2005

Capacity Utilization versus Congestion Levels in wavelength Routing for large scale WDM Networks

Quoc Viet Phung
Edith Cowan University

Daryoush Habibi
Edith Cowan University

Hoang N. Nguyen
Edith Cowan University

Kungmang Lo
Edith Cowan University

10.1109/APCC.2005.1554177
This is an Author's Accepted Manuscript of: Phung, Q.V. , Habibi, D. , Nguyen, H. N., & Lo, K. (2005). Capacity Utilization versus Congestion Levels in wavelength Routing for large scale WDM Networks. Proceedings of Asia Pacific Conference on Communications. (pp. 822-826). Perth. IEEE. Available here

2005 Asia-Pacific Conference on Communications, Perth, Western Australia, 3 - 5 October 2005.

Capacity Utilization versus Congestion Levels in Wavelength Routing for Large Scale WDM Networks

Quoc V. Phung, Daryoush Habibi, *Senior Member*, *IEEE*, Hoang N. Nguyen, and Kungmeng Lo
School of Engineering and Mathematics
Edith Cowan University
Perth - Australia

***Abstract*— Wavelength Routing (WR) has been a key issue in WDM optical networks which carry huge amount of traffic aggregated from Internet Protocol (IP), Asynchronous Transfer Mode (ATM) and SDH/SONET layers. The problem, however, has been proved NP-hard. The *time complexity* and the *optimality of solutions* are two conflicting metrics. In addition, for optimization purposes, capacity utilization and network congestion level can be compromised to reduce the blocking probability of future connections. In this paper, we propose a heuristic approach for WR in large scale networks based on a balanced model developed for the capacity utilization and the congestion level. This is a two step approach, i.e K shortest paths (KSP) algorithm and a Path Selection Algorithm (PSA), and is applicable not only at network design phase but also for online provisioning where a number of traffic connections may be requested simultaneously. We investigate the *time complexity* and the *optimality of solutions* as metrics for comparing our approach and the ILP formulation for wavelength routing. Simulation results show that our approach yields very promising results in terms of the optimality of solutions whilst also applicable to very large scale networks, say 500 nodes or more.**

I. Introduction

Wavelength Routing (WR) in WDM networks is becoming more and more critical due to the rapidly increasing bandwidth required by its client layers such as IP, ATM and SONET/SDH. This problem is applicable in both phases of network design and network operation. In the network design phase, long-term/estimated traffic demands are joined and routed simultaneously to obtain optimal solutions. On the other hand, traffic demands in the operation phase are requested in sequence and established without any backtracking, hence it is difficult to achieve optimal solutions. WR can be modeled as an optimization problem, proven to be NP-hard [1]. Solutions to the WR problem can be either *optimal* or *near-optimal*. Optimal solutions can be obtained from Integer Linear Programming (ILP) formulation [2] or exhaustive searching [3]. These approaches are, however, *intractable* with larger network sizes (network nodes and links) and larger number of traffic connections. Heuristic approaches are often alternative solutions for computational time problem but offer near-optimal solutions.

WR, as an optimization problem, is subject to specific network performance objectives such as capacity utilization, congestion level or traffic delay. These objectives can conflict with one another. For example, when minimization of the capacity utilization is the only objective, the congestion level index or the total number of capacity used in some links may reach to their upper limits, although the total number of wavelength channels used in the network is minimized. Such links are blocked to future connections, and hence, in some cases the requests may be refused even though there are still many wavelength channels available.

In this paper, we propose a method which can compromise in a controlled manner between the two metrics of network performance, referred as *the capacity utilization* and denoted by f_{sum} and the *the network congestion level*, denoted by f_{max}. The aim is to reduce the blocking probability for future connections. Next, we devise a new heuristic approach to the wavelength routing problem for large scale networks. This is a two-step approach, namely K shortest paths (KSP) and Path Selection Algorithm (PSA), and is subject to a cost function, which can be a function of the number of wavelengths required (the capacity utilization), the congestion level, or the combination of these objectives constructed from the proposed method.

The rest of this paper is organized as follows. Section II reviews solutions to the WR problem. We propose our heuristic approach in Section IV. Simulation results of the proposed approach are presented and compared with ILP formulation in section V. Finally, Section VI summarizes the paper and proposes some future directions for this research.

II. Related work

The WR problem has been proven to be NP-hard [1]. Finding optimal solutions in literature has commonly been categorized into two approaches: the ILP formulation [4], [5] and the heuristics based on graph theory [6]–[9]. The ILP formulation is a classical model for obtaining the optimal solution and is only applicable to small networks. In [4], the authors propose an approximate approach using *randomized rounding* which consists of three phases: non-integral multi-commodity flow problem, path stripping and randomization. Non-integral multi-commodity flow relaxes $0-1$ integer formulations and solves the relaxed problem by linear programming. Path stripping converts the edge flows of commodity i to a set of paths τ_i that may carry the flow commodity i in the optimal case. Finally, randomization selects suitable path for commodity i from τ_i by casting a $|\tau_i|$ dice with face probability equal to the weights of the paths in τ_i. With a different approach, Tornatore M. *et al* [5] have proposed an ILP model based on source formulation that significantly reduces the number of decision variables and constraints compared to the classical ILP formulation. This formulation can result in optimal solutions but may still be intractable with increasing network size.

On the other hand, in heuristic solutions, finding shortest path (SP) [6] is the simplest method to obtain the optimal solution. However, this solution may result in a high level of congestion. The search with k-shortest paths in [7] offers a better performance but the computational complexity is extremely high. In recent years, several heuristic approaches for searching the optimal solution for the WR problem have been proposed [8], [9]. For example, in [9], the algorithm initially sets up lightpaths with SP algorithms, and then a re-routing process is performed to reduce the maximum number of wavelengths used. The authors have achieved significant results in terms of time complexity, but they do not consider the optimality of the solution.

III. A COMPROMISE BETWEEN THE CAPACITY UTILIZATION AND THE CONGESTION LEVEL

As mentioned in Section I, there are some intrinsic properties when provisioning the connections over the network as follows.

- There may be more than one route for connections which achieve the same minimum value of f_{sum}. The objective is then to select the solution with the lowest congestion.
- Similarly, there may be different routes which achieve the same minimum value of f_{max} and the objective will be to choose the solution that has the smallest number of capacity units used.

For this reason, we propose a method that can be controlled to strike a compromise between the two conflicting performance objectives. The cost function for this is given by:

$$f_{com} = Af_{sum} + Bf_{max} \tag{1}$$

where f_{com} is an integrated objective function of f_{sum} and f_{max}, and A and B are two non-negative constants. Let M and W be the number of network links and the maximum number of wavelength channels in links respectively, $0 \leq f_{sum} \leq MW$ and $0 \leq f_{max} \leq W$. We investigate the dependence of f_{com} on f_{sum} and f_{max} for different values of the constants A and B. Three possible cases are considered here.

- **Case 1:** $B = k$ and $A > kW$, $k \neq 0$

Let $A = kA' \quad \Rightarrow \quad A' > W$. The objective function f_{com} is given as: $f_{com} = k(A'f_{sum} + f_{max})$.

We assume that f_{sum} increases by 1, i.e. $f'_{sum} = f_{sum} + 1$ and f_{max} varies within its limited range, denoted as f'_{max}. Then the new value for f_{com}, denoted by f'_{com}, is calculated as: $f'_{com} = k(A'f_{sum} + f'_{max} + A')$.

Therefore $f'_{com} - f_{com} = k\,[A' - (f_{max} - f'_{max})]$.

Since $0 \leq f_{max}, f'_{max} \leq W$ and $A' > W$, then $-W \leq (f_{max} - f'_{max}) \leq W$.

$$\therefore \quad (f'_{com} - f_{com}) > 0 \text{ or } f'_{com} > f_{com}$$

The result implies that f_{com} is changed significantly even with a small change in f_{sum}. In this case, corresponding to each unit increasing in f_{sum}, f_{com} increases by a factor of A. Thus, f_{com} is still increasing even when f_{max} decreases by the maximum value (W).

Therefore, we conclude that in order to minimize f_{com}, we need to first minimize f_{sum} and then f_{max}.

- **Case 2:** $A = k$ and $B > kMW$

Let $B = kB' \quad \Rightarrow \quad B' > MW$. The objective function f_{com} is given as: $f_{com} = k(f_{sum} + B'f_{max})$.

Similar to the first condition, we examine the case when f_{max} is increased by 1, and f_{sum} varies within its range (represented as f'_{sum}).

$$f'_{com} - f_{com} = B' - (f_{sum} - f'_{sum})$$

Since $0 \leq f_{sum}, f'_{sum} \leq MW$ and $B' > MW$,
then $-MW \leq (f_{sum} - f'_{sum}) \leq MW$.

$$\therefore \quad (f'_{com} - f_{com}) > 0 \text{ or } f'_{com} > f_{com}$$

Therefore, in order to minimize f_{com}, we first minimize f_{max} and then minimize f_{sum}.

In summary, f_{com} can be represented as:

$$f_{com} = f_{sum} + \alpha f_{max} \tag{2}$$

where α is called *the weighting factor* and equals to $\frac{B}{A}$

If $\alpha < \frac{1}{W}$, then the first priority is to minimize the total capacity usage. However, if $\alpha > MW$, then network congestion level is minimized first

- **Case 3:** $\frac{1}{W} < \alpha < MW$

There is no mathematical model for the priority levels of the congestion level metric and the capacity utilization metric when the weighting factor is in range from $\frac{1}{W}$ to MW. The priority can be intuitively estimated as shown in Fig. 1. For the congestion level metric, the priority is highest for $\alpha \leq \frac{1}{W}$ and reduce to its lowest value for $\alpha \geq MW$. Conversely, the priority trends for capacity utilization is from the lowest to the highest when the weighting factor varies from $\frac{1}{W}$ to MW.

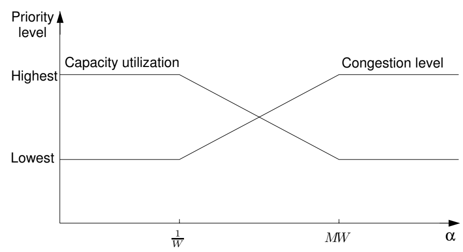

Fig. 1. Priority relationship between capacity efficiency and congestion levels

IV. Proposed heuristic approach to optimal solutions

In this sections, we present our approach that achieves near-optimal solutions for large scale networks. This consists of two steps: K-Shortest Paths (KSP) and Path Selection Algorithm (PSA). In KSP, for each connection, we generate the set of K shortest paths as candidate solutions for that connection. The purpose of the PSA step is to select suitable paths in all sets of K shortest paths of traffic connections with respect to a specific objective function, ie. the capacity utilization, the congestion level or the proposed combined objective. The PSA returns only one suitable route among the K shortest paths for each connection.

A. Network model and notations

We model our routing problem as follows.

- Let an undirected graph $G(V, E)$ be a physical network topology, where V is the set of N vertices representing network nodes, and $E =$ is set of M undirected edges representing the bi-directional optical fibers. Let W be the weight of each link e_l, representing the maximum number of wavelengths in the link.
- Let $T = \{t_1, t_2, \ldots, t_D\}$ be the set of D traffic requirements (traffic connections) over the network, where t_i denotes the connection between node pair (s_i, d_i).
- $P_i = \{p_i^{(1)}, p_i^{(2)}, \ldots, p_i^{(K)}\}$ is the set of K shortest paths of connection t_i, where $p_i^{(j)}$ denotes the j^{th} path of connection t_i.
- We refer the objective function, which is to minimize the total utilized capacity, to as $CapMin$.
- We refer the objective function, which is to minimize the congestion level in the network, to as $CongMin$.
- We refer the integrated objective function, which combines the above two objectives, to as $CombMin$.

B. KSP Algorithms

K shortest paths is a classical network and graph programming problem that has been widely studied [10]–[12]. Eppstein [10] proposes a k shortest path algorithm between a node pair in digraph in time $O(m + n\log n + k)$, where n is number of nodes, m is number of edges and k is number of paths. Martins *et al.* [12] present two algorithms for the k shortest paths problem, one based on label setting algorithm and another based on label correcting algorithm. Their results show that these algorithms perform better than the algorithm in [10]. In our study, we adopt the K shortest paths algorithm in [11], an improved version of [12]. This algorithm runs with time complexity $O(K \times m)$, and requires a space of $O(K \times n + m)$ for $k \leq \frac{m}{n}$, or $O(K \times n)$ if $K > \frac{m}{n}$), and is outline in Algorithm 1.

This KSP determines the K candidate routes for connections $t_i \in T$. A candidate p_i^j of connection t_i is represented as path-link form $p_i^j = \{b_l^{ij} | e_l \in E\}$, defined by:

$$b_l^{ij} = \begin{cases} 1 & \text{if path } p_i^{(j)} \text{ uses link } e_l \\ 0 & \text{otherwise} \end{cases}$$

Algorithm 1 K-Shortest Paths - KSPs

Input : An undirected graph $G(V, E)$, a pair of source and destination nodes (s, t), and K: the number of shortest paths required.

Output: A set of K-shortest paths over the network G from s to t.

1: To assure the possible repetition of the algorithm in a path between a pair of source-destination nodes (s, t), the given network is enlarged with a super source node S, and super destination nodes T, with zero cost arcs (S, s) and (t, T). We find the shortest tree from source node s to other nodes in the network and mark the shortest path $p_1 = \{s_0(= s), s_1, \ldots, s_{r-1}, s_r(= t)\}$ from s to t as the first shortest path.

2: Determine the first node s_h in p_1 such that s_h has more than a single incoming arc. If s'_h, of which the incoming arcs are the incoming arcs of s_h, except those coming from s_{h-1}, does not exist, then generate the node s'_h, else determine the next node s_i in p that has not alternate yet. The shortest path from s to s'_h ($d(s, s'_h)$) is calculated as:

$$d(s, s'_h) = \min_x\{d(s, x) + d(x, s'_h)\}$$

where (x, s'_h) are incoming arcs of s'_h.

3: For each $s_j \in \{s_i, \ldots, s_{r-1}\}$, generate s'_j following the same rules as s'_h, but with one more incoming arc of (s'_{j-1}, s'_j). Clearly, the shortest path from s to s'_j is the second shortest path from s to s_j. Therefore $p_2 = \{s_0, \ldots, s'_i, \ldots, s'_{r-1}, s_r(= t)\}$ is the second shortest path. Repeat step 2 for the shortest path $p_k (k = 2, 3, \ldots)$ to find the next shortest path until $k = K$.

C. Path Selection Algorithm (PSA)

The goal of PSA is to select suitable paths from the outcomes of the KSP step. The selection process has to satisfy the following conditions:

- The cost of the objective function has to be minimized. The cost function is the weighted sum of the capacity utilization and the congestion level.
- For each connection, only one path is selected.
- The total number of wavelengths used per link does not exceed the link capacity.

PSA is a greedy algorithm in which a route that has a highest cost amongst the K candidate routes will be removed until all above conditions are satisfied. The details are given in Algorithm 2.

V. Simulation Results

We examine the performance of our heuristic approach in three contexts: 1) We examine the efficiency of the proposed integrated objective function and compare it to $CapMin$ and $CongMin$. 2) We verify the capacity utilization and the congestion level of our approach in comparison to shortest path algorithm (SP). 3) The computational time of our heuristic is evaluated in large scale networks.

Algorithm 2 Path Selection Algorithm (PSA)

Input : An undirected graph $G(V, E)$, a path constant matrix P of traffic T, and a cost function c.

Output: Set of paths provisioned for T on graph G.

1: Let $f_l = \sum_i \sum_j b_l^{ij}, \forall e_l \in E$. Then $f_{sum} = \sum_l f_l$ is the total number of wavelength channels used in the network; and $f_{max} = \max(f_l)$ is the maximum congestion level.

2: Let $f^{i,k}$ be the value of the objective function after removing unmarked paths p_i^k out of P.
Find a path $p_{i_0}^{k_0}$ in P so that $f^{i_0,k_0} = \min_{i,k} f^{i,k}$.

3: If P_{i_0} contains only one path then we select and mark this path for connection t_{i_0}, otherwise the path is removed out of P.

4: Repeat step 2 and 3 until all $P_i, i = [1 \dots D]$ are selected.

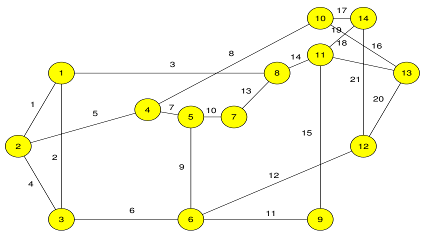

Fig. 2. NSFNET with 14 nodes and 21 links

A. Capacity utilization versus the network congestion level

Simulations are carried out on an undirected network NSFNET with 14 nodes, 21 links and $W = 16$ as shown in Fig. 2. We examine the routing using two popular objective functions ($CapMin$ and $CongMin$) and our proposed objective function ($CombMin$). With respect to the $CombMin$ objective, we consider two scenarios in which either $CongMin$ or $CapMin$ has higher optimization priority when $\alpha = \frac{1}{W+1} < \frac{1}{W}$ or $\alpha = (MW+1) > MW$ respectively. We use $CombCapMin$ to indicate higher priority for capacity utilization, and $CombCongMin$ to denote that the higher priority is assigned to network congestion. We use the ILP model to obtain pairs of $\{f_{sum}, f_{max}\}$ for each objective in the NSFNET network and random traffic matrices with $D = \{60, 70, 89, 90\}$. Fig. 3 and Fig. 4 show the simulation results in which we compare the resource utilization, f_{sum}, and the congestion levels, f_{max}. It is logically observed that the capacity utilization and the congestion levels increase when the number of traffic requests increases.

- **Capacity utilization:**

It can be seen in Fig. 3 that the capacity utilization in the network increases with the rising of traffic requirements. It is worth noting that the capacity utilization corresponding to $CombCapMin$ coincides with those corresponding to $CapMin$, and has better capacity utilization than $CombCongMin$ and $CongMin$. These results validate the

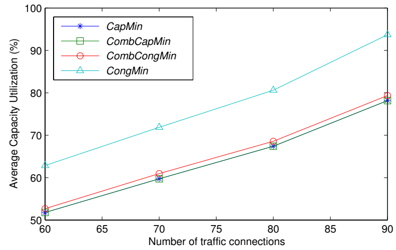

Fig. 3. Capacity utilization versus the number of required connections

theory that we developed in Section III, that is, if the weighting factor α is less than $\frac{1}{W}$, then the capacity utilization objective has a higher priority in optimization. In this simulation, we choose $\alpha = \frac{1}{W+1} < \frac{1}{W}$. Therefore, we can rank the network capacity utilization resulting from these algorithms as: $CapMin = CombCapMin > CombCongMin > CongMin$.

- **Congestion level:**

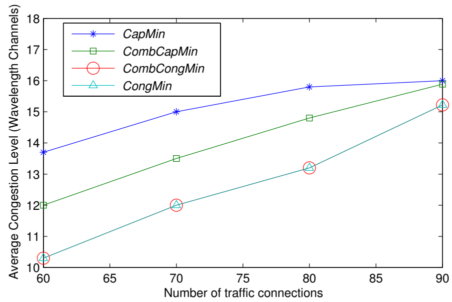

Fig. 4. Congestion levels versus the number of required connections

In contrast to the capacity metric, the congestion levels achieved, as shown in Fig. 4, can be ranked from higher efficiency to lower efficiency as: $CongMin = CombCongMin > CombCapMin > CapMin$. However, as discussed, the required capacity of $CongMin$ is more than the other $CapMin$ and $CombMin$ with the same network environment and traffic requirements. In addition, we observe from Fig. 4 that the congestion level achieved with $CombCapMin$ is a little bit higher than $CongMin$ and $CombCong$, but always lower than $CapMin$. This is again in agreement with the theory which we proposed in Section III.

Furthermore, we notice from Fig. 4 that with the higher network load, the congestion levels achieved with the three objective functions are very close due to bandwidth limitations, but the capacity utilization achieved with $CapMin$ and $CombMin$ is still much lower than $CongMin$. The explanation is that all possible optimal solutions found with $CongMin$ result in the same congestion level, and the ILP solver picks an arbitrary optimal solution. This may lead to a higher number of wavelengths used with $CongMin$, compared to $CapMin$ or $CombMin$.

In brief, a balance between capacity utilization and congestion level can be controlled through our proposed integrated objective function by adjusting the weighting factors.

B. A comparison between our heuristic and the shortest path algorithm

Traffic demands of D connections are varied from $[50 \ldots 105]$ for $K = 5$. The capacity utilization and the congestion level are measured and compared between the shortest path (SP) algorithm and our approach. These results are shown in Fig. 5.

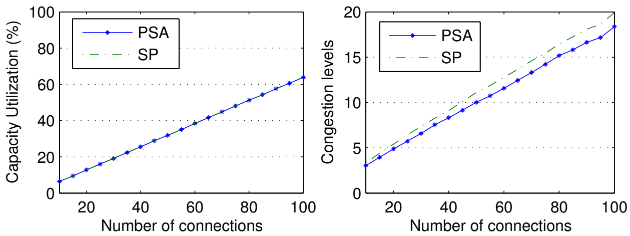

Fig. 5. Shortest Path (SP) versus Path Selection Algorithm (PSA)

We observe that the PSA and the SP result in the same minimal capacity utilization while the congestion achieved with the PSA is better than the SP algorithm. With the same network configuration where the number of wavelengths is limited to 16 in each link, in this simulation, our heuristic is able to serve for traffic pattern of around 90 connections on average which is much better in comparison to those resulting from the SP algorithm with just over 70 connections.

C. The time complexity in large scale networks

We randomly generate physical networks with $N = \{200, 300, 400, 500\}$ nodes, average nodal degree of 3, and traffic demands $D = 400$. We assume that the connectivity of the network is assured. The time complexity is measured with $K = \{2, 3, 4\}$ as in Fig. 6.

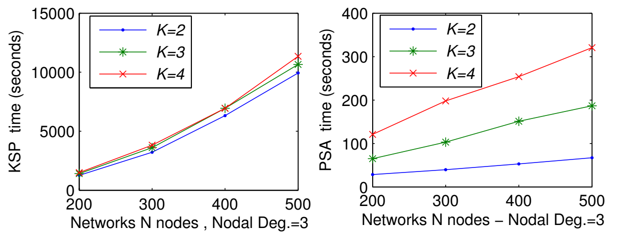

Fig. 6. The computational time of our approach in large scale networks

The total computational time of our approach includes the time taken by the KSP algorithm on the left figure and the time of the PSA algorithm on the right. It is worth noting that the PSA takes much less time than the KSP algorithm. For example, for N=500 and $K = 2$, the computational time of the PSA is about 70 seconds, it is about 10000 seconds for the KSP. In addition, the the PSA computational time does not increase much with the number of network nodes. For example, with N=200 and K=4 it is around 120 seconds and increases to only 320 seconds for N=500 and K=4. Therefore, the computational time of our approach is mostly dependant on the KSP algorithms, but the PSA algorithm has the advantage that it can deal with very large scale networks.

VI. Conclusion

Although wavelength routing problem has been extensively studied in the literature, it still remains a critical problem and has been proved to be NP-complete. Time complexity and optimality of solutions are two conflicting metrics, and it is important to find approaches that can balance these. In this paper we proposed a new fast heuristic approach for wavelength routing in large scale network in which we can obtain optimal solutions but still keep the congestion levels low.

We are currently tackling the same problem in the context of network survivability, particularly with path protection, in which two distinct lightpaths (or two disjoint paths) have to be established from a source node to a destination node of a connection, one being the primary path and the other being the backup path. This is called a path-pair. We believe that the same results could be achieved for network survivability if we can devise an efficient algorithm to find K-shortest path-pairs instead of K-shortest paths as presented in this paper.

References

[1] I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath communication: An approach to high-bandwidth optical WANs," *IEEE Transactions on Communiations*, vol. 40, no. 7, pp. 1171–1182, 1992.

[2] H. Zang, J. P. Jue, and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks," *Optical Networks Magazine*, vol. 1, no. 1, pp. 47–60, 2000.

[3] E. Hyytia and J. Virtamo, "Wavelength assignment and routing in WDM networks," *in Nordic Teletraffic Seminar (NTS)*, vol. 14, 1998.

[4] D. Banerjee and B. Mukherjee, "A practical approach for routing and wavelength assigment in large wavelength-routed optical netwroks.," *IEEE Journal on Selected Area in Communications*, vol. 14, no. 5, pp. 903–908, 1996.

[5] M. Tornatore, G. Maier, and A. Pattavina, "WDM network optimization by ILP based on source formulation," in *Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE Infocom 2002*, vol. 3, pp. 1813– 1821, 2002.

[6] C. Chen and S. Banerjee, "A new model for optimal routing in alloptical networks with scalable number of wavelength converters," in *Proceedings of the IEEE Global Telecommunications Conference*, vol. 2, (Singapore), pp. 993–997, 1995.

[7] M. Alanyali and E. Ayanoglu, "Provisioning algorithms for WDM optical networks," *IEEE/ACM Transactions on Networking*, vol. 7, no. 5, 1999.

[8] H. V. Madhyastha and N. Balakrishnan, "An efficient algorithms for virtual-wavelength-path routing minimizing average number of hop," *IEEE Journal On Selected Areas in Communications*, vol. 21, no. 9, 2003.

[9] J. H. Siregar, H. Takagi, and Y. Zhang, "Fast routing and wavelength assignment heuristics for large-scale wdm optical networks," *IEICE Transactions on Communications*, vol. E86-B, no. 12, 2003.

[10] D. Eppstein, "Finding the *k* shortest paths," *SIAM Journal on Computing*, vol. 28, no. 2, pp. 652–673, 1998.

[11] E. D. Q. V. Martins and J. L. E. D. Santos, "A new shortest paths ranking algorithm," *Investigao Operacional*, vol. 20, no. 1, pp. 47–62, 2000.

[12] E. D. Q. V. Martins, M. M. B. Pascoal, and J. L. E. D. Santos, "The *k* shortest paths problem," tech. rep., CISUC, June 1998.