Efficient Routing in All-Optical Networks

Prabhakar Raghavan* Eli Upfal†

Abstract

Communication in all-optical networks requires novel routing paradigms. The high bandwidth of the optic fiber is utilized through *wavelength-division multiplexing:* a single physical optical link can carry several logical signals, provided that they are transmitted on different wavelengths. We study the problem of routing a set of requests (each of which is a pair of nodes to be connected by a path) on sparse networks using a limited number of wavelengths, ensuring that different paths using the same wavelength never use the same physical link.

The constraints on the selection of paths and wavelengths depend on the type of photonic switches used in the network. We present efficient routing techniques for the two types of photonic switches that dominate current research in all-optical networks. Our results establish a connection between the expansion of a network and the number of wavelengths required for routing on it.

*IBM T.J. Watson Research Center, Yorktown. This work was supported in part by grant MDA 972-92-C-0075 from ARPA.

†The Weizmann Institute, Israel, and IBM Almaden Research Center, California. Work at the Weizmann Institute supported in part by the Norman D. Cohen Professorial Chair of Computer Science.

0

1. Introduction

The subject of this paper is the design of algorithms for an emerging generation of networks known as *all-optical networks* [12, 15, 16, 17, 23, 27]. These networks promise data transmission rates several orders of magnitudes higher than current networks. The key to high speeds in these networks is to maintain the signal in optical form, thereby avoiding the prohibitive overhead of conversion to and from the electrical form. (Traditional networks use the electrical form to switch signals along routes, and to restore signal strength. Signals can be modulated electronically at a maximum bit rate of the order of 10 Gbps, while the optical fiber bandwidth is about 10 THz [28]). The high bandwidth of the optic fiber is utilized through *wavelength-division multiplexing*: two signals connecting different source-destination pairs may share a link, provided they are transmitted on carriers having different frequencies (i.e., wavelengths) of light.

The major applications for such networks are in video conferencing, scientific visualization and real-time medical imaging, high-speed supercomputing and distributed computing [16, 27, 30]. The books by Green [27] and by McAulay [21] give a comprehensive overview of the physical theory and applications of this technology. What are the algorithmic issues concerning routing in such networks? The answer depends on the exact physical mechanism used to switch signals along routes through the network. Two types of photonic switching devices dominate current research. Both appear likely to be used in networks of the future, depending on the cost, scale and applications involved. Our goal is to devise algorithms for both technologies, and to understand the difficulties in devising algorithms for each.

1.1. Two Models

We model the network as an undirected graph. A *request* consists of a pair of nodes, an *instance* consists of a set of requests. For the bulk of this paper, the requests in the instance are specified all at once; however, in Section 4 we indicate why some of our algorithms work in the *dynamic* setting in which requests appear and disappear.

A *solution* consists of a setting for the switches in the network, and an assignment of a *wavelength* to each request. A solution has to guarantee that there is a path between the pair of nodes in each request, and that no edge will carry two different signals on the same wavelength. For our algorithmic purposes, a wavelength will be an integer in the interval $[1, w]$ for some positive integer w. Generally, we wish to minimize the quantity w, since the cost of switching and amplification devices depends on the number of wavelengths they handle. We will also consider variants in which the number of available wavelengths is fixed, necessitating communication in a sequence of rounds each of which routes some of the requests.

Clearly, if an intermediate node could change the wavelength on which a signal is transmitted, routing an instance using the minimum number of wavelength would be equivalent to the problem of integer multicommodity flow. Unfortunately, current (or any foreseeable) technologies cannot implement such a photonic switch. This necessitates the study of novel routing techniques that can be implemented with less powerful but feasible switches.

The first type of switch we consider is the *generalized* switch based on accousto-optic filters [14]. Here, signals for different requests may travel on an edge into a node v (on different wavelengths, of course) and then exit v along different edges. Thus, the photonic switch can differentiate between several wavelengths coming in along an edge and direct each of them to a different output of the switch. The only constraint on the solution is that no two paths sharing any edge have the same wavelength.

The *elementary* switch cannot direct different frequencies coming into a node along different outgoing edges [7, 19] [1]. It is considerably easier (than the generalized switch) to build, is faster to switch, and it can currently carry a larger number of different wavelengths. It is simplest to think of each node as partitioning its incident edges into subsets; within a subset, all signals flowing on any edge flood all other edges in that subset, but the signals on different subsets remain insulated from each other. Thus a signal may flood edges not on its planned path, blocking the use of its wavelength on these edges as well. Formally, a configuration of the network is a partition of its *edges* into subsets, with the following constraints: (i) each request is assigned to one subset, and there is a path in that subset connecting the endpoints of that request; (ii) no more than w requests are assigned to any subset. Thus each subset corresponds to a region flooded by any signal routed through that region, and constraint (ii) above ensures that the number of wavelengths is within the permissible bounds.

The actual process of setting up switches and routes (as well as wavelength assignment) in networks employing either type of switch is done using an electronic backbone control network. The reader may wonder at the use of a relatively slow (electronic) network to set up these high-speed connections. In fact, the major applications for such networks require connections that last for relatively long peri-

[1] The terms "elementary" and "generalized" are borrowed from [1]; a large set of names prevails in the communications and physics literature.

ods once set up; thus the initial overhead is acceptable as long as sustained throughput at high data rates is subsequently available. Consequently, in this paper we will generally view the algorithmic process as central rather than distributed. However, where appropriate, we will point out how our algorithms can be implemented with local control and/or can deal with dynamic request sequences (even though this does not appear to be a primary focus of current algorithmic research in this area). Eventually we imagine that centralized algorithms will be implemented in an approximate, distributed form (much as the internet implements an approximation to a shortest path algorithm, in a distributed fashion). The following facts are easy to verify, and may strengthen the reader's intuition about the two types of switches. They tell us that generalized switches are at least as powerful as elementary switches, and may be far more powerful.

Fact 1: Given a routing of an instance on a network with elementary switches, we can route this instance on a network with the same topology but using generalized switches using the same number of wavelengths.

Fact 2: There is an n-node network and a permutation that requires $\Omega(n)$ wavelengths using elementary switches, but $O(1)$ wavelengths suffice using generalized switches.

1.2. Related Previous Work

Barry and Humblet [9, 8], Pieris and Sasaki [24, 25] and Pankaj [22] have given lower bounds on the number of wavelengths required for permutation routing in any network, independent of topology, with a given number of generalized switches. Pankaj [22] went on to consider lower and upper bounds for a few specific networks; for example, he gives an upper bound of $O(log^2 n)$ wavelengths for permutation routing on the hypercube. In addition, a number of papers in the communications literature [7, 15, 28] have formulated the routing problem for both elementary and generalized switches as combinatorial optimization problems. Aggarwal, Bar-Noy, Coppersmith, Ramaswami, Schieber and Sudan [1] gave bounds on the number of switches required without taking into account the network topology, as a function of the number of wavelengths available. In addition, they proved results on routing in non-blocking permutation networks using generalized switches. Other related work includes algorithms for integer multicommodity flow [18, 20, 26] and online call assignment [3, 4, 5, 6]. However, there is substantial evidence (see Theorem 17 of [1], for instance) that our problem is considerably harder than integer multicommodity flow, due to the severe additional restrictions imposed by the path-coloring requirements.

1.3. Our Results

In this paper we address the main problem left open by the work in [1]: that of obtaining provably good routing algorithms for arbitrary networks. We focus on sparse, bounded degree networks. We measure the quality of routing algorithms by the number of wavelengths and rounds used for routing an arbitrary *k-relation*: an instance in which each node is a source and destination of no more than k messages.

We show that the number of wavelengths required in the worst-case is closely related to the *edge-expansion* of the graph: the minimum, over all subsets S of vertices, $|S| \leq n/2$, of the ratio of the number of edges leaving S to the size of S. The lower bound is given in Section 2.1, and the upper bound for generalized switches is in Section 2.2. We also show that our algorithm has a natural variant that works for elementary switches, but with an increase in the number of rounds; to our knowledge, these are the first provably good results for elementary switches.

We then turn to a number of specific topologies including trees, cycles and meshes of various dimensions, exploiting the properties of these graphs to obtain better upper bounds in both models. Trees, cycles and meshes arise often in practice. Some researchers [30] have proposed embedding a logical mesh in a physical network (a one-time, possibly computationally intensive task) and then routinely performing the routing on the logical mesh to exploit its regularity. Meshes of various dimensions are also used in parallel computers, where all-optical networks have been proposed [16]. Finally, in Section 4, we indicate how our algorithms for arbitrary graphs work without modification in a dynamic setting in which requests appear and disappear over time.

2. Routing With Generalized Switches

2.1. Lower Bound

We give a lower bound on the number of wavelengths for routing in an arbitrary graph G, in terms of its *edge-expansion* $\beta(G)$, and the number of rounds. Note that this lower bound applies *a fortiori* to networks with elementary switches.

Theorem 1: For every $\beta \leq 1$ and k, there is a graph G with edge-expansion $\beta(G) = \beta$, and a k-relation $\mathcal{R}$, such that routing $\mathcal{R}$ on a network with topology G and generalized switches requires $\Omega(k/(\beta^2))$ wavelengths.

Proof: The proof uses the mesh-like graph in Figure 1, adapted from [1]. However, in

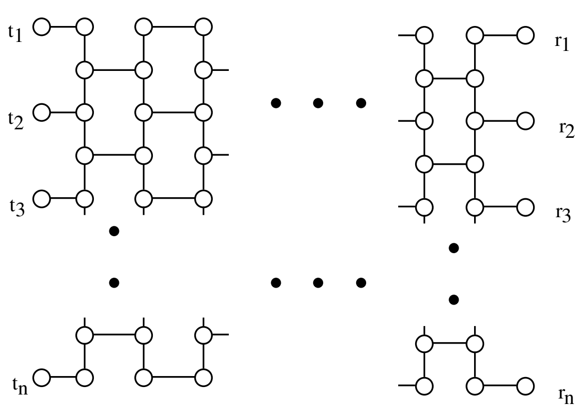

Figure 1: *The graph for the lower bound.*

our case, the graph has $B = \lceil 1/\beta \rceil$ rows and columns. Further, we augment it by connecting each of the r_i and t_i by a single edge to a distinct constant-degree expander having B nodes. Each node in the expander connected to t_i requests k connections to the corresponding node in the expander connected to r_{B-i}. It is easy to verify now that the edge-expansion of this graph is β, and that the route for each of the $O(kB^2)$ requests shares an edge with the route for every other request. □

2.2. Arbitrary Bounded Degree Graphs

We now give an algorithm for routing with generalized switches in arbitrary bounded-degree graphs, and analyze its performance on an arbitrary *k-relation.* We do so in the general setting in which the number of available wavelengths is traded off for the number of rounds of routing.

Our solution for general graphs is based on a random walk technique. A *random walk* on the undirected graph $G = (V, E)$ is a Markov chain $\{X_t\} \subseteq V$ associated to a particle that moves from vertex to vertex according to the following rule: the probability of a transition from vertex i, of degree d_i, to vertex j is $1/d_i$ if $\{i, j\} \in E$, and 0 otherwise. (For technical reasons we assume that the particle actually remains where it is with probability 1/2 at each step, and moves with probability 1/2 only. This technicality is ignored for the remainder of the paper.) Let Q denote the transition probability matrix of this random walk on G, and assume that the absolute value of all the eigenvalues of Q other than the largest one is bounded by λ. (All eigenvalues of Q are real.) The stationary distribution of the random walk, denoted π, (or $\pi(G)$) is given by $\pi_v = d_v/(2|E|)$. A trajectory W of length τ is a sequence of vertices $[w_0, w_1, \ldots, w_\tau]$ such that $(w_t, w_{t+1}) \in E$. The Markov chain $\{X_t\}$ induces a probability distribution on trajectories in the obvious way.

Algorithm:

Input: An n-node bounded degree network $G = (V, E)$; a k-relation $\mathcal{R} = \{(a_1, b_1), ..., (a_\ell, b_\ell)\}$.

Output: Routing of the relation $\mathcal{R}$ on the network G.

1. Let $L = -3\frac{\log kn}{\log \lambda}$.

2. For each $(a_i, b_i) \in \mathcal{R}$

   (a) Choose a node $r_i \in V$ uniformly at random.

   (b) Choose a trajectory W_i' (resp. W_i'') of length L from a_i to r_i (resp. b_i to r_i) according to the distribution on trajectories, conditioned on the endpoints being a_i and r_i (resp. b_i and r_i).

   (c) Connect a_i to b_i by the path P_i defined by W_i' followed by W_i''.

   (d) Use a wavelength that is not used by any other transmission sharing an edge with the path P_i.

Analysis of the Algorithm:

Theorem 2: With high probability the algorithm uses $O(kL^2)$ wavelengths.

Sketch of the proof: Let $Q_{v,w}^{(\tau)}$ denote the probability that a random walk is at vertex w at step τ given that it started at v. It is known that

$$Q_{v,w}^{(\tau)} = \pi(w) + O(\lambda^\tau \sqrt{\pi(w)/\pi(v)}). \qquad (1)$$

Following [13], our analysis of the algorithm relies heavily on the fact that the trajectories W_i' (resp. W_i'') have the same distribution (up to negligible factors) as independent random random walks of length L from a_i (resp. b_i). The difference is that we pick the endpoint of the trajectory using π (the stationary probability) instead of $Q_{a_i,\cdot}^{(L)}$. However, since $L = -3\frac{\log kn}{\log \lambda}$,

$$|Q_{v,w}^{(\tau)} - \pi(w)| = O((kn)^{-2}),$$

for all v, w, and the difference is negligable. (See [13] for a complete proof.)

Since the paths are generated by random walks, and the starting points of the paths are chosen by the stationary distribution (up to a constant factor), the expected number of paths traversing an edge is bounded by $O(kL)$, and the expected number of different paths sharing an edge with the path of a given message is bounded by $O(kL^2/T)$. Thus, using the Chernoff bound we prove that with high probability there is no path that shared edges with more than $O(kL^2)$ other paths in any iteration. $\square$

To relate the performance of the algorithm to the lower bound in Theorem 1, we use the relation between the edge-expansion β of a bounded degree graph and the value of its second largest eigenvalue in absolute value λ [2, 29]:

$$1 - O(\beta^{-1}) \leq \lambda \leq 1 - O(\beta^{-2}).$$

Applying this relation, for graphs for which the left hand side is equality

$$kL^2 = k(\log n)^2(\log \lambda)^2 \leq (k\beta^2)(\log n)^2,$$

so that the number of wavelengths used by the algorithm is within a polylog factor of the optimal number. For other graphs the number of wavelengths used is at most the square of the optimum. In the next two subsections we present better solutions for some of these graphs.

2.3. Trees and Rings

Two important topologies in practice are trees and rings. We study these graphs here, as well as the related *tree of rings* found widely in practice.

Theorem 3: Given any tree, there is a deterministic algorithm that routes any set of requests on that tree using no more than $(3/2)w_{opt}$ wavelengths, where w_{opt} is the minimum possible number of wavelengths for that set of requests.

Comments: Note that our algorithm is provably good on trees of arbitrary degree and for every set of requests, whether or not it a k-relation for fixed k. For trees (as well as rings), w_{opt} may be linear in n; in both cases, our results can be "slimmed down" to suit a given bound on w by randomly partitioning the requests into (an easily computed number of) rounds. This is nearly optimal; details are omitted.

Proof: We only give a very brief outline here. The idea is to reduce the problem of requests on a tree to a derived set of requests on the star graph (a tree consisting of one central node to which each of the remaining nodes is connected by a "spoke").

Once we are down to the star graph, we note that assigning wavelengths to the requests corresponds to edge-coloring a multigraph, each node of which corresponds to a spoke in the star. This can be done with at most $3d/2$ colors (and no better in general) [11], where d is the maximum node degree in the multigraph. Finally, it turns out that in the reduction to multigraph edge-coloring, d remains a lower bound on w_{opt}. □

For rings, we invoke slightly different techniques to obtain:

Theorem 4: There is a polynomial-time algorithm that, for any set of requests on a ring, uses no more than $2w_{opt}$ wavelengths, where w_{opt} is the minimum possible number of wavelengths for that set of requests.

The *tree of rings* is a network constructed as follows: start from a tree, and replace each node of the tree by a cycle. Each edge corresponds to the corresponding cycles sharing a node. The tree of rings is a common interconnection pattern in local-area networks: there is a main ring, with several sub-rings dangling from it, sub-subrings from the sub-rings, and so on. By combining the algorithms of Theorems 3 and 4, we can give an algorithm that is within a factor of 3 of the optimal number of wavelengths on any tree of rings.

2.4. d-Dimensional Meshes

Let M_d denote an n node d dimension mesh, i.e. the set of vertices of M_d is:

$$\{\bar{a} = (a_1, ..., a_d) \mid 1 \leq a_i \leq n^{1/d}, \quad i = 1, ..., d\},$$

and two vertices are connected by an edge iff their Hamming distance is one.

Theorem 5: There is a probabilistic algorithm that routes any k-relation on M_d and with high probability uses $O(kdn^{1/d})$ wavelengths.

Proof: Let $L(\bar{a}, i)$ denote the chain of vertices in M_d with all their coordinates other than the ith coordinate equal to the coordinates of $\bar{a}$, i.e.,

$$L(\bar{a}, i) = \{(a_1, ..., a_{i-1}, x, a_{i+1}, ..., a_n) \mid 1 \leq x \leq n^{1/d}\}.$$

We route a message from its origin to its destination in $2d-1$ segments. The first $d-1$ segments are random, the last d take the message to its destination. The first segment starts at the origin of the message. For $i = 1, ..., d-1$, if segment i, starts at node $\bar{a}$, then segment i connects node $\bar{a}$ to a random node $(a_1, ..., a_{i-1}, r, a_{i+1}, ..., a_d)$, on $L(\bar{a}, i)$, where r is chosen randomly and uniformly in the range $[1, ..., n^{1/d}]$.

Let $\bar{a}$ be the start point of segment $2d-i$, $1 = 1, ..., d$, and assume that the $d-i+1$ coordinate in the destination of the message is b_i, then segment i connects node $\bar{a}$ to node $(a_1, ..., a_{i-1}, b_i, a_{i+1}, ..., a_d)$ on $L(\bar{a}, i)$.

For the analysis of the number of wavelengths used by the algorithm we assume that all messages that pass an edge in $L(\bar{a}, i)$ must use different wavelengths (although they may use different edges of this chain).

To simplify the presentation we transform the analysis to an acyclic directed network H that consists of two degree $n^{1/d}$ butterflies connected back to back. The network H has $2d-1$ levels and $n^{(d-1)/d}$ nodes per level. Nodes at levels i and $2d-i$ correspond to the chains $L(\bar{x}, i)$ in M_n. A node in level i is connected to the $n^{1/d}$ nodes in level $i+1$ lying on chains that can be reached by changing the corresponding coordinate. Clearly the number of wavelengths needed in the algorithm is bounded by the maximum number of messages that share vertices with a given message when routed on the network H.

The routing starts with up to $kn^{(d-1)/d}$ messages per input of H. The probability that when routing on H any message shares vertices with more than $12kdn^{1/d}$ other messages is bounded by

$$kne^{-2\frac{knd}{n^{\frac{d-1}{d}}}} \leq \frac{1}{n}.$$

□

3. Routing With Elementary Switches

3.1. Arbitrary Bounded Degree Graphs

Let G be an n node network with generalized switches. Let G' be a network with the same topology as G but with elementary switches. Clearly, given a routing of a k-relation on G in one round using w wavelengths, one can route the k-relation on G' in w rounds using one wavelength per round, simply by routing the set of messages assigned to each wavelength in a separate round.

When the assumption that the whole communication request is known in advance is unrealistic (see Section 4), we can use a probabilistic algorithm that randomly assigns routes to messages.

The idea is to adapt the algorithm of Section 2.2, but to route a request in a round with probability $1/(9kL^2)$. As a result, the expected number of other routes intersecting a given route is $1/9$. Consider a new graph in which each route is represented by a node, with an edge being present if the corresponding routes intersect. An argument from random graph theory can now be invoked to show that every component of intersecting routes has size $O(\log n)$ with high probability, so that $O(\log n)$ wavelengths suffice. Details are omitted. Recall that $\lambda(G)$ is the second largest eigenvalue in absolute value, of the network G:

Theorem 6: With high probability the algorithm routes any k-relation in

$$O(k(\log n)^2(\log\lambda(G))^2)$$

rounds, usings $c\log n$ wavelengths per round for some constant c.

3.2. Routing on Meshes

Theorem 7: There is a probabilistic algorithm that for any $w \leq \sqrt{nk}$ routes any k-relation with high probability in $O((\frac{k}{w} + \sqrt{\frac{k}{w}})\sqrt{n}\log n)$ rounds using $O(w)$ wavelengths per round.

Proof: If $w < 24k$ we can partition the routing problem to $O(k/w)$ sub-problems each of routing a w-relation. Thus, it suffices to consider the case $w > 24k$, so let $f = \sqrt{w/(6k)} > 2$. The algorithm has two phases. In the first phase we match pairs of columns that have at least f messages to route between them. In the second phase we partition the columns into sets of f, and route all the remaining messages between nodes in each of these sets. Let H be a multi-graph on $\sqrt{n}$ nodes; each node corresponds to one column in the mesh. Let m_{ij} be the number of messages from nodes in column i to nodes in column j of the mesh. Connect node i to node j in H with $\lfloor m_{ij}/(kf)\rfloor$ edges. Since we are routing a k-relation, the degree of a node in H is bounded by $\sqrt{n}k/kf = \sqrt{n}/f$, and the edges of H can be colored with $(3/2)(\sqrt{n}/f)$ colors. In each round of the first phase of the algorithm, using $kf < w$ wavelengths, we route all the messages that correspond to edges with a given color in H (these form a matching). Since there are no more than $\sqrt{n}/2$ such edges we can dedicate one row of the mesh to all requests corresponding to one edge of H. After $(3/2)(\sqrt{n}/f)$ rounds, no two columns of the mesh have more than kf messages to route between them.

The second phase of the algorithm consists of $3\sqrt{n}\log n/f = \sqrt{nk}\log n/\sqrt{w}$ rounds. In each round we randomly partition the $\sqrt{n}$ columns into $\sqrt{n}/f$ sets of f columns each, dedicate to each set a row, and try route all the messages between nodes in one set. If a set has no more than w messages between nodes in its columns then all messages are delivered, else we assume that no message in that set has been delivered. We say that a set is *good* at a given round if it has fewer than w messages between the nodes in its columns, so that successful transmission ensues.

Assume that at round τ columns i and j are in the same set A. What is the probability that this set is good at this round? Consider the other $f-2$ columns in that set. They were chosen at random from among all sets of $f-2$ columns that do not include columns i and j. The expected number of messages in a random set of $(f-2)$ columns that do not include i and j is bounded by $k(f-2)\sqrt{n}\frac{f-2}{\sqrt{n}-2} \leq 2k(f-2)^2$. Thus, with probability at least $1/2$ the $f-2$ other columns in A do not contribute more than $4k(f-2)^2$ messages, and the total number of messages in A is bounded by $4k(f-2)^2 + 2kf(f-2) + kf \leq 6kf^2 = w$. Thus, with probability at least $1/2$ the set is good.

To show that all the messages are delivered in $O(\sqrt{n}\log n/f)$ rounds we need to show that each pair of columns appears at least once in a good set. The probability that two columns do not share a good set in $3\sqrt{n}\log n/f$ rounds in bounded by

$$(\sqrt{n})^2(1-\frac{f}{2\sqrt{n}})^{3\sqrt{n}\log n/f} = n^{-\Omega(1)}.$$

Thus, with high probability all messages are transmitted. □

Using a more complicated version of the above technique we can obtain similar results for meshes of higher dimensions. Details are omitted.

4. Extensions

An important practical consideration is that of routing when the requests are *dynamic:* there is a sequence of time steps at each of which we either have a new request or are told that an existing request ceases to exist. Even in the case of non-optical networks, algorithmic work on this problem has been relatively recent [3, 4, 5, 6], and has taken the direction of extending known work on "offline" multicommodity flow; however, there is substantial evidence that adding wavelength constraints makes the problem much harder than online multicommodity flow.

Our offline algorithms for arbitrary graphs in both switch models (Sections 2.2 and 3.1) are *randomized oblivious algorithms:* a request chooses its route independently of other requests, and with high probability we do not use too many wavelengths. We pause to observe a very useful property of such algorithms: an adversary may specify (before the execution of the algorithm) a sequence of insertions and deletions of requests, subject to the set of requests being a k-relation at every point in time. Then, the bound on wavelengths remains valid for each step regardless of the adversary's choices; summing probabilities, we can tolerate this for a request sequence of length polynomial in n. All we require is that the adversary be *oblivious* [10]: the sequence is prescribed without knowledge of the actual random choices made by the algorithm. We omit a detailed proof of the following theorem.

Theorem 8: Let S be a sequence of request insertions and deletions of length $n^{O(1)}$, such that the requests valid at any time are a k-relation. Then, the number of wavelengths used with generalized switches is $O(k(\log n \log \lambda)^2)$, where λ is the second largest eigenvalue (in absolute value) of the adjacency matrix of the network.

An analogous result (building on Theorem 6) can be shown for elementary switches. The main difference here is that even though the number of wavelengths required at any time remains $O(\log n)$, an inserted request may require "recoloring" currently established paths. This happens when the path created for a newly inserted request joins two components of paths, both of which use the same wavelength. Using the *backwards analysis* trick common in computational geometry, we can show that the expected number of existing paths recolored at an insertion is a (small) constant, and with high probability is $O(\log n)$.

The major problem left open by our work is to tighten the gap between the upper and lower bounds for arbitrary graphs, particularly those of poor expansion. This would be especially interesting for elementary switches, for which our current algorithmic tools (as well as lower bounds) appear to be weak.

Acknowledgements: We thank Rajiv Ramaswami and Larry Rudolph for giving us insight into the practical aspects of optical communications.

References

[1] A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, and M. Sudan. Efficient routing and scheduling algorithms for optical networks. In *Proc. ACM-SIAM SODA.*, pages 412–423, January 1994.

[2] Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.

[3] J. Aspnes, Y. Azar, A. Fiat, S. Plotkin, and O. Waarts. On-line machine scheduling with applications to load balancing and virtual circuit routing. In *Proc. 25th Annual ACM Symposium on Theory of Computing*, pages 623–631, May 1993.

[4] B. Awerbuch, Y. Azar, and S. Plotkin. Throughput competitive on-line routing. In *Proc. 34th IEEE Annual Symposium on Foundations of Computer Science*, November 1993. To appear.

[5] B. Awerbuch, Y. Azar, S. Plotkin, and O. Waarts. Competitive routing of virtual circuits with unknown duration. In *Proc. 5th ACM-SIAM Symposium on Discrete Algorithms*, January 1994. To appear.

[6] Y. Azar, B. Kalyanasundaram, S. Plotkin, K. Pruhs, and O. Waarts. On-line load balancing of temporary tasks. In *Proc. Workshop on Algorithms and Data Structures*, pages 119–130, August 1993.

[7] K. Bala, T.E. Stern, and K. Bala. Algorithms for routing in a linear lightwave network. In *Proc. INFOCOM*, pages 1–9. IEEE, 1991.

[8] R. A. Barry and P. A. Humblet. On the number of wavelengths and switches in all-optical networks. To appear in *IEEE Trans. Comm.*, 1993.

[9] R. A. Barry and P. A. Humblet. Bounds on the number of wavelengths needed in WDM networks. In *LEOS'92 Summer Topical Mtg. Digest*, pages 114–127, 1992.

[10] S. Ben-David, A. Borodin, R.M. Karp, G. Tardos, and A. Wigderson. On the power of randomization in on-line algorithms. In *Proc. 22nd Annual ACM Symposium on Theory of Computing*, pages 379–388, 1990.

[11] C. Berge. *The Theory of Graphs and its Applications.* John Wiley, 1962.

[12] C. Brackett. Dense wavelength division multiplexing networks: Principles and applications. *IEEE J. Selected Areas in Comm.*, 8:373–380, August 1990.

[13] A.Z. Broder, A. Frieze, and E. Upfal. Existence and construction of edge disjoint paths on expander graphs. In *Proc. 24th ACM STOC*, pages 140–149. ACM, 1992.

[14] K-W. Cheng. Accousto-optic tunable filters in narrowband WDM networks. *IEEE JSAC*, 8:1015–1025, 1990.

[15] N.K. Cheung, K. Nosu, and G. Winzer. IEEE JSAC: special issue on dense WDM networks. *IEEE JSAC: Special Issue on Dense WDM Networks*, 8, 1990.

[16] P. E. Green. *Fiber-Optic Communication Networks.* Prentice Hall, 1992.

[17] H.S. Hinton. Architectural considerations for photonic switching networks. *IEEE J. Selected Areas in Comm.*, 6:1209–1226, August 1988.

[18] P. Klein, S. Plotkin, C. Stein, and É. Tardos. Faster approximation algorithms for the unit capacity concurrent flow problem with applications to routing and finding sparse cuts. *SIAM Journal on Computing*, 1992. Accepted for Publication.

[19] M. Kovacevic and M. Gerla. Rooted routing in a linear lightwave networks. In *Proc. INFOCOM*, pages 39–48. IEEE, 1992.

[20] T. Leighton, F. Makedon, S. Plotkin, C. Stein, S. Tragoudas, and É. Tardos. Fast approximation algorithms for multicommodity flow problem. *J. Comp. and Syst. Sci.*, 1992. Accepted for Publication.

[21] A.D. McAulay. *Optical Computer Architectures.* John Wiley, 1991.

[22] R.K. Pankaj. *Architectures for linear lightwave networks.* PhD thesis, MIT, 1992.

[23] S. Personick. Review of fundamentals of optical fiber systems. *IEEE J. Selected Areas in Comm.*, 3:373–380, April 1983.

[24] G. R. Pieris and G. H. Sasaki. A linear lightwave Beneš network. Submitted to IEEE/ACM Trans. on Networking.

[25] G. R. Pieris and G. H. Sasaki. Scheduling transmissions in broadcast and select networks. Unpublished manuscript, 1993.

[26] S. Plotkin, D. Shmoys, and É. Tardos. Fast approximation algorithms for fractional packing and covering problems. In *Proc. 32nd IEEE Annual Symposium on Foundations of Computer Science*, pages 495–504, October 1991.

[27] R. Ramaswami. Multi-wavelength lightwave networks for computer communication. *IEEE Communications Magazine*, 31:78–88, 1993.

[28] M. Settembre and F. Matera. All optical implementations of high capacity TDMA networks. *Fiber and Integrated Optics*, 12:173–186, 1993.

[29] A. Sinclair and M.R. Jerrum. Approximate counting, uniform generation and rapidly mixing Markov chains. *Information and Computation*, 82:93–133, 1989.

[30] R.J. Vetter and D.H.C. Du. Distributed computing with high-speed optical networks. *IEEE Computer*, 26:8–18, February 1993.