Optimal Resource Allocation and Fairness Control in All-Optical WDM Networks

Kayvan Mosharaf, *Student Member, IEEE*, Jérôme Talim, and Ioannis Lambadaris, *Member, IEEE*

***Abstract*—This paper investigates the problem of optimal wavelength allocation and fairness control in all-optical wavelength-division-multiplexing networks. A fundamental network topology, consisting of a two-hop path network, is studied for three classes of traffic. Each class corresponds to a source-destination pair. For each class, call interarrival and holding times are exponentially distributed. The objective is to determine a wavelength allocation policy in order to maximize the weighted sum of users of all classes (i.e., class-based utilization). This method is able to provide differentiated services and fairness management in the network. The problem can be formulated as a Markov decision process (MDP) to compute the optimal allocation policy. The policy iteration algorithm is employed to numerically compute the optimal allocation policy. It has been analytically and numerically shown that the optimal policy has the form of a monotonic nondecreasing switching curve for each class. Since the implementation of an MDP-based allocation scheme is practically infeasible for realistic networks, we develop approximations and derive a heuristic algorithm for ring networks. Simulation results compare the performance of the optimal policy and the heuristic algorithm, with those of complete sharing and complete partitioning policies.**

***Index Terms*—Dynamic wavelength allocation, Markov decision process (MDP), monotonic optimal policy, wavelength-division multiplexing (WDM).**

I. Introduction

WAVELENGTH-DIVISION-MULTIPLEXING (WDM) networks, using wavelength routing, is a promising candidate to handle the huge bandwidth demand of future backbone wide-area networks. In wavelength routing networks, each optical path must be established with a specific wavelength between each source-destination pair. This is known as *wavelength continuity constraint* and can be relaxed by using wavelength converters (WCs) at intermediate nodes [1]. The routing and wavelength assignment (RWA) problem is an important issue in WDM networks to choose a suitable route and wavelength among the possible selections. RWA is usually divided into two separate problems: wavelength assignment problem and routing problem.

Many heuristic algorithms such as Random, Least-Used, Most-Used, and First-Fit wavelength assignments have been already proposed to assign wavelengths in WDM networks with dynamic traffic [1], [2]. The objective of these algorithms is typically to minimize the overall call blocking probability or maximize the overall utilization in a single-class network. Some analytical studies have been performed to calculate the blocking probabilities or network utilization for different assignment strategies [3]–[6]. In these studies, it can be seen that all users have the same priority (class). As a result, the differentiation among different users is not possible.

In this paper, we investigate the wavelength allocation problem and fairness control issue for different classes of users with dynamic traffic. With the objective of maximizing the weighted sum of class-based utilization, we define a Markov decision process (MDP) model, based on which the optimal wavelength allocation policy is determined [7]. In many admission control and resource allocation problems in telecommunications, it has been shown that under some conditions, the optimal policy of a MDP exists and it is stationary and monotone [8], [9]. In [10], the multimodularity, submodularity, and convexity properties are investigated in queueing systems. These properties imply the monotonicity of optimal policy in the context of stochastic control. In this study, by using the multimodularity of the cost function and the induction method, we prove that the optimal policy is a nondecreasing switching curve. Moreover, the policy iteration algorithm [7] is deployed to determine numerically the optimal policy. Using the properties of the optimal policy, we develop a simple heuristic algorithm to provide fairness in WDM ring networks.

The rest of the paper is organized as follows. In Section II, we describe the problem and make some assumptions regarding the network. In Section III, the problem is formulated in an MDP framework with the discounted cost model. Section IV shows the multimodularity of the value function and provides the structure of optimal policy. Section V introduces our heuristic allocation algorithm and Section VI compares the performance of our proposed allocation algorithms with other standard policies. Conclusions are presented in Section VII.

II. Problem Description

We first consider a two-hop path network topology for a single fiber circuit-switched wavelength routing network, as depicted in Fig. 1(a). This fundamental topology can represent the link-load correlation model [5], [6]. The total number of available wavelengths in the system is W. Traffic is divided into three classes. Each class corresponds to a different source-destination pair. Class 1 (respectively, Class 3) consists of the users that use hop H_1 (respectively, H_2); Class 2 includes the customers that use both hops H_1 and H_2. If a WC is deployed at node 2, then a Class 2 call is accepted whenever there is at least one available wavelength in both hops. Without a WC at this node, the same wavelength must be available in both hops to accommodate a Class 2 call. Any arriving call is blocked

Manuscript received April 21, 2004; revised November 11, 2004.

The authors are with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada (e-mail: mosharaf@sce.carleton.ca).

Digital Object Identifier 10.1109/JSAC.2005.851791

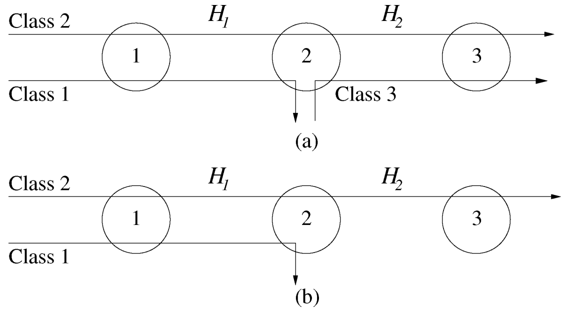

Fig. 1. Two-hop network topology. (a) Three-class system. (b) Two-class system.

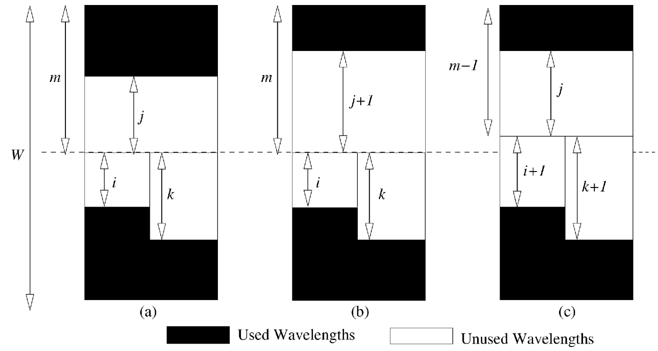

Fig. 2. Possible states after the departure of a Class 2 call. (a) Original state. (b) Final state after taking action $a = 0$. (c) Final state after taking action $a = -1$.

when all wavelengths along its path are used. Blocked calls do not interfere with the system. Arrivals of Class ℓ calls, $\ell = 1$, 2, 3, are distributed according to Poisson process with rate λ_ℓ. The call holding time of a Class ℓ call is exponentially distributed with mean μ_ℓ^{-1}.

Wavelength allocation policy is a particular problem related to resource allocation policies. In general, current wavelength allocation strategies are deploying heuristic algorithms such as complete sharing (CS) and complete partitioning (CP) [12]. When implementing CS, no wavelength is reserved for any class. In addition, an arriving call will be accepted if at least one wavelength is available throughout all the hops along its path. Although the global network utilization is high in this case, this approach is greedy, and it is suboptimal if different classes of users provide different rewards for the same grade of service. When deploying CP policy, each class is assigned a constant number of wavelengths that cannot be used by calls from the other classes. Hence, it supports service differentiation and controls class-based blocking probabilities. However, CP policy may not maximize the overall utilization of the available resources.

To improve the system performance in a dynamic environment, it would be essential to assign a certain number of wavelengths to each class as a function of the current number of customers from different classes. This paper investigates a dynamic wavelength allocation policy, which is called dynamic partitioning (DP) hereafter. It consists of determining the appropriate number of wavelengths allocated to each class taking into account the current state of the system, with the objective of maximizing the weighted sum of the number of calls for each class. This approach can take advantage of both CP and CS policies.

III. MDP Formulation

In this section, we present an MDP approach to the problem described in Section II. In order to model the system, we first define the state of the system and then introduce the appropriate cost function.

Let n_ℓ denote the number of Class ℓ calls in the system and m be the number of wavelengths currently allocated to Class 2 calls and being used in both hops H_1 and H_2. We first formulate the problem when there is a WC at node 2. Thus, for any (n_1, n_2, n_3), one can derive $i = W - m - n_1$, $j = m - n_2$, and $k = W - m - n_3$ as the numbers of available wavelengths reserved for Class 1, Class 2, and Class 3 users, respectively [refer to Fig. 2(a)]. Therefore, the four-component vector (i, j, k, m) completely characterizes the system. Let

$$S = \{(i, j, k, m) |\, 0 \le i \le W - m,\ 0 \le j \le m,\ 0 \le k \le W - m,\ 0 \le m \le W\}$$

be the state space of the system and let s_t denote the state of the system at time t. Based on the statistical assumptions, $\{s_t, t \ge 0\}$ is a continuous-time Markov chain whose transitions are either the event of a call request arrival or call departure.

When removing the WC from node 2, the *wavelength continuity constraint* must be satisfied. Let Λ_1 and Λ_2 denote the sets of available wavelengths on H_1 and H_2, respectively. Let q denote the number of wavelengths that are available on both H_1 and H_2 (i.e., $q = |\Lambda_1 \cap \Lambda_2|$). Variable q shows the number of Class 2 calls that can be accepted in the system. In a system without WC, if $q = 0$, then a Class 2 call will be blocked. In this case, the state space of the system is

$$S' = \{(i, j, k, m, q) |\, 0 \le i \le W - m,\ 0 \le j \le q,\ 0 \le k \le W - m,\ 0 \le m \le W,\ 0 \le q \le m\} \quad (1)$$

where $i = W - m - n_1$, $j = m - n_2$, and $k = W - m - n_3$.

One can note that for a given q, if $(i, j, k, m, q) \in S'$ then $(i, j, k, m) \in S$. Thus, both systems with and without WC have similar structures. In the following, we only focus on a system with WC.

To simplify the notation, we drop index t from s_t and introduce the following operators.

- $D_\ell : S \mapsto S$, $\ell = 1, 2, 3$; *departure operator*, describing the change of the state of the system at a Class ℓ departure time.
  - $D_1 s = (i+1, j, k, m)$, $D_2 s = (i, j+1, k, m)$, $D_3 s = (i, j, k+1, m)$.
- $A_\ell : S \mapsto S$, $\ell = 1, 2, 3$; *arrival operator*, describing the change of the state of the system at a Class ℓ arrival time.
  - $A_1 s = ((i-1)^+, j, k, m)$, $A_2 s = (i, (j-1)^+, k, m)$, $A_3 s = (i, j, (k-1)^+, m)$, where $x^+ = \max(0, x)$ and $s = (i, j, k, m)$.

According to the above operators, we can define the set of possible events as $\mathcal{E} = \{A_1, A_2, A_3, D_1, D_2, D_3\}$.

Investigating DP policy involves the determination of wavelength allocation as a function of the current state of the system

$s=(i,j,k,m)$, and the *event* $e \in \mathcal{E}$. The objective, then, is to maximize the usage of the optical resources. Equivalently, this can be translated into maximizing weighted sum of the number of call from different classes. This problem can be formulated as an MDP [13]. Let us describe the model more accurately in the following paragraphs.

Decision epochs take place only at departure times, when one call terminates and one wavelength becomes available. *The released wavelength may be reserved for the same class or switched to use by calls of other classes.* This is shown in Fig. 2, where the initial state is $s(i,j,m)$. In Fig. 2(b), the final state of the system is presented after a departure of Class 2 call, when we decide to keep the wavelength for the same class. Fig. 2(c) depicts the state of the system when the policy maker decides to reserve the wavelength for Class 1 and Class 3 users. Thus, the decision results in leaving m unmodified [Fig. 2(b)] or *decreasing* it by 1 [Fig. 2(c)]. In a similar fashion, when considering the termination of a Class 1 (or 3) call, the decision will result in *increasing* the value of m by 1, or leaving it unchanged.

When the system is in state s and the event e has just occurred, the decision maker takes *action* a from a set of possible actions $\mathcal{A}(s,e)$, where

$$\mathcal{A}(s,e)=\begin{cases}\{\emptyset\}, & \text{if } e=A_1 \text{ or } A_2 \text{ or } A_3\\ \{0,+1\}, & \text{if } e=D_1 \text{ or } D_3\\ \{-1,0\}, & \text{if } e=D_2\end{cases}.$$

Let $P_a,\ a \in \mathcal{A}(s,e)$ be the *policy operator* describing the change of state of the system when applying action a after a departure.

- $D_1P_0s=\big(i+1,j,k,m\big)$.
- $D_1P_1s=\big(i,(j+1)^-,(k-1)^+,(m+1)^-\big)$.
- $D_2P_0s=\big(i,j+1,k,m\big)$.
- $D_2P_{-1}s=\big((i+1)^-,j,(k+1)^-,(m-1)^+\big)$.
- $D_3P_0s=\big(i,j,k+1,m\big)$.
- $D_3P_1s=\big((i-1)^+,(j+1)^-,k,(m+1)^-\big)$, where $x^+=\max(0,x)$, $x^-=\min(W,x)$, and $s=(i,j,k,m)$.

This initial continuous-time MDP can be converted into an equivalent discrete-time MDP by applying the *uniformization* technique [13]. To do so, we introduce a random sampling rate ν defined as $\nu := W(\mu_1+\mu_2+\mu_3)+\lambda_1+\lambda_2+\lambda_3$. In the original MDP, the system is observed at the times of actual transitions. After applying uniformization technique, the system may be observed between two consecutive actual transitions, which are called "fictitious" or "dummy" transitions. When considering the discrete-time MDP, only one single transition can occur during each time slot. Let s_n denote the state of the system for the equivalent discrete-time MDP during time slot n. A transition can correspond to an event of: 1) Class 1 call arrival or departure; 2) Class 2 call arrival or departure; 3) Class 3 call arrival or departure; and 4) fictitious (or dummy) transition.

To complete the MDP description, we need to define the reward (or cost) function [13]. As mentioned previously, our objective is to determine a wavelength allocation policy that maximizes the weighted sum of n_1, n_2, and n_3. Therefore, we define the one-step reward function which expresses the instant reward when the state of the system is s

$$\begin{aligned}R(s)&=\alpha n_1+\beta n_2+\delta n_3\\ &=(\alpha+\delta)W-(\alpha i+\beta j+\delta k+(\alpha+\delta-\beta)m),\ s\in S\end{aligned} \quad (2)$$

where α, β, and δ are the weights assigned to Classes 1, 2, and 3 users, respectively. Note that one can maximize $R(s)$ by minimizing $\alpha i+\beta j+\delta k+(\alpha+\delta-\beta)m$. Consequently, instead of maximizing reward function (2), we can minimize one-step cost function defined by

$$C(s)=\alpha i+\beta j+\delta k+(\alpha+\delta-\beta)m=\mathbf{B}\cdot s^T,\ s\in S \quad (3)$$

where vector $\mathbf{B} := (\alpha,\beta,\delta,(\alpha+\delta-\beta))$.

The choice of the weights α, β, and δ has an impact on wavelength occupancy by each class and on their respective blocking probabilities. By assigning appropriate values to these weights, we can provide service differentiation based on classes of traffic.

We also define policy $\pi=(\pi_1,\pi_2,\ldots\pi_n)$ for n-stage finite-horizon problem such that π_i is the action applied after the ith event. Based on the one-step cost function, we can apply the results of discounted finite-horizon model [7]. In order to estimate the expected cost under a policy π and at the time step n, we can define n-stage finite-horizon γ-discounted value function as [7]

$$V_n^\pi(s)=E^\pi\left[\sum_{l=0}^{n-1}\gamma^l C(s_l)\mid s_0=s\right] \quad (4)$$

where $0<\gamma<1$ is the discount factor, $\pi=(\pi_1,\pi_2,\ldots,\pi_n)$ is the allocation policy, and E^π denotes the conditional expected cost given that the initial state is s, while the decision maker applies the policy π. The optimal policy $\Pi_n(s)$ and the minimum value function $V_n(s)$ are given by

$$\Pi_n(s)=\arg\min_\pi V_n^\pi(s) \text{ and } V_n(s)=\min_\pi V_n^\pi(s).$$

The optimal value function and the optimal policy can be computed by using the following recursive scheme, known as the relative value iteration algorithm [7]

$$V_{n+1}(s)=\min_\pi\left[C(s)+\gamma\sum_{s'}P_{ss'}^a V_n(s')\right] \quad (5)$$

where $P_{ss'}^a=P(s_{n+1}=s'\mid s_n=s,\pi_n=a)$ is the transition probability to jump from state s to state s' when applying action a.

Using the uniform sampling rate ν, introduced in the uniformization procedure, the transition probabilities can be written [11]

$$P_{ss'}^a\times\nu=\begin{cases}\lambda_1, & \text{if } s'=A_1s\\ \lambda_2, & \text{if } s'=A_2s\\ \lambda_3, & \text{if } s'=A_3s\\ n_1\mu_1, & \text{if } s'=D_1P_0s,\ a=0\\ n_1\mu_1, & \text{if } s'=D_1P_1s,\ a=1\\ n_2\mu_2, & \text{if } s'=D_2P_0s,\ a=0\\ n_2\mu_2, & \text{if } s'=D_2P_{-1}s,\ a=-1\\ n_3\mu_3, & \text{if } s'=D_3P_0s,\ a=0\\ n_3\mu_3, & \text{if } s'=D_3P_1s,\ a=1\\ F, & \text{if } s'=s\end{cases}$$

where $F = W(\mu_1+\mu_2+\mu_3)-n_1\mu_1-n_2\mu_2-n_3\mu_3$. Replacing $P^a_{ss'}$ in (5) yields

$$\begin{aligned} V_{n+1}(s) = &C(s) + \frac{\gamma}{\nu}\Big[F\,V_n(s) + \lambda_1 V_n(A_1 s) \\ &+ \lambda_2 V_n(A_2 s) + \lambda_3 V_n(A_3 s) \\ &+ \mu_1 n_1 \min\{V_n(D_1P_0s), V_n(D_1P_1s)\} \\ &+ \mu_2 n_2 \min\{V_n(D_2P_0s), V_n(D_2P_{-1}s)\} \\ &+ \mu_3 n_3 \min\{V_n(D_3P_0s), V_n(D_3P_1s)\}\Big]. \end{aligned} \quad (6)$$

From (6), it can be noticed that at a Class 1 (respectively, Class 3) call termination time, the optimal action is $a = 0$, if $V_n(D_1P_0s) < V_n(D_1P_1s)$ (respectively, $V_n(D_3P_0s) < V_n(D_3P_1s)$), and $a = 1$, otherwise. Similarly, after a Class 2 departure, the policy maker takes action $a = 0$ if $V_n(D_2P_0s) < V_n(D_2P_{-1}s)$, and $a = -1$, otherwise.

IV. Structure of the Optimal Policy

In this section, we will prove that for the discounted cost problem, the optimal allocation policy is a switching curve of threshold type. We first show this property for two classes of users, which is less complex and easier to present. In what follows, we use the term *two-class system* when system includes Classes 1 and 2. *Proposition 1*, which is based on value iteration and induction method, proves that the optimal value function for a two-class system is multimodular. In *Theorem 1*, using multimodularity of value function, we show that optimal policy is a monotonic nondecreasing switching curve. In *Proposition 2* and *Theorem 2*, we generalize the results of the two-class system for the original problem with three classes.

Consider a two-class system shown in Fig. 1(b). Let m be the number of wavelengths allocated to Class 2 calls. Therefore, the state of the system can be expressed by $s = (i, j, m)$. We consider similar definition and operators for this system. The optimal value function for this case can be shown by

$$\begin{aligned} V^*_{n+1}(s) = &\mathbf{B}^* \cdot s^T + \gamma[\lambda_1 V^*_n(A_1 s) + \lambda_2 V^*_n(A_2 s) \\ &+ (W(\mu_2+\mu_1) - n_2\mu_2 - n_1\mu_1)V^*_n(s) \\ &+ \mu_1 n_1\ \min\{V^*_n(D_1P_0s), V^*_n(D_1P_1s)\} \\ &+ \mu_2 n_2\ \min\{V^*_n(D_2P_0s), V^*_n(D_2P_{-1}s)\}] \end{aligned} \quad (7)$$

where vector $\mathbf{B}^* := (\alpha, \beta, (\alpha-\beta))$ and $\mathbf{B}^* \cdot s^T$ is the one-step cost.

If the value function has some properties such as *convexity*, *submodularity*, or more generally *multimodularity*, then its computation time will be remarkably decreased and we can show that the associated optimal policy will have a simple switching curve. To show the multimodularity of the value function, we first introduce a class of multimodular functions $\mathcal{F}$ as follows [10].

Definition 1: A function $f \in \mathcal{F}$ such that $f : \mathbb{N}^3 \mapsto \mathbb{R}$ is multimodular, if for all $(i, j, m) \in \mathbb{N}^3$

$$\begin{aligned} &f(i,j,m) + f(i,j+1,m+1) \\ &\quad\le f(i+1,j,m) + f(i-1,j+1,m+1) \end{aligned} \quad (8)$$

$$\begin{aligned} &f(i,j,m) + f(i-1,j,m+1) \\ &\quad\le f(i,j-1,m) + f(i-1,j+1,m+1) \end{aligned} \quad (9)$$

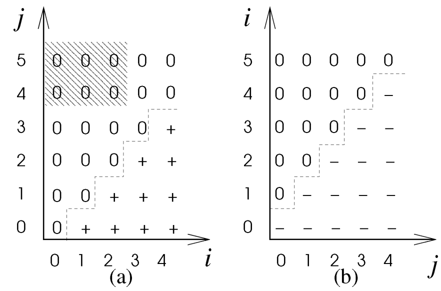

Fig. 3. Illustrative example. Optimal policy for a two-class system, $W = 10$, $k = 5$, $\lambda_1 = \lambda_2 = 5$, $\alpha = 1$, and $\beta = 0.5$. (a) $e = D_1$. (b) $e = D_2$. Actions $a = +1, 0$, and -1 are depicted by "+," "0," and "−," respectively.

$$\begin{aligned} &f(i,j,m) + f(i-1,j+1,m) \\ &\quad\le f(i,j,m-1) + f(i-1,j+1,m+1). \end{aligned} \quad (10)$$

To interpret the multimodularity of f, as an example. we rewrite (8) as follows:

$$\begin{aligned} &f(i-1,j+1,m+1) - f(i,j+1,m+1) \\ &\quad\ge f(i,j,m) - f(i+1,j,m). \end{aligned} \quad (11)$$

This equivalent inequality shows that function f has monotonic increasing differences. In the following, we show that the value function is multimodular and satisfies inequalities (8)–(10). This leads us to show that the optimal policy is a monotonic switching curve.

Proposition 1: The value function for a two-class system $V^*_{n+1}(i, j, m)$, is multimodular in i, j, and m.

Proof: The proof of this proposition, which is based on induction, is given in Appendix A. ∎

We now derive the structure of the optimal policy and show that it is nondecreasing switching curve. More specifically, we will prove that after a termination of Class ℓ call, there is a nondecreasing switching curve which partitions the state space into two regions. One is *transferring region* (i.e., for all states belonging to this region, the optimal policy transfers the released wavelength to the other class), whereas the other is *keeping region* (i.e., in all states within this region, the optimal policy keeps the released wavelength to Class ℓ calls). As an illustrative example, the structure of the optimal policy is shown in Fig. 3.

Theorem 1: For each class of users in a two-class system, the optimal wavelength allocation policy is a monotonic nondecreasing switching curve.

Proof: To prove this theory, we consider the two following cases.

Case 1: Optimal policy after a Class 1 departure: Suppose that the state of the system is $s = (i, j, m)$ and a Class 1 call terminates. We first claim that if the optimal action is to keep the released wavelength for Class 1 (i.e., $a = 0$), then for all states $s' = (i', j, m)$ with $i' < i$, the optimal action will be equal to zero, as well, which is in agreement with intuition. Because of taking action $a = 0$ as optimal action in state s, it can be seen that: $V^*_n(i+1, j, m) - V^*_n(i, j+1, m+1) < 0$. By using multimodularity of $V^*_n(s)$, inequality (8), one can show that

$$\begin{aligned} &V^*_n(i,j,m) - V^*_n(i-1,j+1,m+1) \\ &\quad\le V^*_n(i+1,j,m) - V^*_n(i,j+1,m+1) < 0. \end{aligned} \quad (12)$$

The above inequality shows that if a Class 1 call departs while system state is $(i-1, j, m)$, then the policy maker again selects action $a = 0$.

Now, suppose that the system's state is $s'' = (i, j', m)$ with $j' > j$ (i.e., the number Class 2 users in state s'' is smaller than the one in state $s = (i, j, m)$). In this case, we can use (9) to show that

$$V_n^*(i+1, j+1, m) - V_n^*(i, j+2, m+1) \leq V_n^*(i+1, j, m) - V_n^*(i, j+1, m+1) < 0. \quad (13)$$

Therefore, we prove that if the state of the system is $s = (i, j, k)$ and the controller takes action $a = 0$ after a Class 1 termination as the optimal action, then for all states $s' = (i', j, m)$ and $s'' = (i, j', m)$ with $i' < i$ and $j' > j$, action $a = 0$ will be taken as the optimal action.

Case 2: Optimal policy after a Class 2 departure: An analogous result can be proved for Class 2. ∎

As an example, one can consider Fig. 3(a). Suppose that the state of the system is $s = (3, 4, 5)$ and action $a = 0$ is taken in this state, then for all states with $i' < 3$ and $j' > 4$, the optimal action is $a = 0$.

We now add Class 3 users to the problem and determine analytically the properties of the optimal policy. The problem formulation can be extended for a three-class system by generalizing *Definition 1*, *Theorem 1*, and *Proposition 1*.

Definition 2: A function $f \in \mathcal{F}$ such that $f : \mathbb{N}^4 \mapsto \mathbb{R}$ is multimodular, if for all $(i, j, k, m) \in \mathbb{N}^4$

$$f(i,j,k,m) + f(i,j+1,k-1,m+1) \leq f(i+1,j,k,m) + f(i-1,j+1,k-1,m+1) \quad (14)$$

$$f(i,j,k,m) + f(i-1,j,k-1,m+1) \leq f(i,j-1,k,m) + f(i-1,j+1,k-1,m+1) \quad (15)$$

$$f(i,j,k,m) + f(i-1,j+1,k,m+1) \leq f(i,j,k+1,m) + f(i-1,j+1,k-1,m+1) \quad (16)$$

$$f(i,j,k,m) + f(i-1,j+1,k-1,m) \leq f(i,j,k,m-1) + f(i-1,j+1,k-1,m+1). \quad (17)$$

Proposition 2: The value function, $V_{n+1}(i, j, k, m)$, is multimodular in i, j, k, and m.

Proof: The proof is quite similar to that of *Proposition 1*. It is given in Appendix B. ∎

Using the result of *Proposition 2*, we prove that the optimal policy is monotonic.

Theorem 2: The optimal policy for a three-class system is a monotonic nondecreasing switching curve.

Proof: The major contribution of this section will be proved in this part. Suppose that the state of the system is $s = (i, j, k, m)$ and a Class 2 call departs from the network and the optimal action is $a = 0$, then we can see that: $V_n(i, j+1, k, m) - V_n(i+1, j, k+1, m-1) < 0$. According to multimodularity of $V_n(s)$, inequality (15), one can derive

$$V_n(i,j,k,m) - V_n(i+1,j-1,k+1,m-1) \leq V_n(i,j+1,k,m) - V_n(i+1,j,k+1,m-1) < 0. \quad (18)$$

This illustrates that if a Class 2 call departs while the state of the system is $(i, j-1, k, m)$, then the optimal action is again $a = 0$. As a result, the optimal action remains the same (i.e., $a = 0$) for all states $s' = (i, j', k, m)$ with $j' < j$. Now, suppose that the state of the system is $(i+1, j, k, m)$, then by using (14), we get the following inequality

$$V_n(i+1,j+1,k,m) - V_n(i+2,j,k+1,m-1) \leq V_n(i,j+1,k,m) - V_n(i+1,j,k+1,m-1) < 0$$

which confirms that the controller takes action $a = 0$ as the optimal action in state $s'' = (i', j, k, m)$ with $i' > i$. Using (16), it can be shown that for all states $s''' = (i, j, k', m)$ with $k' > k$, action $a = 0$ will be taken as the optimal action. An analogous proof can be provided for Classes 1 and 3 calls. This property shows that for each class of users, the optimal policy is nondecreasing switching curve. ∎

Hence, the optimal policy is described by three switching curves; each switching curve corresponds to one of the classes. For Class 1 users, we can show that for each (j, k, m), there exists a minimum value for i, such that $V_n(i+1, j, k, m) < V_n(i, j+1, k-1, m+1)$. This value is the boundary between *transferring* and *keeping* regions.

Formally, we define Φ_1, Φ_2, and Φ_3 as the switching curves for Classes 1, 2, and 3, respectively

$$\Phi_1(j,k,m) = \min_i\{(i,j,k,m) \in S : V_n(i+1,j,k,m) < V_n(i,j+1,k-1,m+1)\}$$

$$\Phi_2(i,k,m) = \min_j\{(i,j,k,m) \in S : V_n(i,j+1,k,m) < V_n(i+1,j,k+1,m-1)\}$$

$$\Phi_3(i,j,m) = \min_k\{(i,j,k,m) \in S : V_n(i,j,k+1,m) < V_n(i-1,j+1,k,m+1)\}.$$

Recursively, we can determine the sequence of n-stage value functions $\{V_1(s), V_2(s), \ldots V_n(s)\}$, and the limit of this sequence when n goes to infinity. Lippman in [14] shows that $V(s) := \lim_{n\to\infty} V_n(s)$ exists and it is the solution of the infinite horizon discounted cost problem. Besides, $V(s)$ is the unique solution to the dynamic programming equation given by (6). Note that the switching curve structure of the optimal policy holds for the infinite horizon discounted case.

The optimal policy can be derived numerically by implementing policy iteration algorithm [7]. This algorithm has two phases.

1) Policy determination phase which returns the structure of the optimal policy in $O(|S|^3)$ steps, where $|S|$ is the total number of states of the system [16]. It can be shown that for a three-class system with W wavelengths $|S| = O(W^5)$.
2) Policy improvement phase which optimizes the determined policy from the previous phase in few iterations. For instance, the policy improvement phase of a three-class 20-wavelength network with $\gamma = 0.9$ converges in five iterations.

Without loss of generality, we assume that $\mu_1 = \mu_2 = \mu_3 = 1$, unless otherwise stated. In this part, we will show the structure of the optimal policy for a three-class system which is derived from policy iteration algorithm.

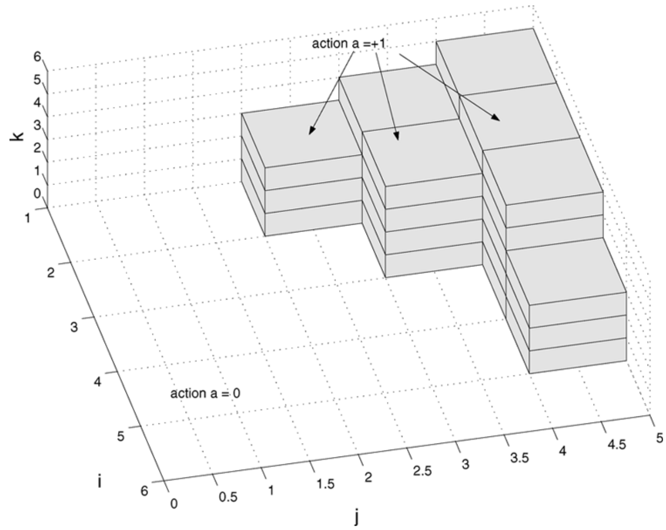

Fig. 4. Optimal policy for three classes of traffic, $W = 10$, $k = 4$, $\lambda_1 = \lambda_2 = \lambda_3 = 5$, $\alpha = 1$, $\beta = \mathbf{0.1}$, $\delta = 0.1$, and $e = D_1$.

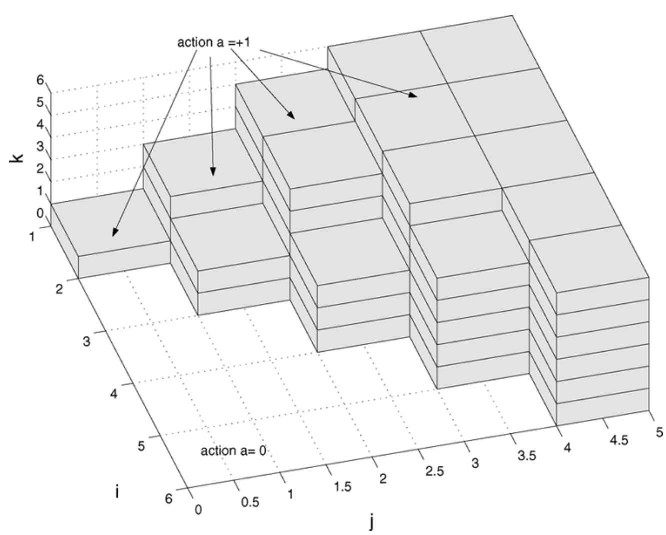

Fig. 5. Optimal policy for three classes of traffic, $W = 10$, $k = 4$, $\lambda_1 = \lambda_2 = \lambda_3 = 5$, $\alpha = 1$, $\beta = \mathbf{0.5}$, $\delta = 0.1$, and $e = D_1$.

Fig. 4 depicts the optimal policy for $W = 10$, $k = 4$, $\lambda_1 = \lambda_2 = \lambda_3 = 5$, $\alpha = 1$, $\beta = 0.1$, $\delta = 0.1$, at a Class 1 call termination time. In this figure, each cube represents action $a = 1$. Note that the policy is a monotonic three-dimensional (3-D) switching curve, dividing the state space into two subsets. The structure of the policy reflects the fact that the three classes of calls are competing for the available wavelengths. In Fig. 5, we only change the value of β from 0.1 to 0.5, and calculate the optimal policy for the same parameters, as in Fig. 4. Comparison of Figs. 4 and 5 show that by increasing β, the decision maker gives more resources to Class 2 users (i.e., more cubes are represented).

V. Multithreshold (MT) Heuristic Algorithm for WDM Ring Networks

Since implementation of an MDP-based DP allocation policy requires polynomial time, it is infeasible to use it for realistic networks. Hence, we use the properties of DP policy to devise a simpler heuristic allocation algorithm. In Section III when DP policy is used, we observe the following.

1) Optimal actions are taken based on *available wavelengths* for each class.
2) Wavelengths are *shared* as much as possible. By adjusting threshold m, they are dynamically partitioned between classes.

We consider these two facts and propose a multithreshold (MT) allocation scheme.

We introduce our heuristic algorithm for a symmetric unidirectional WDM ring network with N nodes with or without WCs. We assume that each link carries one fiber with W wavelengths. The network includes $N(N-1)$ classes of traffic streams characterized by their source-destination. We modify the definition of classes as follows. A call from Class c_h^r, $r = 1, 2, \ldots, N$, and $h = 1, 2, \ldots, N-1$ originates at node r and passes through h hops from origin to destination. In addition to the assumptions stated in Section II, we assume that classes with the same hop count (i.e., the number of hops used from origin to destination) are assigned the same weighting factor, and the same arrival rate (i.e., $\lambda_{c_h^1} = \lambda_{c_h^2} = \cdots = \lambda_{c_h^N}$, $h = 1, 2, \ldots, N-1$, where h is the hop count). Therefore, we can merge classes with the same hop count into a single class. In the following, we refer to classes with hop count h as Class C_h.

In such a network when CS policy is deployed, the system is unfair because classes of calls with smaller hop counts experience lower blocking rates than the ones with greater hop counts. To improve fairness, MT allocation will be deployed. When we use this allocation policy, similar to DP policy, while all classes *share* the wavelengths as much as possible, classes with higher hop counts are protected from the ones with lower hop counts based on *available wavelengths* on each link of the ring network.

To do so, we define $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_{N-1})$ a vector representing the thresholds of MT allocation. Threshold τ_h, $h = 1, 2, \ldots, N-1$, is associated with Class C_h traffic. In general, MT allocation scheme is capable to control blocking performance and threshold vector $\boldsymbol{\tau}$ can take any value. In particular, for solving fairness problem in a ring network we have to ensure that calls with higher hop counts are protected from the ones with lower hop counts. Therefore, we assume that $\tau_1 \geq \tau_2 \geq \cdots \geq \tau_{N-1}$. Since calls from class C_{N-1} experience the highest blocking probability in the ring, they are always accepted in the system. As a result, we set $\tau_{N-1} = 0$. Let l_n, $n = 1, 2, \ldots, N$, be the link between node n and node $n+1$ and ω_n denote the number of available wavelengths on link l_n. An arriving call of class C_h will be assigned an available wavelength if ω_n on all the links along its path is greater than or equal to τ_h.

As an example, consider a four-node unidirectional ring network shown in Fig. 6. One can note that on each link, the total number of available wavelengths W is partitioned into three sets. As shown in Fig. 6, available wavelengths on links l_1 and l_2 are greater than τ_1. Thus, all C_1, C_2, and C_3 calls which pass through l_1 and l_2 can occupy free wavelengths of these links. Since $\tau_2 < \omega_3 < \tau_1$, Class C_1 calls that use link l_3 will be blocked. On links l_4, we have $\omega_4 < \tau_2$ and as a result, only C_3 calls are allowed to occupy an idle wavelength on this link.

The next step toward MT allocation implementation is to determine the threshold vector $\boldsymbol{\tau}$, so that all classes of calls experience the same blocking probability. In [17], we defined an objective function to be minimized. It can be observed that this objective function inherits some type of pseudoconvexity properties from the value function expressed by (5). Using the properties of the objective function, we proposed and implemented a fast simulation-based algorithm to find the optimal thresholds. Due to space limitation, details of threshold determination will not be presented in this paper. Refer to [17] for complete implementation.

VI. Performance Comparison

In this section, we compare the performance of our proposed DP and heuristic allocation policies, with those of CS and CP policies. In order to implement CP policy for the two-hop network shown in Fig. 1, one can divide the total number of wavelengths W into two parts. Let M be the number of wavelengths dedicated to Class 2 and $W-M$ be the number of wavelengths reserved for Classes 1 and 3. Note that M is a constant value. Using Erlang's B formula, we can compute p_u^ℓ, the probability of having u users of Class ℓ in the system

$$p_u^\ell = \frac{\frac{1}{u!}\left(\frac{\lambda_\ell}{\mu_\ell}\right)^u}{\sum_{n=0}^{T_\ell}\frac{1}{n!}\left(\frac{\lambda_\ell}{\mu_\ell}\right)^n}$$

where T_ℓ is the total number of wavelengths reserved for Class ℓ. One can notice that $p_{T_\ell}^\ell$ is the probability that all dedicated wavelengths to Class ℓ are busy (i.e., $p_{T_\ell}^\ell$ is the blocking probability of Class ℓ calls). Using p_{W-M}^1, p_M^2, and p_{W-M}^3, we can derive the expected number of calls of each class in the system

$$N_\ell(M) = \left(\frac{\lambda_\ell}{\mu_\ell}\right)(1-p_{T_\ell}^\ell).$$

We also define M^* as

$$M^* := \arg\max_{M\in\{1,\dots W-1\}} \alpha N_1(M) + \beta N_2(M) + \delta N_3(M). \tag{19}$$

To compare the performance of DP, CP, and CS policies, we simulate the system by deploying the optimal policy implemented in Section IV to evaluate the performance metric of DP policy. For CP, the simulation result is carried out for two independent M/M/$(W-M^*)$/$(W-M^*)$ queues associated with Class 1 and Class 3 and one M/M/M^*/M^* queue related to Class 2. Finally, we simulate the system without any allocation policy to evaluate the performance of CS policy.

We first study a two-class system. Figs. 7 and 8 depict the average-time reward function $(\alpha n_1 + \beta n_2)$ versus the total offered load ρ, where $\rho = \lambda_1/\mu_1 + \lambda_2/\mu_2$. In both examples, the parameters are set as follows: $W = 10$, $\lambda_1 = \lambda_2$, $\alpha = 1$, and $6 \le \rho \le 40$. The difference between the two examples lies in the value of β. In Fig. 7, we set $\beta = 0.1$ which is significantly smaller than α. It shows that for low load, all the policies have similar performance. As the load increases, DP policy shows much better performance, in particular when compared with CS

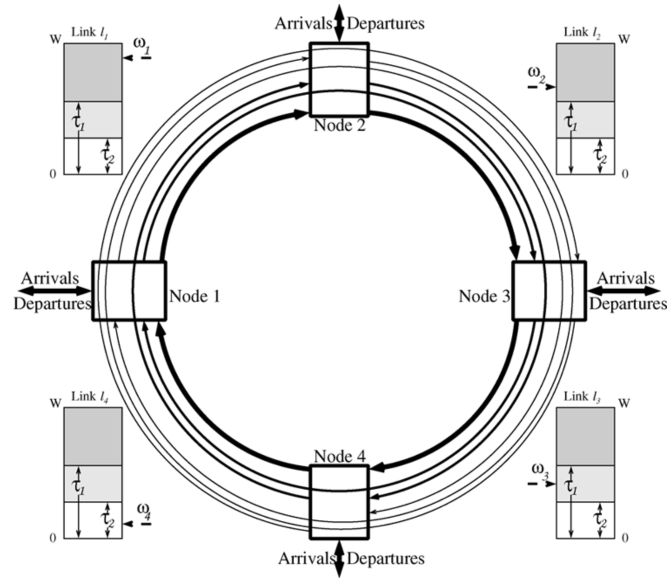

Fig. 6. Four-node unidirectional ring network.

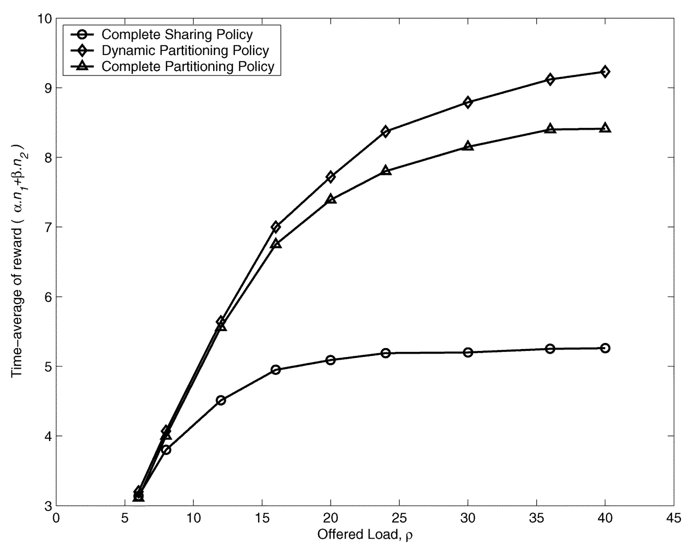

Fig. 7. Performance comparison of DP, CP, and CS policies. Performance metric is the time-average of $\alpha n_1 + \beta n_2$, $W = 10$, $\lambda_1 = \lambda_2$, $6 \le \rho \le 40$, $\alpha = 1$, and $\beta = \mathbf{0.1}$.

policy. Fig. 8 illustrates the performance of the system when $\beta = 0.5$. Comparison of Figs. 7 and 8 show that DP policy outperforms CP policy and in particular CS policy as the difference between α and β increases.

Another important performance metric is the weighted sum of blocking rates. Applying DP and CS policies, we simulate and determine this quantity for a system with $W = 10$, $\lambda_1 = \lambda_2$, $\alpha = 1$, $\beta = 0.1$, and $6 \le \rho \le 80$. To determine the relative performance improvement of DP policy when compared with CS policy, we calculate the relative performance ratio $(\mathrm{DP}_p - \mathrm{CS}_p)/\mathrm{CS}_p$, where DP_p and CS_p represent the blocking performance of DP and CS policies, respectively. This quantity is plotted versus the offered load in Fig. 9. One can see that DP policy have higher performance, up to 45% for intermediate offered load (e.g., $\rho \simeq 15$), which is a fact observed in networks and is in agreement with intuition.

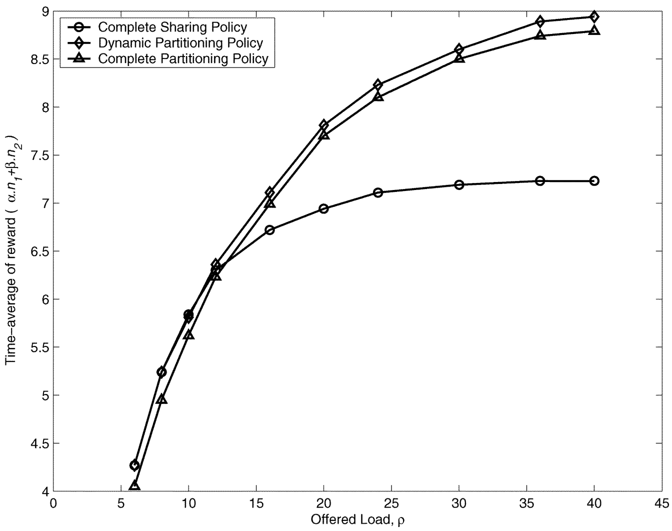

Fig. 8. Performance comparison of DP, CP, and CS policies. Performance metric is the time-average of $\alpha n_1 + \beta n_2$, $W = 10$, $\lambda_1 = \lambda_2$, $6 \leq \rho \leq 40$, $\alpha = 1$, and $\beta = \mathbf{0.5}$.

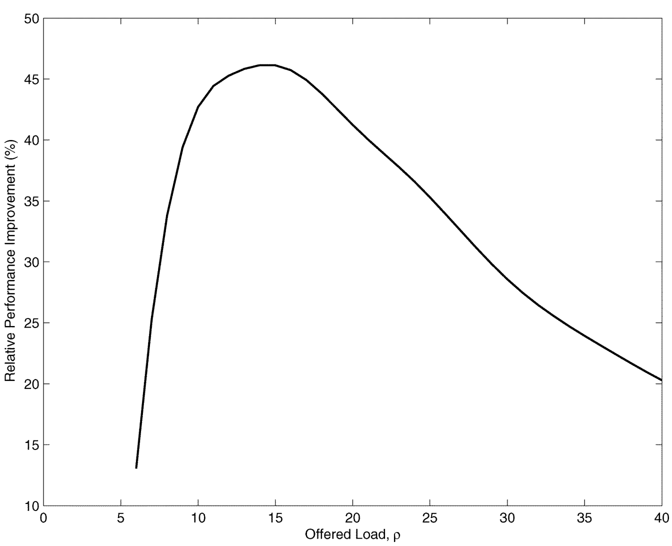

Fig. 9. Relative performance comparison between DP and CS policies. Performance metric is the weighted sum of blocked calls, $W = 10$, $\lambda_1 = \lambda_2$, $6 \leq \rho \leq 80$, $\alpha = 1$, and $\beta = 0.1$.

As mentioned in Section V, one potential problem in wavelength-routed WDM networks is fairness. We implemented a CS policy for a two-hop three-class system with a WC at node 2 and with the following parameters: $W = 10$, $\lambda_1 = \lambda_2 = \lambda_3$, $\rho = \lambda_1/\mu_1 + \lambda_2/\mu_2 + \lambda_3/\mu_3$, and $9 \leq \rho \leq 45$. The simulation result depicted in Fig. 10 shows that Class 2 calls experience more blocking than Class 1 and Class 3 calls. We apply the DP policy to improve the unfairness problem in the network. Through the simulation, we observe that for $\alpha = \delta = 1$ and $\beta = 1.95$, all the classes can experience the same blocking probabilities. Another approach to equalize the blocking probabilities is to use CP policy to partition the wavelengths among the users of different classes. In this particular example, we dedicate half of the wavelengths to Class 2 calls and the other half to Classes 1 and 3. Fig. 10 shows the blocking probabilities when we deploy CS, CP, and DP policies.

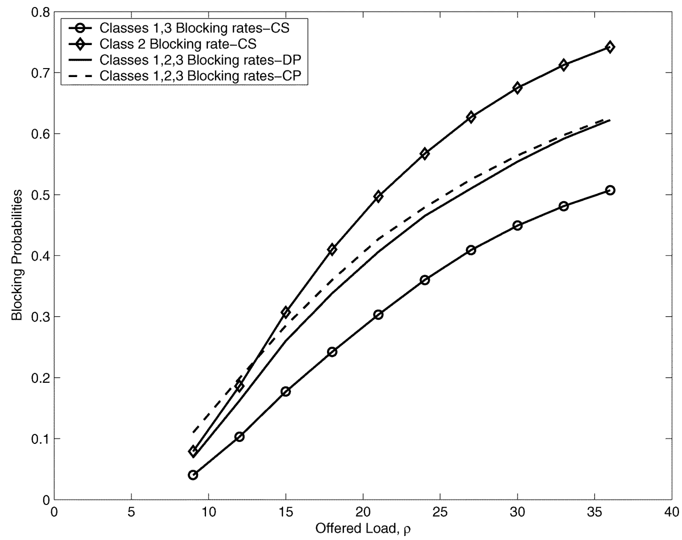

Fig. 10. Performance comparison of DP, CP, and CS policies for three classes of users. Performance metric is blocking probabilities of different classes, $W = 10, \lambda_1 = \lambda_2 = \lambda_3, 9 \leq \rho \leq 45, \alpha = \delta = 1$, and $\beta = 1.95$.

We now evaluate the performance of our heuristic wavelength allocation scheme. The simulation results are carried out for a four-node unidirectional ring network with and without WCs. The performance metric used in the numerical comparison deals with fairness and blocking probabilities.

Suppose that Class C_{high} and C_{low} calls experience the highest and lowest blocking probabilities in the network, respectively. We define fairness ratio as $f_r := B_{\text{high}}/B_{\text{low}}$, where B_x is the blocking rates of Class C_x. The closer the f_r to 1, the better the fairness is. In our simulations, the arrival rate of each class is set inversely proportional to its hop count (i.e., $\lambda_{C_1} = 2\lambda_{C_2} = \cdots = (N-1)\lambda_{C_{N-1}}$). As a result, on each link the expected wavelength request rate of each class is the same. To achieve fairness, by using MT allocation scheme each class experiences the same blocking probability.

In order to employ CP policy, one can partition the total number of available wavelengths into separate sets, each of which is dedicated to one of the origin-destination pairs. Note that nonoverlapping origin-destination pairs can share the same set of wavelengths. For example, consider the four-node ring network depicted in Fig. 6. Since class C_1 calls do not share any link, they can use the same set of wavelengths. Class C_2 calls passing through $\{l_1, l_2\}$ (i.e., c_2^1) and also C_2 calls using $\{l_3, l_4\}$ (i.e., c_2^3) can share the same set of wavelengths. Similarly, c_2^2 and c_2^4 calls can use one set of wavelengths. Four different sets are required for *overlapping* C_3 calls. As a result, the total wavelengths is partitioned into seven sets; set $\mathcal{S}_1$ is dedicated to class C_1 calls, set $\mathcal{S}_2$ is assigned to c_2^1 and c_2^3 calls and $\mathcal{S}_3$ is to classes c_2^2 and c_2^4. Four sets, $\mathcal{S}_4$–$\mathcal{S}_7$, are needed for class C_3 calls. Let w_h denote the number of wavelengths in set $\mathcal{S}_h$. Note that all sets associated with class C_h have the same number of wavelengths. Therefore, we have $w_2 = w_3$, $w_4 = w_5 = w_6 = w_7$, and $W = \sum_{h=1}^{7} w_h$. Using Erlang's B formula, we can compute the blocking probability of a call which uses wavelengths of set $\mathcal{S}_h$. We can find appropriate values for w_h in a way that all seven classes of calls experience the same blocking rate.

TABLE I
CLASS-BASED BLOCKING PROBABILITIES AND FAIRNESS RATIOS FOR A FOUR-NODE 40-WAVELENGTH RING NETWORK WITHOUT WCS. ρ_L IS THE TOTAL OFFER LOAD PER LINK

ρ_L	Policy	B_1	B_2	B_3	f_r	$\boldsymbol{\tau}$
20	CS	.000008	0.00021	0.00064	80.0	-
20	MT	0.00049	0.00045	0.00043	**1.14**	(4,1,0)
25	CS	0.00033	0.00795	0.02473	75.0	-
25	MT	0.01410	0.01561	0.01505	**1.10**	(5,1,0)
30	CS	0.00201	0.04344	0.13807	67.7	-
30	MT	0.07360	0.07063	0.07750	**1.10**	(6,1,0)
35	CS	0.00502	0.10198	0.29933	59.6	-
35	MT	0.13509	0.15496	0.17947	**1.32**	(6,1,0)
40	CS	0.00859	0.13096	0.37806	44.0	-
40	MT	0.16722	0.16821	0.19403	**1.16**	(6,1,0)

TABLE II
CLASS-BASED BLOCKING PROBABILITIES AND FAIRNESS RATIOS FOR A FOUR-NODE 40-WAVELENGTH RING NETWORK WITH WCS IN ALL NODES. ρ_L IS THE TOTAL OFFER LOAD PER LINK

ρ_L	Policy	B_1	B_2	B_3	f_r	$\boldsymbol{\tau}$
25	CS	0.00137	0.00257	0.00400	2.91	-
25	MT	0.00280	0.00156	0.00259	**1.79**	(1,0,0)
30	CS	0.01245	0.02361	0.03522	2.82	-
30	MT	0.02547	0.01609	0.02425	**1.58**	(1,0,0)
35	CS	0.03990	0.07491	0.10968	2.74	-
35	MT	0.07935	0.05349	0.07911	**1.48**	(1,0,0)
40	CS	0.07695	0.14370	0.20794	2.70	-
40	MT	0.15311	0.10807	0.15417	**1.42**	(1,0,0)
45	CS	0.11677	0.21677	0.30387	2.60	-
45	MT	0.22698	0.16142	0.23046	**1.42**	(1,0,0)

We study the fairness objective in the ring network shown in Fig. 6, when MT, CP, and CS policies are used. Table I shows the class-based blocking probabilities for a 40-wavelength ring network without WCs. As Table I reports, deploying MT policy results in a fairness ratio close to 1. As expected, when CS policy is used the fairness ratio is high, $f_r \in [44, 80]$. For low-traffic load, the blocking rates are relatively small and the competition for access to the network is low. Conversely, when the offered load is high, calls from different classes compete for utilizing resources. As a result, by increasing traffic load, the optimal τ_1 increases in order to block more C_1 calls and protect C_2 and C_3 calls.

Now, we compare MT and CP policies. This comparison is based on the number of wavelengths required to provide a certain grade-of-service (GoS) in terms of blocking probabilities. For instance, when $\rho = 30$ and CP is used, using Erlang's B formula, we computed $w_1 = 13$, $w_2 = w_3 = 8$, and $w_4 = w_5 = w_6 = w_7 = 6$. Therefore, 53 wavelengths ($W = 13 + 2 \times 8 + 4 \times 6$) is required for CP implementation, which shows 32.5% growth in terms of network resource cost. In this case, we have $f_r = 1.2$ and $B_1 = 0.084$, $B_2 = 0.070$, and $B_3 = 0.072$.

We also investigate the impact of using WCs on fairness. Table II presents the performance of MT and CS policies for the above example when all nodes are equipped by WCs. As reported in Table II, even if MT policy is not used, utilizing WCs considerably improves the fairness ratio.

Comparison of Tables I and II illustrate the following.

1) For systems without WCs, using MT allocation policy results in a very good fairness ratio (i.e., $f_r \approx 1$).
2) Deploying WCs not only decreases the overall blocking probabilities but also improves the fairness ratio. For example, the overall blocking rate of the system using MT allocation without WCs is equal to 0.0733 for $\rho = 30$, whereas this metric for the network with WCs is 0.0226.
3) In both cases, with and without WCs, MT allocation mechanism is able to improve the fairness ratio. One can see that $f_r \in [1.10, 1.16]$ and $f_r \in [1.42, 1.78]$ for the networks with and without WCs, respectively.
4) The optimal thresholds in the network with WCs are smaller than the ones in the network without WCs.

VII. CONCLUSION AND FUTURE WORKS

We have described an approach to the problem of dynamic wavelength allocation in all-optical WDM networks. First, the problem has been formulated for a two-hop path in an MDP framework and the optimal policy is obtained using the policy iteration method. It is proved that the optimal policy which maximizes the reward function is a monotonic nondecreasing switching curve. Properties of the optimal policy enabled us to propose a simple heuristic allocation algorithm to provide fairness in WDM ring networks with and without WCs. The simulation results, carried out for two-hop tandem and four-node ring networks, show that both DP and heuristic algorithms yield significant performance improvement compared with CP and CS policies.

APPENDIX A
PROOF OF PROPOSITION 1: MULTIMODULARITY OF $V_n^*(s)$

To simplify the notation, the following intermediate operators are introduced:

$$J_n^{+1}(i,j,m) = \min\{V_n^*(i+1,j,m), V_n^*(i,j+1,m+1)\} \quad (20)$$

$$J_n^{-1}(i,j,m) = \min\{V_n^*(i,j+1,m), V_n^*(i+1,j,m-1)\}. \quad (21)$$

One can interpret J_n^{+1} and J_n^{-1} as the costs just after event e and exactly before taking action a. Using these operators, the dynamic programming (7) can be rewritten as

$$\begin{aligned} V_{n+1}^*(s) = \mathbf{B}^* \cdot s^T + \frac{\gamma}{\nu}[&\lambda_1 V_n^*(A_1 s) + \lambda_2 V_n^*(A_2 s) \\ &+ (W(\mu_1+\mu_2) - n_1\mu_1 - n_2\mu_2)V_n^*(s) \\ &+ \mu_1 n_1 J_n^{+1}(s) + \mu_2 n_2 J_n^{-1}(s)]. \end{aligned} \quad (22)$$

We will prove that $V_{n+1}^* \in \mathcal{F}$ by induction in three steps. In *Step 1*, we show that $V_0^* \in \mathcal{F}$. Assuming $V_n^*(s) \in \mathcal{F}$, we establish in *Step 2*, that both J_n^{+1} and $J_n^{-1} \in \mathcal{F}$. Finally, in *Step 3*, we conclude that $V_{n+1}^* \in \mathcal{F}$.

A. Step 1

One can easily verify that $V_0^* = \mathbf{B}^* \cdot s^T$ satisfies inequalities (8)–(10) and as a result $V_0^* \in \mathcal{F}$.

B. Step 2

Assuming that $V_n^*(s) \in \mathcal{F}$, one can show that both J_n^{+1} and $J_n^{-1} \in \mathcal{F}$. Let us first prove that $J_n^{+1} \in \mathcal{F}$. In doing so, we first show that J_n^{+1} satisfies (8) and

$$J_n^{+1}(i,j,m) + J_n^{+1}(i,j+1,m+1) \leq J_n^{+1}(i+1,j,m) + J_n^{+1}(i-1,j+1,m+1) \quad (23)$$

is valid for all four possible values of $J_n^{+1}(i+1,j,m)$ and $J_n^{+1}(i-1,j+1,m+1)$ as follows.

Case 1: Assume that $V_n^*(i+2,j,m) < V_n^*(i+1,j+1,m+1)$ and $V_n^*(i,j+1,m+1) < V_n^*(i-1,j+2,m+2)$. According to J^+ operator, it can can be noticed that

$$\begin{aligned} J_n^{+1}(i+1,j,m) &= V_n^*(i+2,j,m) \\ &= \min\{V_n^*(i+2,j,m), V_n^*(i+1,j+1,m+1)\} \quad (24) \\ J_n^{+1}(i-1,j+1,m+1) &= V_n^*(i,j+1,m+1) \\ &= \min\{V_n^*(i,j+1,m+1), V_n^*(i-1,j+2,m+2)\}. \quad (25) \end{aligned}$$

Using (20), we can deduce

$$J_n^{+1}(i,j,m) + J_n^{+1}(i,j+1,m+1) \leq V_n^*(i+1,j,m) + V_n^*(i+1,j+1,m+1). \quad (26)$$

On the other hand, the multimodularity of $V_n^*(s)$ implies that

$$V_n^*(i+1,j,m) + V_n^*(i+1,j+1,m+1) \leq V_n^*(i+2,j,m) + V_n^*(i,j+1,m+1). \quad (27)$$

According to the assumption of this case, the right-hand side of inequality (27) is equal to

$$V_n^*(i+2,j,m) + V_n^*(i,j+1,m+1) = J_n^{+1}(i+1,j,m) + J_n^{+1}(i-1,j+1,m+1). \quad (28)$$

By combining (26)–(28), we get

$$J_n^{+1}(i,j,m) + J_n^{+1}(i,j+1,m+1) \leq J_n^{+1}(i+1,j,m) + J_n^{+1}(i-1,j+1,m+1). \quad (29)$$

Case 2: The assumption for this case is

$$\begin{aligned} J_n^{+1}(i+1,j,m) &= V_n^*(i+2,j,m) \\ J_n^{+1}(i-1,j+1,m+1) &= V_n^*(i-1,j-2,m+2). \quad (30) \end{aligned}$$

Using (20), we derive the following inequality:

$$J_n^{+1}(i,j,m) + J_n^{+1}(i,j+1,m+1) \leq V_n^*(i+1,j,m) + V_n^*(i,j+2,m+2). \quad (31)$$

Similar to the previous case, the multimodularity of the value function implies that

$$\begin{aligned} V_n^*(i+1,j,m) + V_n^*(i+1,j+1,m+1) &\leq V_n^*(i+2,j,m) + V_n^*(i,j+1,m+1) \quad (32) \\ V_n^*(i,j+1,m+1) + V_n^*(i,j+2,m+2) &\leq V_n^*(i+1,j+1,m+1) + V_n^*(i-1,j+2,m+2). \quad (33) \end{aligned}$$

After summing (32) and (33), we obtain

$$\begin{aligned} V_n^*(i+1,j,m) + V_n^*(i,j+2,m+2) &\leq V_n^*(i+2,j,m) + V_n^*(i-1,j+2,m+2) \quad (34) \\ V_n^*(i+2,j,m) + V_n^*(i-1,j+2,m+2) &= J_n^{+1}(i+1,j,m) + J_n^{+1}(i-1,j+1,m+1). \quad (35) \end{aligned}$$

Combining (31), (34), and (35) implies (23).

Case 3: The assumption is

$$\begin{aligned} J_n^{+1}(i+1,j,m) &= V_n^*(i+1,j+1,m+1) \\ J_n^{+1}(i-1,j+1,m+1) &= V_n^*(i-1,j+2,m+2). \quad (36) \end{aligned}$$

Using (20), we can show that

$$J_n^{+1}(i,j,m) + J_n^{+1}(i,j+1,m+1) \leq V_n^*(i,j+1,m+1) + V_n^*(i,j+2,m+2). \quad (37)$$

From multimodularity of $V_n^*(s)$ and the assumption of this case, it can be concluded that

$$\begin{aligned} V_n^*(i,j+1,m+1) + V_n^*(i,j+2,m+2) &\leq V_n^*(i+1,j+1,m+1) + V_n^*(i-1,j+2,m+2) \\ V_n^*(i+1,j+1,m+1) + V_n^*(i-1,j+2,m+2) &= J_n^{+1}(i+1,j,m) + J_n^{+1}(i-1,j+1,m+1). \quad (38) \end{aligned}$$

By combining (37) and (38), it can be shown that inequality (23) is satisfied.

Case 4: The assumption for this case is

$$\begin{aligned} J_n^{+1}(i+1,j,m) &= V_n^*(i+1,j+1,m+1) \\ J_n^{+1}(i-1,j+1,m+1) &= V_n^*(i,j+1,m+1). \quad (39) \end{aligned}$$

Using (20) and the assumption of Case 4, we can obtain

$$\begin{aligned} J_n^{+1}(i,j,m) + J_n^{+1}(i,j+1,m+1) &\leq V_n^*(i,j+1,m+1) + V_n^*(i+1,j+1,m+1) \\ V_n^*(i,j+1,m+1) + V_n^*(i+1,j+1,m+1) &= J_n^{+1}(i+1,j,m) + J_n^{+1}(i-1,j+1,m+1). \quad (40) \end{aligned}$$

So far, we have proved that for all possible cases, inequality (23) is satisfied. An analogous proof can be applied (by induction with respect to n) to get the following inequalities:

$$J_n^{+1}(i,j,m) + J_n^{+1}(i-1,j,m+1) \le J_n^{+1}(i,j-1,m) + J_n^{+1}(i-1,j+1,m+1)$$
$$J_n^{+1}(i,j,m) + J_n^{+1}(i-1,j+1,m) \le J_n^{+1}(i,j,m-1) + J_n^{+1}(i-1,j+1,m+1). \quad (41)$$

Hence, we established that $J_n^{+1}(s)$ and similarly $J_n^{-1}(s)$ are multimodular.

C. Step 3

Since $V_{n+1}(s)$ is a linear combination of $J_n^{+1}(s)$, $J_n^{-1}(s)$, and $V_n(s)$ with positive coefficients, then $V_{n+1}(s)$ is also multimodular.

Appendix B
Proof of Proposition 2: Multimodularity of $V_n(s)$

Let us generalize the intermediate operators as

$$J_n^{+1}(i,j,k,m) = \min\{V_n(i+1,j,k+1,m), V_n(i,j+1,k,m+1)\} \quad (42)$$
$$J_n^{-1}(i,j,k,m) = \min\{V_n(i,j+1,k,m), V_n(i+1,j,k+1,m-1)\}. \quad (43)$$

By substituting these operators in the dynamic programming equation given in (6), we can see that

$$V_{n+1}(s) = C(s) + \frac{\gamma}{\nu}\Big[\lambda_1 V_n(A_1 s) + \lambda_2 V_n(A_2 s) + \lambda_3 V_n(A_3 s) + \mu_1 n_1\, J_n^{+1}(s) + \mu_2 n_2\, J_n^{-1}(s) + \mu_3 n_3\, J_n^{+1}(s) + F\, V_n(s)\Big]. \quad (44)$$

Similar to the proof of *Proposition 1*, it can be shown by induction that $V_{n+1} \in \mathcal{F}$. We can easily show that $V_0 = C(s) \in \mathcal{F}$. Assume $V_n(s) \in \mathcal{F}$ and one can establish that both $J_n^{+1}, J_n^{-1} \in \mathcal{F}$. As a result, it can be shown that $V_{n+1} \in \mathcal{F}$. Let us prove that J_n^{+1} satisfies (14)–(17). We first show that

$$J_n^{+1}(i,j,k,m) + J_n^{+1}(i,j+1,k-1,m+1) \le J_n^{+1}(i+1,j,k,m) + J_n^{+1}(i-1,j+1,k-1,m+1) \quad (45)$$

is valid for all four possible values of $J_n^{+1}(i+1,j,k,m)$ and $J_n^{+1}(i-1,j+1,k-1,m+1)$. Let us provide more details for one of these four cases.

Case 1: Assume that: $V_n(i+2,j,k+1,m) < V_n(i+1,j+1,k,m+1)$ and $V_n(i,j+1,k,m+1) < V_n(i-1,j+2,m+2)$, then

$$J_n^{+1}(i+1,j,k,m) = V_n(i+2,j,k+1,m)$$
$$J_n^{+1}(i-1,j+1,k-1,m+1) = V_n(i,j+1,k,m+1).$$

Using (42), we can derive

$$J_n^{+1}(i,j,k,m) + J_n^{+1}(i,j+1,k-1,m+1) \le V_n(i+1,j,k+1,m) + V_n(i+1,j+1,k,m+1). \quad (46)$$

$V_n(s) \in \mathcal{F}$ implies that

$$V_n(i+1,j,k+1,m) + V_n(i+1,j+1,k,m+1) \le V_n(i+2,j,k+1,m) + V_n(i,j+1,k,m+1). \quad (47)$$

According to the assumption of this case, the right-hand side of inequality (47) is equal to

$$V_n(i+2,j,k+1,m) + V_n(i,j+1,k,m+1) = J_n^{+1}(i+1,j,k,m) + J_n^{+1}(i-1,j+1,k-1,m+1). \quad (48)$$

By combining (46)–(48), it can be shown that

$$J_n^{+1}(i,j,k,m) + J_n^{+1}(i,j+1,k-1,m+1) \le J_n^{+1}(i+1,j,k,m) + J_n^{+1}(i-1,j+1,k-1,m+1). \quad (49)$$

The proof of *Cases 2, 3*, and *4* are similar to *Cases 2, 3,* and *4* of *Proposition 1*, respectively.

Correspondingly, it can be shown that J_n^{+1} satisfies (15)–(17) as well. To this end, we established that $J_n^{+1}(s) \in \mathcal{F}$. In much the same way, we can show that $J_n^{-1}(s) \in \mathcal{F}$. Since $V_{n+1}(s)$ is a linear combination of $J_n^{+1}(s)$, $J_n^{-1}(s)$, and $V_n(s)$ with positive coefficients, then $V_{n+1}(s)$ is also multimodular.

References

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

[2] L. Pezoulas, M. J. Fransisco, I. Lambadaris, and C. Huang, "Performance analysis of a backward reservation protocol in networks with sparse wavelength conversion," in *Proc. ICC3*, vol. 2, May 2003, pp. 1468–1473.

[3] R. A. Barry and P. A. Humblet, "Models of blocking probability in all-optical networks with and without wavelength changers," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 5, pp. 858–867, Jun. 1996.

[4] A. Birman, "Computing approximate blocking probabilites for a class of all-optical networks," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 5, pp. 852–857, Jun. 1996.

[5] S. Subramaniam, M. Azizoglu, and A. K. Somani, "All-optical networks with sparse wavelength conversion," *IEEE/ACM Trans. Netw.*, vol. 4, no. 4, pp. 544–557, Aug. 1996.

[6] Y. Zhu, G. N. Rouskas, and H. G. Perros, "A path decomposition approach for computing blocking probabilities in wavelength-routing networks," *IEEE/ACM Trans. Netw.*, vol. 8, no. 6, pp. 747–762, Dec. 2000.

[7] M. L. Puterman, *Markov Decision Processes*. New York: Wiley, 1994.

[8] I. Lambadaris, P. Narayan, and I. Viniotis, "Optimal service allocation amoung two heterogeneous traffic types with no queueing," in *Proc. IEEE 26th CDC*, Los Angeles, CA, Dec. 1987, pp. 209–213.

[9] K. Mosharaf, J. Talim, and I. Lambadaris, "A Markov decision process model for dynamic wavelength allocation in WDM networks," in *Proc. IEEE GLOBECOM*, San Francisco, CA, Dec. 2003, pp. 2590–2594.

[10] E. Altman, B. Gaujal, and A. Hordijk, "Multimodularity, convexity and optimization properties," *Math. Oper. Res.*, vol. 25, pp. 324–347, 2000.

[11] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification, and Adaptive Control*. Englewood Cliffs, NJ: Prentice-Hall., 1986.

[12] K. W. Ross and D. H. K. Tsang, "The stochastic knapsack problem," *IEEE Trans. Commun.*, vol. 37, no. 7, pp. 740–747, Jul. 1989.

[13] D. P. Bertsekas, *Dynamic Programming: Deterministic and Stochastic Models*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[14] S. A. Lippman, "Semi-Markov decision processes with unbounded rewards," *Manage. Sci.*, vol. 13, pp. 717–731, 1973.

[15] A. Birman and A. Kershenbaum, "Routing and wavelength assignment methods in single-hop all-optical networks with blocking," in *Proc. IEEE INFOCOM*, Apr. 1995, pp. 431–438.

[16] M. L. Littman, L. D. Thomas, and L. Kaelbling, "On the complexity of solving Markov decision problems," in *Proc. 11th Conf. on Uncertainty Artif. Intell.*, Aug. 1995, pp. 394–402.

[17] K. Mosharaf, I. Lambadaris, and J. Talim, "Multithreshold wavelength allocation policy for fairness control in WDM ring networks," Carleton Univ., Ottawa, ON, Canada, Tech. Rep., SCE-04–13, 2004.

Kayvan Mosharaf (S'03) was born in Isfahan, Iran, in 1971. He received the B.Sc. degree from the Isfahan University of Technology, Isfahan, Iran, and the M.Sc. degree from the University of Tehran, Tehran, Iran, in 1994 and 1997, respectively, both in electrical engineering. He is currently working towards the Ph.D. degree in the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada.

His research interests include optical WDM networks, modeling and performance evaluation, traffic engineering, and Markov decision processes.

Jérôme Talim received the Ph.D. degree in computer science from INRIA, University of Nice, Sophia Antipolis, France, in 1999.

He held a Postdoctoral position at Cambridge University in 1999, where he investigated distributed system paradigms. He also worked at Vrije Universiteit, Amsterdam, in 2000 on call center performance evaluation. He is an Assistant Professor at Carleton University, Ottawa, ON, Canada, since 2001. His research interests include stochastic processes, computer networks, and distributed system.

Ioannis Lambadaris (M'02) was born in Thessaloniki, Greece. He received the Diploma degree in electrical engineering from the Polytechnic School, Aristotle University of Thessaloniki, Thessaloniki, Greece, in 1984, the M.Sc. degree in engineering from Brown University, Providence, RI, in 1985, and the Ph.D. degree in electrical engineering from the Department of Electrical Engineering, Systems Research Center (SRC), Institute for Systems Research (ISR), University of Maryland, College Park, in 1991.

After finishing his graduate education, he was a Research Associate at Concordia University, Montreal, QC, Canada, from 1991 to 1992. From September 1992 and July 1997, he was an Assistant Professor with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada. Currently, he is an Associate Professor in the same department. His interests lie in the area of applied stochastic processes and their application for modeling and performance analysis of computer communication networks. His current research concentrates on quality-of-service (QoS) control for IP and evolving optical networks architectures. His research is done in close collaboration with his students and colleagues in the Broadband Networks Laboratory (www.sce.carleton.ca/bbnlab/bnlhome.shtml).

Dr. Lambadaris received a Fellowship from the National Fellowship Foundation of Greece (1980–1984) during his undergraduate studies. He also received the Technical Chamber of Greece Award (ranked first in graduating class). He was a recipient of a Fulbright Fellowship (1984–1985) for graduate studies in the U.S. While at Carleton University, he received the Premiers Research Excellence Award, and the Carleton University Research Excellence Award (2000–2001), for his research achievements in the area of modeling and performance analysis of computer networks.