication cost of links, and function $d: E \rightarrow \mathbb{R}^+$, specifies the transmission delay over links. Binary function $w: (E, M) \rightarrow \{0, 1\}$ is used to dictate whether a wavelength is used over a link. In graph G, there are n nodes, m edges, and γ wavelengths in each edge. For some node $v_i \in V$, $1 \leqslant i \leqslant n$, v_i is an *MC* node when $\theta(v_i) > 1$; otherwise, $\theta(v_i) = 1$. Moreover, the light splitting capacities of *MC* nodes may be different. For some λ_k over link e_j, $1 \leqslant k \leqslant \gamma$, $1 \leqslant j \leqslant m$, $w(e_j, \lambda_k) = 1$ indicates that wavelength λ_j can be used to route data; otherwise, $w(e_j, \lambda_k) = 0$.

A request r with a delay bound Δ is represented as (s, D, Δ) with destination set $D = \{\{d_1, d_2, \ldots, d_q\}$ and indicates that the data needs to be routed from a certain source s to all destinations d_l, $1 \leqslant l \leqslant q$, where $s \in V$, $D \subseteq V - \{s\}$ is a set of destinations, $|D| = q$, and the transmission delay of routing data to all destinations must be bounded by the delay

bound Δ. For different sources, destinations and emergence levels, the delay bounds may be different. A tighter bound will result in fewer candidate routes and make the request more likely to be suspended. For most of the cases, the delay bound of a request may be determined through previous experiences concerning the specified source, destinations, and application domain.

Suppose a light-tree uses wavelength λ_k to route data to some or all destinations in D. Denote such a scenario by (T_k, λ_k). The following conditions will be satisfied:

(1) each leaf node is reachable from the root s;
(2) $v_i \in V(T_k)$, in-degree $in(T_k, v_i) = 1$;
(3) $v_i \in V(T_k)$, out-degree $out(T_k, v_i) \leqslant \theta(v_i)$;
(4) $e_j \in E(T_k)$, $w(e_j, \lambda_k) = 1$;

where $V(T_k)$ and $E(T_k)$ are the nodes and the edges in tree T_k, and $in(T_k, v_i)$ and $out(T_k, v_i)$ represent the number of inbound edges to node v_i and the number of outbound edges from node v_i in T_k. The first condition ensures that T_k is a connected graph rooted at s. The second condition ensures that T_k is a tree. A light-tree can be viewed as a routing topology from the root; therefore, the third condition ensures that each internal node must provide light splitting capacity sufficient for splitting the input signal to transmit the signal to all associated nodes. The fourth condition ensures that the same wavelength used in each link of T_k can be used to route data. The communication cost and the transmission delay of T_k for routing the data to all destinations are denoted by $c(T_k)$ and $d(T_k)$ and can be written as

$$\text{communication cos t}: \quad c(T_k) = \sum_{e_j \in E(T_k)} c(e_j),$$

$$\text{transmission delay}: \quad d(T_k) = \max_{d_l \in D \cap V(T_k)} d(P_{T_k}(s, d_l)),$$

where $P_{T_k}(s, d_l)$ represents a light-path from s to d_l in T_k and $d(P_{T_k}(s, d_l)) = \sum_{e_j \in P_{T_k}(s, d_l)} d(e_j)$.

Because we consider the case where only one request is processed at a time, an *MRWA-DC* problem instance can be represented by (G, r). There can be several light-trees associated with different wavelengths to satisfy the request. The set of derived light-trees is called a light-forest and denoted by $\Gamma = \{(T_k, \lambda_k) | 1 \leqslant k \leqslant \gamma\}$. The case T_k is empty implies that λ_k is not used for request r. Forest Γ can be viewed as a *feasible solution* of (G, r) or a *feasible light-forest* when the following conditions are satisfied:

(1) destination constraint: $D \subseteq \bigcup_{k=1}^{\gamma} V(T_k)$;
(2) delay constraint: $d(T_k) \leqslant \Delta$, $1 \leqslant k \leqslant \gamma$.

The destination constraint and the delay constraint describe that all destinations will be reached and that the transmission delays of all light-trees are bounded by Δ, respectively. If no feasible light-forest can be obtained, then the request cannot be successfully met. Because each non-empty light-tree in Γ corresponds to a unique wavelength, the number of non-empty light-trees in a feasible forest can be viewed as the wavelength consumption for rerouting request r. The wavelength consumption can be thus represented by $\sum_{1 \leqslant k \leqslant l} n\_emp(T_k)$, where $n\_emp(T_k) = 1$ indicates T_k is not empty; otherwise,

$n\_emp(T_k) = 0$. The goal of the *MRWA-DC* problem is to find a feasible forest Γ that minimizes the multicast cost f defined by

$$f(\Gamma) = \alpha \sum_{1 \leqslant k \leqslant l} c(T_k) + \beta \sum_{1 \leqslant k \leqslant l} n\_emp(T_k),$$

where α and β reflect the relative importance between communication cost and wavelength consumption. In the following section, an ILP formulation will be proposed to solve the problem.

3. ILP formulation

If no confusion would arise, notation v, e, λ and d are used to represent v_i, e_j, λ_k and d_l. Nevertheless, $d(e)$ represents the communication cost of link e. The notation used in our ILP formulation is given as follows:

$I(v)$: set of inbound edges to node v in V,
$O(v)$: set of outbound edges from node v in V,
$y_e^{\lambda,d}$: binary variable indicating whether wavelength λ over link e is used for the light-path from s to d, i.e., $y_e^{\lambda,d} = 1$, if yes; 0, otherwise;
x_e^{λ} : binary variable indicating whether wavelength λ over link e is used for the request, i.e., $x_e^{\lambda} = 1$, if yes; 0, otherwise;
z^{λ} : binary variable indicating whether wavelength λ is used for the request, i.e., $z\lambda = 1$, if yes; 0, otherwise.

The *MRWA-DC* problem can be formulated as follows:
Objective function

$$\text{Minimize} \quad \alpha \sum_{\lambda \in M} \sum_{e \in E} x_e^{\lambda} c(e) + \beta \sum_{\lambda \in M} z^{\lambda}$$

Subject to

(c1) $\quad \forall d \in D, \quad \sum_{\lambda \in M} \sum_{e \in I(s)} y_e^{\lambda,d} - \sum_{\lambda \in M} \sum_{e \in O(s)} y_e^{\lambda,d} = -1 \quad$ (source constraints),

(c2) $\quad \forall d \in D, \quad \sum_{\lambda \in M} \sum_{e \in I(d)} y_e^{\lambda,d} - \sum_{\lambda \in M} \sum_{e \in O(d)} y_e^{\lambda,d} = 1 \quad$ (target constraints),

(c3) $\quad \forall d \in D, \forall v \in V, v \neq s, d, \forall \lambda \in M, \quad \sum_{e \in I(v)} y_e^{\lambda,d} - \sum_{e \in O(v)} y_e^{\lambda,d} = 0 \quad$ (wavelength continuity constraints),

(c4) $\quad \forall v \in V, \forall \lambda \in M, \quad \sum_{e \in I(v)} x_e^{\lambda} \leqslant 1 \quad$ (input constraints),

(c5) $\quad \forall v \in V, \forall \lambda \in M, \quad \sum_{e \in O(v)} x_e^{\lambda} \leqslant \theta(v) \quad$ (capacity constraints),

(c6) $\forall d \in D, \forall e \in E, \forall \lambda \in \mathrm{M}, \quad x_e^{\lambda} \geqslant y_e^{\lambda,d}$ (link usage constraints),

(c7) $\forall e \in E, \forall \lambda \in M \quad x_e^{\lambda} \leqslant z^{\lambda} \cdot w(e, \lambda)$ (wavelength usage constraints),

(c8) $\forall d \in D, \forall \lambda \in M, \quad \sum_{e \in E} y_e^{\lambda,d} d(e) \leqslant \Delta$ (delay constraints),

(c9) $\forall d \in D, \forall e \in E, \forall \lambda \in M, \quad y_e^{\lambda,d} \in \{0,1\}$ (0 − 1 constraints),

(c10) $\forall e \in E, \forall \lambda \in M, \quad x_e^{\lambda} \in \{0,1\}$ (0 − 1 constraints),

(c11) $\forall \lambda \in M, \quad z^{\lambda} \{0,1\}$ (0 − 1 constraints).

In the above formulation, source constraints (c1), target constraints (c2), and wavelength continuity constraints (c3) ensure that a light-path using some specified wavelength from the source to each d can be found, and that each node except the source and destinations should pass a signal (information) from its input port to its output port. Since at most one signal can enter the input port using the same wavelength, we need input constraints (c4). Moreover, the number of split output signals must be bounded by the light-splitting capacity. The capacity constraints (c5) ensure that $\sum_{e \in O(v)} x_e^{\lambda}$, the number of outbound edges from v for routing the signals of the output port to other nodes using λ, should be smaller than or equal to $\theta(v)$. Therefore, the constraint $x_e^{\lambda} \geqslant y_e^{\lambda,d}$ must be satisfied by the set of constraints (c6). Constraints (c7) ensure that only the link with unused wavelengths ($w(e, \lambda) = 1$) can be used to route the request. Therefore, $z\lambda = 1$ and $w(e, \lambda) = 1$, if $x_e^{\lambda} = 1$. According to the above discussion, $\cup_{y_e^{\lambda,d}=1} e$ (i.e., $\cup_{\forall e \in E, y_e^{\lambda,d}=1} e$) will be proved to be a light-path from s to d using wavelength λ or a null set, and its transmission delay can be represented by $\sum_{e \in E} y_e^{\lambda,d} d(e)$. The set of constraints (c8) is used to specify the delay bound constraint. Constraints (c9) and (c11) are used to define binary variables. In the first term in the objective function, $x_e^{\lambda} = 1$ means that wavelength λ over link e needs to route the request, so $\sum_{e \in E} x_e^{\lambda} c(e)$ is the communication cost of using wavelength λ. Therefore, the first term represents the total communication cost. The second term in the objective function denotes the total wavelength consumption.

In the ILP formulation, the numbers of variables, $y_e^{\lambda,d}$, x_e^{λ}, and z^{λ}, are $mq\gamma$, $m\gamma$, and γ, respectively. Therefore, the total number of variables is $(q+1)m\gamma+\gamma$. The numbers of constraints (1)–(8) are q, q, $(n-2)q\gamma$, $n\gamma$, $n\gamma$, $mq\gamma$, $m\gamma$, and $q\gamma$, respectively. The total number of constraints in the ILP is $(nq + mq + 2n + m - q)\gamma + 2q$.

Using the ILP formulation, the solution would be proved in following properties to be an optimal light-forest to solve the *MRWA-DC* problem.

Property 1. For each $d \in D$, there exists one edge $e \in O(s)$ and one wavelength $\lambda \in M$, such that $y_e^{\lambda,d} = 1$ in each solution.

Proof. According to constraints (c1), we have $\sum_{\lambda \in \mathrm{M}} \sum_{e \in O(s)} y_e^{\lambda,d} \geqslant 1$ for each $d \in D$ in the solution satisfying all constraints in the formulation. Therefore, there exists at latest one specific link e in $O(s)$, and one wavelength λ in M such that $y_e^{\lambda,d} = 1$ for each d.   □

Property 2. There exists only one light-path from s to d using some wavelength λ.

Proof. According to Property 1, $\forall~d \in D$, $\exists~e_1 \in O(s)$ and $\lambda \in M$ such that $y_{e_1}^{\lambda,d} = 1$. We assume that there is no light-path from s to d; that is, the terminal node in the light-path is not equivalent to d. Without loss of generality, suppose the light-path is $\hat{p}^{\lambda} = \langle e_1, e_2, \ldots, e_{a-1} \rangle$, where e_j is the edge from v_j to v_{j+1} and $v_a \neq d$. For all e_j, $1 \leqslant j \leqslant a-1$, we have $y_{e_l}^{\lambda,d} = 1$. Nevertheless, for $v_a \neq d$ such that $\hat{p}^{\lambda}$ ends at v_a, we have $\sum_{e \in O(v_a)} y_e^{\lambda,d} = 0$. It implies $\sum_{e \in I(v_a)} y_e^{\lambda,d} - \sum_{e \in O(v_a)} y_e^{\lambda,d} \geqslant 1$ and (c3) is violated. A contradiction arises. Therefore, we may conclude that $v_a = d$ and $\hat{p}^{\lambda}$ is a light-path from s to d using wavelength λ. If more than one light-path exists for routing the data to d, then the value of the objective function is not minimum. The proof is complete. $\square$

Thus, $\cup_{y_e^{\lambda,d}=1} e$ (represented as $\hat{p}^{\lambda}(s,d)$) may be a light-path or an empty set when λ is used or not used from s to d. According to Property 2, there are exactly q non-empty light-paths among $\hat{p}^{\lambda}(s,d)$ for all $\lambda \in M$ and for all $d \in D$. Because the objective is to minimize the multicast cost, it can be seen that a loop never exists in $\hat{p}^{\lambda}(s,d)$. To simplify our discussion, directions of a light-path and a light-tree are ignored in the rest of the paper.

Property 3. A graph obtained by merging all light-paths using the same wavelength is a light-tree.

Proof. Assume that $\cup_{d \in D} \hat{p}^{\lambda}(s,d)$, the union of all light-paths using λ, is not a tree; that is, there exists at least one cycle in $\cup_{d \in D} \hat{p}^{\lambda}(s,d)$. Suppose the cycle is formed by two different sub-light-paths between two specific nodes u and v in two light-paths. Therefore, there are two input signals using λ entering v which will cause input constraints (c4) to be violated. Moreover, the set of capacity constraints (c5) ensures that the number of split signals of each internal node in $\cup_{d \in D} \hat{p}^{\lambda}(s,d)$ is not greater than its light splitting capacity. Therefore, $\cup_{d \in D} \hat{p}^{\lambda}(s,d)$ is a light-tree. $\square$

For each $e \in E$, because $x_e^{\lambda} = 1$ will let $z^{\lambda} = 1$ satisfied and may increase the objective function value, x_e^{λ} is set to 1 for satisfying (c6) and z^{λ} is set to 1 for satisfying (c7) only; otherwise, the link usage constraints (c6) and wavelength usage constraints (c7) is violated or the value of objective function is not minimum. Suppose that $T^{\lambda} = \cup_{d \in D} \hat{p}^{\lambda}(s,d)$. Since $\hat{p}^{\lambda}(s,d) = \cup_{y_e^{\lambda,d}=1} e$, $T^{\lambda} = \cup_{d \in D} \cup_{y_e^{\lambda,d}=1} e$ represents a light-tree using λ to route the request to each destination $d \in V(T^{\lambda}) \cap D$.

Property 4. A feasible light-forest will be found in each solution of the ILP formulation.

Proof. According to Properties 2 and 3, each destination can imply a light-path using some wavelength and these light-paths using λ can form a light-tree $T^{\lambda} = \cup_{x_e^{\lambda}=1} e$. For the delay constraints (c8) ensuring $\forall l$, $1 \leqslant l \leqslant q$, $d_l \in D$, $\sum_{e \in E} y_e^{\lambda,d_l} d(e) \leqslant \Delta$ and $\hat{p}^{\lambda}(s,d_l) = \cup_{y_e^{\lambda,d_l}=1} e$ being a light-path in T^{λ}, $d(\hat{p}^{\lambda}(s,d_l)) \leqslant \Delta$ is obtained. $d(T^{\lambda}) = \max_{d_l \in D \cap V(T^{\lambda})} d(\hat{P}^{\lambda}(s,d_l)) \leqslant \Delta$ is satisfied. Therefore, it can be seen that $\Gamma = \{(T^{\lambda}, \lambda) \mid T^{\lambda} = \cup_{x_e^{\lambda}=1} e, \forall \lambda \in M\}$ is a feasible light-forest for rerouting the request. $\square$

Property 5. The communication cost of the light-forest is $\sum_{\lambda \in M}, \sum_{e \in E} x_e^\lambda c(e)$.

Proof. As defined above, $c(T^\lambda) = \sum_{e \in T^\lambda} c(e)$ and $T^\lambda = \cup_{x_e^\lambda = 1} e$. So, we have $c(T^\lambda) = \sum_{x_e^\lambda = 1} c(e) = \sum_{e \in E} x_e^\lambda c(e)$. According to Property 4, we have that the communication cost of Γ is $\sum_{\lambda \in M} c(T^\lambda) = \sum_{\lambda \in M} \sum_{e \in E} x_e^\lambda c(e)$.　□

Property 6. The wavelength consumption is $\sum_{\lambda \in M} z^\lambda$.

Proof. By Property 3, if $x_e^\lambda = 1$, link e will be contained in some light-tree using $\lambda (T^\lambda = \cup_{x_e^\lambda = 1} e)$. That is, the wavelength λ needs to be used to route the request (which implies $z\lambda = 1$) for $x_e^\lambda = 1$ in (c7). That is, $x_e^\lambda = 1$ implies that T^λ is not empty (i.e., $n\_emp(T^\lambda) = 1$) and $z\lambda = 1$. We know that z^λ is set to 1 for satisfying (c7) only; otherwise, (c7) is violated or the objective function value cannot be minimum. Therefore, the wavelength consumption $\sum_{\lambda \in M} n\_emp(T^\lambda)$ is $\sum_{\lambda \in M} z^\lambda$.　□

Property 7. The objective function defined in the ILP formulation is equivalent to the multicast cost function.

Proof. As defined above and Properties 5 and 6, $f(\Gamma) = \alpha \sum_{T^\lambda \in \Gamma} c(T^\lambda) + \beta \sum_{T^\lambda \in \Gamma} n\_emp(T^\lambda) = \alpha \sum_{\lambda \in M} \sum_{e \in E} x_e^\lambda c(e) + \beta \sum_{\lambda \in M} z^\lambda$. The property readily follows.　□

According to Properties 4 and 7, each solution of the ILP formulation must be an optimal light-forest with the minimum multicast cost.

4. Near-k-shortest-path heuristic (NKSPH)

In the previous section, an ILP was presented to solve the *MRWA-DC* problem, but it is not affordable for the networks consisting of more nodes or wavelengths. Therefore, it is required to propose a heuristic algorithm that can route requests in a reasonable time. Let a *wavelength-based graph* $G^\lambda(V, E^\lambda)$ of λ be defined as a graph obtained by removing all edges which are not available in λ, where $E^\lambda = \{e | e \in E, w(e, \lambda) = 1\}$. For each node pair of the source and destinations, a light-path whose transmission delay abides by the delay bound can be found in $G^\lambda(V, E^\lambda)$. We consider the graph obtained by the union of the light-paths for all source–destination pairs. In this graph, there could be infeasible nodes or cycles. Here, the infeasibility of a node means that the number of outbound branches from the node outnumbers its light splitting capacity. We remove some outbound edges from each infeasible node to make it feasible. Further, we delete the edge with the maximal transmission delay in each cycle so as to derive a tree. The resulting tree will be referred as a light-tree and denoted by T. The request can be routed using wavelength λ on the available paths of T to some destinations. For these destinations not reachable on T, another light-tree will be found by applying the above procedure to another wavelength-based graph. The process is repeated until all destinations can be successfully routed. The

set of light-trees developed is called a light-forest, which is a solution to the studied problem. In the above discussion, three sub-problems still remain to be resolved: (1) in what order are the wavelengths selected; (2) how to resolve the infeasibility of nodes; and (3) how to minimize the communication cost of each light-tree.

We know that an optimal solution is not guaranteed in the above iterative process. Therefore, the *NKSPH* proposed in this section does not focus on how to find a near optimal light-tree in iteration but on providing a large-scale adjustable search space. For the three issues that our heuristic needs to address, we use simple strategies to prevent from heavy computing loads. The strategy includes selecting wavelengths in random, preserving the edges connecting to the maximal number of destinations, and selecting the k near shortest light-paths that satisfy the delay bound. Recall that there are q destinations in D and k light-paths are demanded for each destinations. Each iteration of the above iterative process consists in the following seven steps:

(1) Choose a wavelength λ and construct the wavelength-based graph $G^{\lambda}(V, E^{\lambda})$.
(2) In $G^{\lambda}(V, E^{\lambda})$, for each destination find at most k light-paths from the source subject to the delay bound.
(3) For each destination, select one of the k derived light-paths and the source. Unify the q light-paths for all destinations to form a graph. Corresponding to k^{q} possible combinations of light-paths for q destinations, at most k^{q} graphs will be constructed.
(4) For any node violating its capacity constraint (i.e., infeasible node) in each of the q^{k} graphs, reserve the outbound edges whose number is equal to the light splitting capacity of the node and which covers more number of destinations than the eliminated to form a light-tree.
(5) For any node violating the input constraint in each graph, reserve the inbound edge with the minimal transmission delay to eliminate the cycle.
(6) Choose the light-tree with a minimum communication cost out of the q^{k} light-trees to be a candidate T^{λ} and remove the destinations in T^{λ} from D.
(7) Repeat the previous six steps until D is empty.

The details of the *NKSPH* procedure are given below.

NKSPH(*G*, *r*(*s*, *D*, *Δ*), *k*)
{
   1. $\Gamma = $ NULL, $T^{opt} = \varnothing$, $C^{opt} = \infty$, $Dest^{opt} = 0$
   2. **While** $D \neq \varnothing$
   3.   **If** $M = \varnothing$ //No wavelength available for routing the data
   4.   Return NULL
   5.   Randomly select a wavelength λ from M
   6. $G^{\lambda}(V, E^{\lambda})$, where $E^{\lambda} = \{e|\ e \in E, w(e, \lambda) = 1\}$
   7.   **For** each $d_l \in D$,
   8. $P(s, d_l, k_l) = $ Finding-k-Near-Shortest-Path($G^{\lambda}(V, E^{\lambda})$, s, d_l, k, Δ)
         //k_l, $k_l \leqslant k$, is the number of light-paths found between s and d_l
   9.   **End** For-loop
  10.   **For** each T, $T = \cup_{1 \leqslant l \leqslant q} Path_{d_l} e$, $Path_{d_l} \in P(s, d_l, k_l)$ for all l, $1 \leqslant l \leqslant q$
  11.   **For** all node x in T in the breadth first order

12.   **If** $(out(T, x) > \theta(x))$ //x is infeasible with respect to the capacity constraint
13.   delete the first $out(T, x) - \theta(x)$ edges according to non-decreasing order of
                        the numbers of connected destinations.
14   **If** $(in(T, x) > 1)$ //x violates the input constraint
15.   delete the first $in(T, x) - 1$ edges according to non-increasing order of
                    the transmission delay
16.   **End** For-loop
17.   **If** $(D(T) > Dest^{opt}$ and $c(T) < MD(T))$ or //$MD(T) = \sum_{d_l \in V(T) \cap D} \min_{P \in P(s,d_l,K_l)} c(P)$

        $(D(T) = Dest^{opt}$ and $c(T) < C^{opt})$
18. $C^{opt} = c(T)$
19. $Dest^{opt} = D(T)$
20. $T^{opt} = T$
21.   **End** For-loop
22. $\Gamma = \Gamma \cup \{(T^{opt}, \lambda)\}$, $D = D - V(T^{opt})$
23. **End** While-loop
24. **Return** Γ
}

In Step 8, *Finding-k-Near-Shortest-Path*$(G^{\lambda}(V, E^{\lambda})$, s, d_l, k, $\Delta)$ is a procedure call for finding k near shortest light-paths with minimum transmission delays between s and d_l subject to the delay bound Δ. Applying these k-shortest path algorithms developed in Pascoal et al. (2001) and Eppstein (1998) can find k shortest light-paths with a minimum communication cost or a minimum transmission delay. However, these algorithms cannot be used directly to find constrained shortest light-paths. Furthermore, it is difficult to revise these algorithms for deriving constrained shortest light-paths. The facts that these algorithms are complicated to implement or modify to meet the needs and that optimal light-paths are not necessarily optimal suggest the deployment of a simple approach for finding the shortest paths in a timely fashion. The *Finding-k-Near-Shortest-Path* strategy is implemented as an iterative procedure consisting of three steps: (1) apply *Dijkstra's* shortest path algorithm (Dijkstra, 1959) to find a light-path with minimum transmission delay (i.e., the light-path is a constrained shortest light-path); (2) keep the light-path if its transmission delay satisfies the delay constraint; and (3) delete the edge whose transmission delay is minimum in this light-path from $G^{\lambda}(V, E^{\lambda})$. The procedure terminates when k near shortest light-paths have been obtained or no more light-path can be found.

Assume the number of light-paths found for destination d_l is k_l, $k_l \leqslant k$. Let $P(s, d_l, k_l)$ be the set of constructed light-paths. A graph is formed by selecting one light-path $Path_{d_l}$ from $P(s, d_l, k_l)$ for each $d_l \in D$ and then unifying these selected light-paths. Denote this graph by $T = \cup_{1 \leqslant l \leqslant q} Path_{d_l}$. The derived graph may contain cycles or infeasible nodes. In Steps 12–13 and 14–15, all nodes in T are examined to verify the capacity constraint and the input constraint. For each node x violating the capacity constraint, the first $out(T, x) - \theta(x)$ edges in non-decreasing order of the numbers of connections to destinations will be eliminated in Step 13 such that each internal node is feasible. In Step 15, the edge with the minimum transmission delay will be reserved when there is more than one edge connected to the node. All cycles can be detected and removed in Step 14–15. After these steps, a light-tree will be obtained. In Step 17, $D(T)$ represents the number of destinations

contained in T and $MD(T) = \sum_{d_l \in V(T) \cap D} \min_{P \in P(s,d_l,K_l)} c(P)$ represents the sum of minimum communication costs of the destinations in T. The case $c(T) < MD(T)$ implies that T will include at least one MC node to reduce the communication cost. In Steps 17–20, a light-tree T^{opt} with the minimum communication cost or covering more destinations will be kept as a local optimal light-tree by using λ. This approximate solution is denoted by (T, λ). Steps 2–23 will be executed iteratively until no wavelength is available or the destination set D is empty. The former case leads to the report that the request cannot be successfully routed by the algorithm.

5. Experiments

In this paper, our work focuses on how to find an optimum light-forest such that switches in the network can be set up to route a request. The approach used in the experiments for evaluating the performance of our solution model follows that used in Waxman (1988). In this approach, there are n nodes randomly distributed over a rectangular grid. The coordinates of all nodes are integer. For the network topology generated in the experiments, each directed link from u to v is associated with the probability function $P(u, v) = \lambda \exp(-p(u, v)/\gamma\delta)$, where $p(u, v)$ is the distance between u and v and δ is the maximum distance between any two nodes. Note that $0 < \lambda, \gamma \leqslant 1$. In the probability function, a larger value of λ produces a network with higher link densities, and small values of γ increase the densities of short links relative to longer ones. We set $\lambda = 0.7$ and $\gamma = 0.9$, let 15% of nodes be MC nodes with randomly generated light splitting capacities, and set the size of rectangular grid to be 50.

Eight types of networks were tested: 30 switches ($n = 30$), 40 switches ($n = 40$), 50 switches ($n = 50$), 60 switches ($n = 60$), 70 switches ($n = 70$), 80 switches ($n = 80$), 90 switches ($n = 90$), and 100 switches ($n = 100$), for each of which 60 different requests were randomly generated. Each 60 requests are categorized into 3 groups corresponding to 2 destinations ($q = 2$), 3 destinations ($q = 3$), and 4 destinations ($q = 4$). The communication cost and the transmission delay of each link are defined as the distance of two nodes of the link on the grid and a random number between 0.1 and 3, respectively. For each request, the source and the destinations were generated randomly. Nevertheless, the value of delay bound Δ needs to be reasonable for otherwise it is very likely that no feasible light-forest can be found. Delay bound Δ is set to be equal to χ times of the derived minimum transmission delay between the source and all destinations in each request, where χ is a control parameter dictating the tightness between delay bound and minimum transmission delay. For example, $\chi = 3$ means that all delay bounds of requests were set to be $3 \times$ their minimum transmission delay.

The experiments consist of four parts: comparison of different wavelength consumption ratios, performance assessment of the ILP, comparisons between different values of k in *NKSPH*, and comparison between *NKSPH* and ILP. Program codes were implemented in C++ with *ILOG's CPLEX* 7.0 on a personal computer with an Intel P4 2.4GMhz CPU and 2GB RAM.

(1) *Comparison of different wavelength consumption ratios:* The multicast cost depends on different values of α and β. We shall study the effects of different values of β/α. The values of β/α are set to be 0.1 ($\alpha = 1$ and $\beta = 0.1$), 1 ($\alpha = 1$ and $\beta = 1$), 10 ($\alpha = 1$ and $\beta = 10$), 50 ($\alpha = 1$ and $\beta = 50$), and 100 ($\alpha = 1$ and $\beta = 100$). For each combination of β/α and χ, we route 5 requests in the network with 50 nodes ($n = 50$). We keep track of the average

communication cost (*CC*), average wavelength consumption (*WC*), and average elapsed execution time (*ET*) (in seconds) for each 5 requests. The results are summarized in Table 1. Several observations can be made as follows:

(i) For larger values of χ, the impact of different values of $\beta \diagup \alpha$ on execution times is not significant. However, the elapsed execution times increase dramatically as the value of β/α grows and the delay bound becomes tighter (i.e., smaller values of χ). For example, for $\beta/\alpha = 0.1, 1, 10, 50$, and 100, *ET* is 53.87, 47.69, 67.79, 46.53, and 67.04 seconds when $\chi = 3.0$ and *ET* is 886.31, 1652.01, 2953.35, 2644.63, and 3533.81 seconds when $\chi = 1.2$. From these results, we know that elapsed execution time is unacceptably lengthy as the delay bound becomes tighter. For example, when $\chi = 1.2$, the average execution time is more than 3,533.81 s in $\beta/\alpha = 100$. Therefore, the ILP formulation cannot solve the *MRWA-DC* problem well when the specified delay bound of a request is close to the minimum transmission delay.

(ii) The ratio of β to α can be set properly such that a request can be routed by less communication cost or wavelength consumption. For example, the average wavelength consumption slightly decreases from 1.6 to 1.0 as the ratio increases rapidly from 0.1 to 100. For $\beta/\alpha = 0.1$ or 1 in $\chi = 3.0$, the communication cost $CC = 79.4$ is only determined by the wavelength consumption cost. For $\beta/\alpha = 100$ and $\chi = 3.0$, the wavelength consumption cost, the product of wavelength consumption and β ($\beta \cdot \omega = 100 \times 1.0$), is dominated by the communication cost. It is therefore reasonable to adjust the ratio to balance the load of each wavelength and each link. Moreover, dynamic adjustment may be a reasonable policy.

(2) *The performance assessment of ILP:* The test instances in Part (1) are again used here. We set $\alpha = 1$, $\beta = 1$, and $\chi = 1.2$. Table 2 shows the results from routing 60 requests in five networks ($n = 30, 40, 50, 60, 70$). For each combination of network types (n/m: the number of nodes/the number of edges) and the number of destinations (q), the first block shows the number of ILP variables (#*NV*), the number of constraints (#*NR*), the number of non-zero constraint entries (#*NZ*) in these constraints. The first performance index we are interested in is the number of requests that are successfully solved (#*Succ*). Performance indices resulted from the successfully solved instances include the minimum elapsed execution time ($ET_{\min}$), the maximum elapsed execution time ($ET_{\max}$), the average minimal elapsed execution time of the first 3 requests ($ET_{\min 3}$), the average maximal elapsed execution time

Table 1
Experimental results between ILP and *NKSPH*

χ	$\beta/\alpha = 0.1$			$\beta/\alpha = 1$			$\beta/\alpha = 10$			$\beta/\alpha = 50$			$\beta/\alpha = 100$		
	CC	*w*	*ET*	*CC*	*w*	*ET*	*CC*	*w*	*ET*	*CC*	*w*	*ET*	*CC*	*w*	*ET*
3.0	79.4	1.6	53.87	79.4	1.6	47.69	81.2	1.0	67.79	81.2	1.0	46.53	81.2	1.0	67.04
2.0	79.6	1.6	81.99	79.6	1.6	74.50	82.2	1.0	127.27	82.2	1.0	127.46	82.2	1.0	127.18
1.5	84.2	1.8	340.33	84.2	1.8	409.39	84.6	1.2	369.14	86.8	1.0	269.49	86.8	1.0	405.02
1.4	85.0	1.8	539.96	85.0	1.8	505.13	85.4	1.2	586.04	87.6	1.0	767.96	87.6	1.0	1149.81
1.3	88.0	2.0	510.91	88.0	2.0	737.73	88.4	1.4	664.81	88.6	1.0	847.96	88.6	1.0	1088.45
1.2	89.4	2.0	886.31	89.4	2.0	1652.01	89.8	1.4	2953.35	92.0	1.0	2644.63	92.0	1.0	3533.81

Table 2
Experimental results for different networks by using ILP

N/m	q	#NV	#NR	#NZ	#Succ	ET_{min}	ET_{max}	ET_{min3}	ET_{max3}	ET_{other}	ET
30/236	2	3540	2954	14,160	20	0.6	13,256.31	0.79	4842.18	14.50	816.83
	3	4720	4281	20,060	19	0.2	2355.52	0.85	1560.76	24.28	263.18
	4	5900	5608	25,960	19	3.6	82,981.89	9.36	42,145.73	640.36	7094.21
40/456	2	6840	5354	27,360	20	0.1	43,923.91	0.49	17,562.67	69.10	2682.85
	3	9120	7831	38,760	16	8.5	27,638.92	18.39	16,583.96	469.58	3406.43
	4	11,400	10,308	50,160	14	17.1	66,999.05	60.04	29,153.12	1226.44	6960.79
50/588	2	8820	6874	35,280	20	0.8	41,063.38	1.99	14,460.90	125.21	2257.08
	3	11,760	10,061	49,980	19	2.3	11,014.69	18.68	10,014.08	819.63	2144.92
	4	14,700	13,248	64,680	18	25.6	74,096.81	65.25	46,930.01	2590.33	7096.11
60/964	2	14,460	10,834	57,840	20	0.2	1253.83	4.57	1120.02	147.71	272.09
	3	19,280	15,951	81,940	18	10.7	123,882.13	53.64	89,361.25	4710.74	18,042.98
	4	24,100	21,068	106,040	6	202.2	73,268.66	73,268.66	—	—	22,266.95
70/1124	2	16,860	12,634	67,440	18	2.1	84,512.84	10.81	84,512.84	1018.16	5488.86
	3	22,480	18,601	95,540	8	9.8	73,856.50	343.21	—	30,475.86	17,561.86
	4	—	—	—	—	—	—	—	—	—	—

of the first 3 requests (ET_{max3}), the average elapsed execution time of other requests (ET_{other}), and the average overall elapsed execution time (ET). From the numerical results, we have the following observations:

(i) The elapsed execution time is more than 34 h (for $ET_{max} = 123,882.13$ s) for some request ($q = 3$) in the network with 60 nodes ($n = 60$), and the ILP formulation cannot be used to solve the network with more than 70 nodes or requests with more than 3 destinations in a reasonable time. For example, $ET_{min3} = 73,268.66$ s in $n = 60$ and $q = 4$. In general, the execution time is proportional to the number of destinations and the number of nodes.

(ii) For networks with more nodes, to fulfill the request becomes hard. That is, the number of solved requests decreases sharply when the number of nodes in network increases. For example, #Succ = 18 for $n = 60$ and $q = 3$, but #Succ = 8 for $n = 70$ and $q = 3$; furthermore, #Succ = 6 for $n = 60$ and $q = 4$, but almost no request for $n = 70$ and $q = 4$ can be solved. This observation helps explain the phenomenon that $ET_{max} = 123,882.13$ s for $n = 60$ and $q = 3$ is much greater than $ET_{max} = 73,856.50$ s for $n = 70$ and $q = 3$. Amongst the 18 solved instances for $n = 60$, there could be some instances that have used an extraordinarily long execution time, while the 6 solved instances for $n = 70$ might be easier to solve.

(3) *Comparisons for different values of k in NKSPH:* Next we proceed to discuss the efficiency of *NKSPH* for different values of k. Numerical results are summarized in Table 3. For the column corresponding to each request group ($q = 2$, $q = 3$, and $q = 4$), we keep track of feasible solutions found (#Succ), average multicast cost (MC), and average elapsed execution time (ET) over every 20 requests. From the numerical results, we have the following observations:

Table 3
Experimental results for different values of k in *NKSPH*

n	k	$q = 2$			$q = 3$			$q = 4$		
		#Succ	MC	ET	#Succ	MC	ET	#Succ	MC	ET
40	2	20	52.75	0.01	20	93.80	0.02	20	103.15	0.02
	4	20	50.10	0.02	20	89.60	0.03	20	95.75	0.04
	6	20	50.10	0.02	20	88.35	0.04	20	93.20	0.06
	8	20	50.10	0.02	20	88.35	0.05	20	92.35	0.09
	10	20	50.10	0.03	20	90.20	0.06	20	91.95	0.10
	15	20	50.10	0.03	20	90.20	0.10	20	93.15	0.25
	20	20	50.10	0.03	20	90.20	0.08	20	93.15	0.28
60	2	20	56.85	0.03	20	80.30	0.06	20	105.40	0.13
	4	20	53.00	0.05	20	74.00	0.17	20	102.35	0.21
	6	20	52.45	0.07	20	72.75	0.22	20	96.75	0.42
	8	20	52.25	0.15	20	73.35	0.30	20	95.65	0.56
	10	20	52.25	0.11	20	73.35	0.18	20	92.50	0.38
	15	20	52.20	0.11	20	72.65	0.24	20	91.55	0.89
	20	20	52.20	0.11	20	72.65	0.48	20	91.40	1.19
80	2	20	57.05	0.06	20	89.40	0.09	20	116.80	0.20
	4	20	54.40	0.19	20	86.95	0.23	20	112.75	0.32
	6	20	54.40	0.15	20	85.25	0.27	20	111.10	0.68
	8	20	53.95	0.31	20	83.55	0.61	20	108.15	0.83
	10	20	53.95	0.30	20	83.05	0.59	20	107.40	1.07
	15	20	53.95	0.20	20	82.65	0.64	20	109.15	0.92
	20	20	53.95	0.21	20	82.65	0.47	20	108.45	1.15
100	2	20	60.05	0.11	20	94.65	0.19	20	125.20	0.32
	4	20	53.00	0.25	20	91.90	0.44	20	119.65	0.66
	6	20	53.00	0.32	20	89.95	0.56	20	117.90	1.24
	8	20	53.00	0.58	20	88.90	0.71	20	117.35	1.42
	10	20	53.00	0.45	20	88.90	1.03	20	118.30	1.84
	15	20	53.00	0.86	20	88.40	1.13	20	117.15	2.13
	20	20	53.00	0.54	20	88.40	1.14	20	114.65	3.24

(i) The elapsed execution time of the *NKSPH* is related to the value of k, the number of destinations, and the number of nodes in the network. For example, $ET = 0.01$ s for $n = 30$, $q = 2$, and $k = 2$ and $ET = 3.24$ s for larger values of $n = 100$, $q = 4$, and $k = 20$.

(ii) Increasing the value of k usually reduces the multicast cost. Such a phenomenon is significant when the request is associated with more destinations or the network has more nodes. For example, it can reduce multicast cost 4.18% $(119.65 - 114.64/119.65 \times 100\%)$ for $n = 100$, $q = 4$ from $k = 4$ to 20, but there is no improvement for $n = 100$, $q = 2$ from $k = 4$ to 20. Nevertheless, in some cases the increase multicast cost is higher when the request is associated with more destinations and the network with more nodes. For example, $MC = 117.35$ for $n = 100$, $m = 4$, and $k = 8$, but $MC = 118.30$ when $k = 10$. It seems reasonable to conclude that choosing a suitable value of k can not only reduce the multicast cost but also save the execution time.

(4) *Comparisons between NKSPH and ILP:* In order to compare *NKSPH* and ILP, only these requests solved successfully in ILP are addressed in the experiment. For each combination of network types and different method (ILP vs. *NKSPH* with different values of k), all experimental results are summarized into three blocks for three groups of requests ($q = 2$, $q = 3$, and $q = 4$). Each block displays the number of optimal solutions (#*Opt*) produced by our heuristic, average multicast cost deviation (*Dev*) of the found solutions from the optimal ones found by the ILP, *MC*, and *ET*, where *Dev* is defined as (Table 4)

$$Dev = \frac{\sum_r f(\Gamma_r^k) - f(\Gamma_r^{opt})/f(\Gamma_r^{opt})}{\#Opt_{ILP}} \times 100\%,$$

where $f(\Gamma_r^k)$ and $f(\Gamma_r^{opt})$ are the multicast cost of the feasible solution Γ_r^k found for k by *NKSPH* and the multicast cost of the optimal solution Γ_r^{opt} found by the ILP for request r,

Table 4
Experimental results between ILP and *NKSPH*

n	Method	$q = 2$				$q = 3$				$q = 4$			
		#*Opt.*	*Dev*	*MC*	*ET*	#*Opt.*	*Dev*	*MC*	*ET*	#*Opt.*	*Dev*	*MC*	*ET*
30	ILP	20	—	63.72	816.83	19	—	76.68	263.18	19	—	84.53	7094.21
	NKSPH												
	$k = 8$	9	5.38%	68.50	0.02	7	9.47%	83.63	0.04	8	12.44%	94.42	0.10
	$k = 10$	9	5.38%	68.50	0.02	7	9.47%	83.63	0.05	9	12.21%	94.21	0.12
	$k = 15$	9	5.38%	68.50	0.02	7	9.47%	83.63	0.05	9	8.64%	91.42	0.17
	$k = 20$	9	5.38%	68.50	0.02	7	9.47%	83.63	0.05	9	8.64%	91.42	0.18
40	ILP	20	—	49.25	2682.85	16	—	70.44	3406.43	14	—	77.70	6960.79
	NKSPH												
	$k = 8$	17	1.22%	50.10	0.02	10	8.79%	76.06	0.05	5	6.66%	82.60	0.09
	$k = 10$	17	1.22%	50.10	0.03	10	11.76%	78.38	0.06	6	5.67%	81.80	0.10
	$k = 15$	17	1.22%	50.10	0.03	10	11.76%	78.38	0.10	7	3.79%	80.10	0.25
	$k = 20$	17	1.22%	50.10	0.03	10	11.76%	78.38	0.08	7	3.79%	80.10	0.28
50	ILP	20	—	57.20	2257.08	19	—	68.68	2144.92	18	—	84.44	7096.11
	NKSPH												
	$k = 8$	13	8.60%	62.95	0.04	8	12.42%	77.26	0.12	5	17.30%	99.89	0.22
	$k = 10$	13	8.60%	62.95	0.09	8	12.42%	77.26	0.19	5	16.13%	98.56	0.33
	$k = 15$	13	8.60%	62.95	0.05	8	11.62%	76.79	0.14	5	15.38%	97.78	0.44
	$k = 20$	13	8.60%	62.95	0.06	8	11.62%	76.79	0.18	5	15.38%	97.78	0.43
60	ILP	20	—	50.20	272.01	18	—	63.33	18,042.98	6	—	63.33	22,266.95
	NKSPH												
	$k = 8$	12	3.59%	52.25	0.15	6	8.27%	71.67	0.34	4	9.93%	72.33	1.87
	$k = 10$	12	3.59%	52.25	0.11	6	8.27%	71.67	0.20	4	7.76%	70.67	1.26
	$k = 15$	13	3.50%	52.20	0.11	6	7.20%	70.89	0.26	5	4.08%	67.83	2.98
	$k = 20$	13	3.50%	52.20	0.11	6	7.20%	70.89	0.53	5	4.08%	67.83	3.96
70	ILP	18	—	53.33	5488.86	8	—	69.10	17,561.86	—	—	—	—
	NKSPH												
	$k = 8$	11	2.00%	54.56	0.28	4	8.44%	74.50	0.91	—	—	118.65	2.02
	$k = 10$	11	2.00%	54.56	0.26	4	8.16%	74.30	0.58	—	—	115.80	2.48
	$k = 15$	11	2.00%	54.56	0.20	4	8.16%	74.30	0.68	—	—	113.90	3.29
	$k = 20$	11	2.00%	54.56	0.37	4	8.16%	74.30	1.27	—	—	113.90	3.76

and $\#Opt_{ILP}$ is the number of requests solved successfully by the ILP. In the first row, $\#Opt = 20$, $MC = 63.62$ and $ET = 816.83$ dictate the number of requests that were successfully solved, the average multicast cost, and the average-elapsed execution time for $n = 30$, $q = 2$ by using the ILP. From the numerical results, we make the following observations:

(i) The elapsed execution time of NKSPH is much shorter than that of the ILP. For example, $ET = 0.10$ s or $ET = 0.18$ s in $n = 30$, $q = 4$, and $k = 8$ or $k = 20$ by using NKSPH, but $ET = 7094.21$ s by using the ILP.

(ii) Although NKSPH cannot always find the optimal solution, the solutions it has produced are close to optimal ones. For large-scale networks, the greater the value of k is, the more the multicast cost is reduced. For example, $Dev = 9.93\%$ for $n = 60$, $q = 4$, and $k = 8$, and $Dev = 4.08\%$, which is much smaller, for $k = 15$ or 20.

6. Conclusions

In this paper, an extended multicast routing and wavelength assignment problem with delay constraints in WDM networks with heterogeneous light splitting capabilities (MRWA-DC problem) was studied. The objective is to find a light-forest whose multicast cost is minimum. The MRWA-DC problem is NP-hard because it can be reduced from the minimum Steiner tree problem, which is already known to be NP-hard. To solve the problem to optimality, we formulated the problem in an ILP formulation. Because the ILP formulation cannot solve instances of the MRWA-DC problem in large-scale networks, a heuristic has been developed to derive approximate solutions in a reasonable time. Results from our computational study show that the ILP formulation can be used to solve the MRWA-DC problem in networks of a limited number of nodes. Statistics from computational experiments evince that the proposed heuristic can produce near-optimal solutions in an acceptable time.

In our study, we find that how to determine adaptive communication cost ratio, wavelength consumption ratio, and the value of k for different networks could be an interesting topic. Moreover, because a WDM network with wavelength conversion can route requests in a more flexible way, the cost of wavelength conversion needs to be included in multicast cost for finding an efficient light-forest. Nevertheless, for the WDM networks with sparse wavelength conversion, an extra constraint describing a node with/without wavelength conversion needs to be taken into consideration. Therefore, the problem is more difficult and worth further research.

References

Ballardie A. Core Based Tree (CBT) multicast routing architecture. Internet RFC 2201, September 1996.

Chen B, Wang J. Efficient routing and wavelength assignment for multicast in WDM networks. IEEE J Sel Areas Commun 2002;20:97–109.

Chen MT, Tseng SS. Multicast routing under delay constraint in WDM network with different light splitting. International computer symposium 2002 (ICS 2002), Taiwan, ROC.

Chlamtac I, Ganz A, Karmi G. Light-path communications: approach to high bandwidth optical WANs. IEEE Trans Commun 1992;40:1171–82.

Dijstra EW. Anode on two problems in connexion with graph. Numer Math 1959:269–71.

Eppstein D. Finding the $k\psi$ shortest paths. SIAM J Comput 1998;28(2):652–73.

Green PE. Fiber-optic networks. Cambridge, MA: Prentice-Hall; 1992.

Jia X-H, Du D-Z, Hu X-D. Integrated algorithm for delay bounded multicast routing and wavelength assignment in all optical networks. Comput Commun 2001;24:1390–9.

Kadaba BK, Jaffe JM. Routing to multiple destinations in computer network. IEEE Trans Commun 1983;31(3):343–51.

Kosiur D. IP multicasting: the complete guide to interactive corporate networks. Wiley Computer Publishing; 1998.

Kumar MS, Kumar PS. Static light-path establishment in WDM networks—new ILP formulations and heuristic algorithms. Comput Commun 2002;25:109–14.

Krishnaswamy RM, Sivarajan KN. Algorithms for routing and wavelength assignment based on solutions of LP-relations. IEEE Commun Lett 2001;5(10):435–7.

Lowe E. Current European WDM development trends. IEEE Commun Mag 1998;36(2):46–50.

Pascoal M, Martins E, Santos J. A new improvement for a k shortest paths algorithm. Invest Oper 2001;21:47–60.

Rouskas GN, Baldine I. Multicast routing with end-to-end delay and delay variation constraints. IEEE J Sel Areas Commun 1997;15(3):346–56.

Sahasrabuddhe LH, Mukherjee B. Light-trees: optical multicasting for improved performance in wavelength-routed networks. IEEE Commun Mag 1999;37:67–73.

Sreenath N, Krishna N, Reddy M, Mohan G, Siva Ram Murthy C. Virtual source based multicast routing in DWM networks with sparse light splitting. In: Proceedings of IEEE workshop on high performance switching and routing, May 2001.

Waxman BM. Routing of multipoint connections. IEEE J Sel Areas Commun 1988;6:1617–22.

Yan S, Ali M, Deogun J. Routing optimization of multicast sessions in sparse light-splitting optical networks. IEEE GLOBECOM 2001;4:2134–8.

Yang DN, Liao W. Design of light-tree based logical topologies for multicast streams in wavelength routed optical networks. IEEE INFOCOM 2003.

Zhang X, Wei J, Qiao C. Constrained multicast routing in WDM networks with sparse light splitting. In: Proceedings of INFOCOM 2000, May.