

# Load-Balanced Routing in Counter Rotated SONET Rings

Peng-Jun Wan

Department of Computer Science, Illinois Institute of Technology, Chicago, Illinois 60616

Yuanyuan Yang

Department of ECE, State University of New York at Stony Brook, Stony Brook, New York 11794

**Load-balanced routing in SONET rings has attracted much attention recently. Most prior works modeled the SONET rings as undirected rings and the traffic as undirected chords. While this model fits well to the traditional telephony applications, it is inefficient for the explosive Internet traffic and multimedia data communications, which exhibit an unidirectional and asymmetric nature. For these applications, it is proper to model the SONET rings as a pair of counter rotated rings and the traffic as directed chords. In this paper, we first explore general flow properties in counter rotated rings and then introduce flow rounding and unsplitting techniques. Afterward, an optimal integral routing algorithm is provided. Finally, we show the NP-completeness of optimal unsplit routing and present several polynomial-time approximation algorithms. © 2000 John Wiley & Sons, Inc.**

**Keywords:** SONET; counter rotated rings; routing; load balancing; approximation algorithms

## 1. INTRODUCTION

With the explosion of Internet traffic and multimedia data communication, the Synchronous Optical Network (SONET) has been adopted by many network service providers as a faster, more efficient, and less expensive transport technology [1]. While the fiber itself offers a virtually unlimited bandwidth, the add-drop multiplexers (ADMs) determine the actual bandwidth available along any fiber link of the SONET ring [8]. With the rapid growth of data traffic in recent years and the exhaustion of the fiber links in the plant, load-balanced routing

becomes an important problem in the planning of the SONET rings.

The optical transmission in the fibers is unidirectional in nature because of the unidirectional operation of optical amplifiers. Accordingly, a SONET ring usually consists of two working counter rotated fiber rings carrying the traffic in opposite directions, with additional transmission capacity provided for fault protection [4]. In the past, the dominant traffic carried was mainly the voice traffic which is both full-duplex (i.e., bidirectional) and symmetric (i.e., the traffic rates in the two directions are the same). In such an environment, the SONET rings can be modeled simply as undirected rings and traffic as undirected chords. The data traffic, on the other hand, has a different nature from the voice traffic in the sense that it is, in nature, either simplex (i.e., unidirectional) or asymmetrical. To reduce the capacity requirement and improve the utilization of fiber links, it is more proper to model the underlying network as a pair of counter rotated rings and each traffic demand as a directed chord. For example, in Figure 1, we show a request  $i$  in such a ring, where  $s_i$  and  $t_i$  are the source and target of the request, respectively, and  $d_i$  is the demand of the request. The demand can be viewed as a directed chord from  $s_i$  to  $t_i$ . In the actual routing, a portion of the demand  $d_i, x_i$ , can be routed in a clockwise direction and the other portion,  $d_i - x_i$ , can be routed in a counterclockwise direction. This paper focuses on the load-balancing problem in routing multiple requests in such ring networks to achieve a minimum link load.

Various restrictions might be imposed on the routings in practice. An *unsplit routing* is the one in which each demand must be carried either entirely clockwise or entirely counterclockwise. A *split routing*, on the other hand, allows the splitting of a demand into two portions to be carried in two directions. A demand, if allowed to be split, can be split in different ways: A *fractional routing* allows a demand to be split into two arbitrary por-

---

Received May 1999; accepted November 1999

Correspondence to: Y. Yang; e-mail: yang@ece.sunysb.edu

Contract grant sponsor: U.S. National Science Foundation; contract grant number: OSR-9350540.

Contract grant sponsor: U.S. Army Research Office; contract grant number: DAAH04-96-1-0234

© 2000 John Wiley & Sons, Inc.

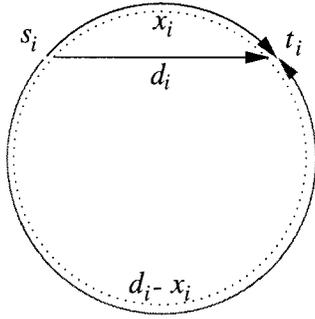


FIG. 1. Request  $i$  in a SONET ring.

tions to be carried in two directions. An *integral routing* only allows a demand to be split into two integral arbitrary portions to be carried in two directions. *Semi-integral routing* is a fractional routing with the additional constraint that the total demands routed in both directions are integers. It is a generalization of the concept of the flush-routing defined in [10] and serves as a bridge between the fractional routing and the integral routing. In real implementations, the integral routing and the unsplit routing are more practical and common.

The load-balanced routing in undirected rings has been well studied recently in [2,5,7–9]. They heavily rely upon many structural and flow properties such as the *cut condition* given by the well-known Okamura–Seymour theorem [6]. While some of them hold in both type of rings, many others such as the cut condition are specific to undirected rings. Therefore, new techniques and approaches are needed to develop optimal routings in counter rotated rings. Furthermore, we observe that some approaches used in undirected rings, such as the one used in [5], can be significantly improved and such improvements are reflected in our solutions to the counter rotated rings. A recent work [10] gave a polynomial-time optimal load-balancing unsplit routing in counter rotated rings when all requests have unit traffic demands. However, this algorithm cannot be generalized to optimal unsplit routing under the arbitrary traffic which is NP-complete. Also, its generalization to the integral routing under arbitrary traffic leads to only a pseudopolynomial time algorithm, which is not acceptable.

The remainder of this paper proceeds as follows: Section 2 describes the terminologies and notations. Section 3 explores basic flow properties in counter rotated rings, which lead to an algorithm for optimal semi-integral routing. Section 4 introduces the concept of parallel routing and an unsplitting technique which are very useful in the design of optimal integral routing and optimal unsplit routing. Section 5 presents a rounding technique which transforms any optimal parallel semi-integral routing into an optimal integral routing. Section 6 shows the NP-completeness of optimal unsplit routing and provides several polynomial-time approximation algorithms. Finally, Section 7 concludes this paper.

## 2. TERMINOLOGIES AND NOTATIONS

We assume that a SONET ring consists of  $n$  nodes labeled clockwise by 0 through  $n - 1$ . All arithmetic involving nodes is performed implicitly using modulo  $n$  operations. For nodes  $s$  and  $t$ , the half-closed arc  $\{s, s + 1, \dots, t - 1\}$  is denoted by  $[s, t)$  and the closed arc  $\{s, s + 1, \dots, t\}$  is denoted by  $[s, t]$ . The readers should be able to tell whether an interval is an arc or a normal interval of real numbers in the context.

The traffic in the ring consists of  $m$  unidirectional requests. The requests are numbered clockwise (starting from node 0) by their sources. The requests sharing the same source are numbered clockwise (starting from the source) by their targets. Furthermore, the requests sharing the same source and target are numbered arbitrarily. As we will see later, such numbering plays an important role in describing the structural and flow properties. The source, target, and demand of the request  $i$  are denoted by  $s_i, t_i$ , and  $d_i$ , respectively. The demand represents the number of time slots required and, therefore, is always an integer. For presentational convenience, we use  $d$  to denote the  $m$ -dimensional vector  $(d_1, d_2, \dots, d_m)$ . A (*feasible*) routing in which the portion of the demand of request  $i$  to be routed clockwise is  $x_i$  is represented by an  $m$ -dimensional vector  $x = (x_1, x_2, \dots, x_m)$ , where  $0 \leq x_i \leq d_i$ . For each vector  $x = (x_1, x_2, \dots, x_m)$ , we use  $\|x\|$  to denote  $\sum_{i=1}^m x_i$ . Therefore,  $\|d\|$  is the total demand of all requests in the ring, and for each routing  $x$ ,  $\|x\|$  is the total demand routed clockwise by  $x$ .

Let  $x$  be any routing. The loads of the link  $k \rightarrow k + 1$  and link  $k + 1 \rightarrow k$  induced by  $x$  are

$$l_k^+(x) = \sum_{i:k \in [s_i, t_i)} x_i, l_k^-(x) = \sum_{i:k \notin [s_i, t_i)} (d_i - x_i),$$

respectively. The clockwise and counterclockwise ring loads of  $x$  are

$$l^+(x) = \max_{0 \leq k \leq n-1} l_k^+(x), l^-(x) = \max_{0 \leq k \leq n-1} l_k^-(x),$$

respectively. Finally, the ring load of the routing  $x$  is

$$l(x) = \max\{l^+(x), l^-(x)\}.$$

The optimal fractional, semi-integral, integral, and unsplit ring load are denoted by  $L_F^*, L_{SI}^*, L_I^*$ , and  $L_U^*$ , respectively. Obviously,

$$L_F^* \leq L_{SI}^* \leq \lceil L_{SI}^* \rceil \leq L_I^* \leq L_U^*.$$

Tighter relationships among them will be derived later in this paper.

Among all these optimal routing problems, only the optimal fractional routing problem is obviously solvable in polynomial time by solving a linear program. For each  $\alpha \in [0, \|d\|]$ , define

$$L(\alpha) = \min_{\sum_{i=1}^m x_i = \alpha, x_i \in [0, d_i], 1 \leq i \leq m} l(x). \quad (1)$$

Then,  $L(\alpha)$  can also be obtained in polynomial time by solving a linear program. Notice that

$$L_{SI}^* = \min_{\alpha \in \{0, 1, \dots, \|d\|\}} L(\alpha).$$

This might suggest one algorithm for the optimal semi-integral routing as follows: For each  $\alpha = 0, 1, \dots, \|d\|$ , find a routing  $x$  with  $\|x\| = \alpha$  and  $l(x) = L(\alpha)$  and then take the best one. However, the time-complexity of such an algorithm is  $\Omega(\|d\|)$ . Noticing that  $\|d\|$  could be as exponentially large as  $n$  or  $m$ , the algorithm is thus pseudopolynomial and is not acceptable.

### 3. BASIC STRUCTURAL AND FLOW PROPERTIES

In this section, we will explore several elegant properties of the function  $L$  defined in Eq. (1) in the last section. To begin with, we first observe the convexity of the function  $L$ :

**Lemma 1.** *The function  $L$  is convex over the interval  $[0, \|d\|]$ .*

**Proof.**  $\forall 0 \leq \alpha_1 < \alpha_2 \leq \|d\|$ , and  $\forall \lambda \in (0, 1)$ , we want to prove that

$$L(\lambda\alpha_1 + (1 - \lambda)\alpha_2) \leq \lambda L(\alpha_1) + (1 - \lambda)L(\alpha_2).$$

Let  $x$  be any routing with  $\|x\| = \alpha_1$  and  $l(x) = L(\alpha_1)$ . Let  $y$  be any routing with  $\|y\| = \alpha_2$  and  $l(y) = L(\alpha_2)$ . Let  $z = \lambda x + (1 - \lambda)y$ . Then,  $z$  is a feasible routing and  $L(\lambda\alpha_1 + (1 - \lambda)\alpha_2) \leq l(z)$  as  $\|z\| = \lambda\alpha_1 + (1 - \lambda)\alpha_2$ . For  $0 \leq k \leq n - 1$ ,

$$\begin{aligned} l_k^+(z) &= \lambda l_k^+(x) + (1 - \lambda)l_k^+(y) \\ &\leq \lambda L(\alpha_1) + (1 - \lambda)L(\alpha_2) \\ l_k^-(z) &= \lambda l_k^-(x) + (1 - \lambda)l_k^-(y) \\ &\leq \lambda L(\alpha_1) + (1 - \lambda)L(\alpha_2). \end{aligned}$$

Therefore,  $l(z) \leq \lambda L(\alpha_1) + (1 - \lambda)L(\alpha_2)$  and, thus, the lemma follows.  $\blacksquare$

Lemma 1 provides a simple way to find optimal semi-integral routing. Let  $\alpha^*$  be the total demand routed clockwise by any optimal fractional routing. Then, Lemma 1 implies that the function  $L$  is non-increasing over the interval  $[0, \alpha^*]$  and nondecreasing over the interval  $[\alpha^*, \|d\|]$ . This follows that  $L_{SI}^* = \min\{L(\lfloor \alpha^* \rfloor), L(\lceil \alpha^* \rceil)\}$ , and it can be obtained as follows:

1. Find an optimal fractional routing  $x$ . If  $\|x\|$  is an integer, return  $x$ .
2. Find a semi-integral routing  $y$  with  $\|y\| = \lfloor \|x\| \rfloor$  and  $l(y) = L(\lfloor \|x\| \rfloor)$ .
3. Find a semi-integral routing  $z$  with  $\|z\| = \lceil \|x\| \rceil$  and  $l(z) = L(\lceil \|x\| \rceil)$ .
4. If  $l(y) \leq l(z)$ , return  $y$ , otherwise return  $z$ .

The following theorem summarizes the above discussions:

**Theorem 2.** *The optimal semi-integral routing problem can be solved in polynomial time; furthermore,  $L_{SI}^* = \min\{L(\lfloor \alpha^* \rfloor), L(\lceil \alpha^* \rceil)\}$ , where  $\alpha^*$  is the total demand routed clockwise by any optimal fractional routing.*

Intuitively, if the total demand routed clockwise by a routing is very small, then the clockwise ring load would be smaller than the counterclockwise ring load. When the total demand routed clockwise by a routing becomes very large, then the counterclockwise ring load would be smaller than the clockwise ring load. Such intuition is verified by the following lemma:

**Lemma 3.** *Let  $\alpha^*$  be the total demand routed clockwise by any optimal fractional routing and  $x$  be any routing. Then,  $l(x) = l^-(x)$  if  $\|x\| \leq \alpha^*$ , and  $l(x) = l^+(x)$  otherwise.*

**Proof.** We prove the lemma by contradiction. We first prove the first part: Assume that  $\|x\| \leq \alpha^*$  and  $l^-(x) < l^+(x) = L(\|x\|)$ . Then,  $\exists 0 < \epsilon \leq 1$ , such that  $\forall 1 - \epsilon \leq \lambda \leq 1, l^-(\lambda x) \leq l^+(\lambda x)$ . For any  $0 < \delta \leq \epsilon \|x\|, (1 - (\delta/\|x\|))x$  is a feasible routing. As  $\|(1 - (\delta/\|x\|))x\| = \|x\| - \delta$  and  $1 - \epsilon \leq 1 - (\delta/\|x\|) \leq 1$ ,

$$\begin{aligned} &L(\|x\| - \delta) \\ &\leq l\left(\left(1 - \frac{\delta}{\|x\|}\right)x\right) = l^+\left(\left(1 - \frac{\delta}{\|x\|}\right)x\right) \\ &= \left(1 - \frac{\delta}{\|x\|}\right)l^+(x) = \left(1 - \frac{\delta}{\|x\|}\right)L(\|x\|) \\ &< L(\|x\|). \end{aligned}$$

This contradicts that the function  $L$  is nonincreasing over the interval  $[0, \alpha^*]$ , which is implied by the concavity of  $L$ .

Now, we prove the second part: Assume that  $\|x\| \geq \alpha^*$  and  $l^+(x) < l^-(x) = L(\|x\|)$ . Then,  $\exists 0 < \epsilon \leq 1$ , such that  $\forall 0 \leq \lambda \leq \epsilon, l^+(x + \lambda(d - x)) \leq l^-(x + \lambda(d - x))$ . For any  $0 < \delta \leq \epsilon(\|d\| - \|x\|), x + (\delta/(\|d\| - \|x\|))(d - x)$  is a feasible routing. As  $\|x + (\delta/(\|d\| - \|x\|))(d - x)\| = \|x\| + \delta$  and  $0 < (\delta/(\|d\| - \|x\|)) \leq \epsilon$ ,

$$\begin{aligned} L(\|x\| + \delta) &\leq l\left(x + \frac{\delta}{\|d\| - \|x\|}(d - x)\right) \\ &= l^-\left(x + \frac{\delta}{\|d\| - \|x\|}(d - x)\right) \\ &= \left(1 - \frac{\delta}{\|d\| - \|x\|}\right)l^-(x) \\ &= \left(1 - \frac{\delta}{\|d\| - \|x\|}\right)L(\|x\|) < L(\|x\|). \end{aligned}$$

This contradicts that the function  $L$  is nondecreasing over the interval  $[\alpha^*, \|d\|]$ .  $\blacksquare$

Now, we are ready to show that the function  $L$  is Lipschitz continuous, that is, there is an  $\epsilon > 0$  such that for any  $0 \leq \alpha, \beta \leq \|d\|$ ,  $|L(\alpha) - L(\beta)| \leq \epsilon|\alpha - \beta|$ .

**Lemma 4.** *Let  $\alpha^*$  be the total demand routed clockwise by any optimal fractional routing. The following statements are true:*

1. If  $\alpha \leq \alpha^*$ , then for any  $0 \leq \delta \leq \alpha$ ,  $L(\alpha) \leq L(\alpha - \delta) \leq L(\alpha) + \delta$ .
2. If  $\alpha \geq \alpha^*$ , then for any  $0 \leq \delta \leq \|d\| - \alpha$ ,  $L(\alpha) \leq L(\alpha + \delta) \leq L(\alpha) + \delta$ .
3. For any  $\alpha$  and  $\delta$  with  $0 \leq \alpha \leq \alpha + \delta \leq \|d\|$ ,  $|L(\alpha + \delta) - L(\alpha)| \leq \delta$ .

**Proof.** Let  $x$  be any routing with  $\|x\| = \alpha$  and  $l(x) = L(\alpha)$ .

1. The first inequality follows from that the function  $L$  is nonincreasing over the interval  $[0, \alpha^*]$ . Next, we prove the second inequality: For any  $0 \leq \delta \leq \alpha$ ,  $(1 - (\delta/\|x\|))x$  is a feasible routing and  $\|(1 - (\delta/\|x\|))x\| = \alpha - \delta$ . Thus, from Lemma 3,

$$\begin{aligned} L(\alpha - \delta) &\leq l\left(\left(1 - \frac{\delta}{\|x\|}\right)x\right) \\ &= l^-\left(\left(1 - \frac{\delta}{\|x\|}\right)x\right) \\ &= \max_{0 \leq k \leq n-1} l_k^-\left(\left(1 - \frac{\delta}{\|x\|}\right)x\right) \\ &= \max_{0 \leq k \leq n-1} \left(l_k^-(x) + \frac{\delta}{\|x\|} \sum_{i: k \notin [s_i, t_i]} x_i\right) \\ &\leq \max_{0 \leq k \leq n-1} (l_k^-(x) + \delta) = l^-(x) + \delta \\ &= l(x) + \delta = L(\alpha) + \delta. \end{aligned}$$

2. The first inequality follows from that the function  $L$  is nondecreasing over the interval  $[\alpha^*, \|d\|]$ . For any  $0 < \delta \leq \|d\| - \alpha$ ,  $x + (\delta/(\|d\| - \|x\|))(d - x)$  is a feasible routing and  $\|x + (\delta/(\|d\| - \|x\|))(d - x)\| = \alpha + \delta$ . Thus, from Lemma 3,

$$\begin{aligned} L(\alpha + \delta) &\leq l\left(x + \frac{\delta}{\|d\| - \|x\|} (d - x)\right) \\ &= l^+\left(x + \frac{\delta}{\|d\| - \|x\|} (d - x)\right) \\ &= \max_{0 \leq k \leq n-1} l_k^+\left(x + \frac{\delta}{\|d\| - \|x\|} (d - x)\right) \\ &= \max_{0 \leq k \leq n-1} \left(l_k^+(x) + \frac{\delta}{\|d\| - \|x\|} \sum_{i: k \in [s_i, t_i]} (d_i - x_i)\right) \\ &\leq \max_{0 \leq k \leq n-1} (l_k^+(x) + \delta) = l^+(x) + \delta \\ &= l(x) + \delta = L(\alpha) + \delta. \end{aligned}$$

3. If  $0 \leq \alpha \leq \alpha + \delta \leq \alpha^*$  or  $\alpha^* \leq \alpha \leq \alpha + \delta \leq \|d\|$ , the inequality follows from (1) and (2), respectively. If  $0 \leq \alpha \leq \alpha^* \leq \alpha + \delta \leq \|d\|$ , we have

$$\begin{aligned} |L(\alpha + \delta) - L(\alpha)| &\leq |L(\alpha^*) - L(\alpha)| + |L(\alpha + \delta) - L(\alpha^*)| \\ &\leq (\alpha^* - \alpha) + (\alpha + \delta - \alpha^*) = \delta, \end{aligned}$$

where the last inequality also follows from (1) and (2). ■

Based on the above lemma, we can bound the difference between the optimal semi-integral ring load and the optimal fractional ring load.

**Corollary 5.**  $L_F^* \leq L_{SI}^* \leq L_F^* + \frac{1}{2}$ .

**Proof.** The first inequality is obvious. We next prove the second inequality: From Lemma 4,

$$\begin{aligned} L(\lfloor \alpha^* \rfloor) &\leq L(\alpha^*) + \alpha^* - \lfloor \alpha^* \rfloor = L_F^* + \alpha^* - \lfloor \alpha^* \rfloor, \\ L(\lceil \alpha^* \rceil) &\leq L(\alpha^*) + \lceil \alpha^* \rceil - \alpha^* = L_F^* + \lceil \alpha^* \rceil - \alpha^*. \end{aligned}$$

Thus, from Theorem 2,

$$\begin{aligned} L_{SI}^* &= \min\{L(\lfloor \alpha^* \rfloor), L(\lceil \alpha^* \rceil)\} \\ &\leq L_F^* + \min\{\alpha^* - \lfloor \alpha^* \rfloor, \lceil \alpha^* \rceil - \alpha^*\} \\ &\leq L_F^* + \frac{1}{2}. \end{aligned}$$

#### 4. PARALLEL ROUTING

In this section, we introduce the concept of parallel routing [10] and the unsplit technique. Two requests  $i$  and  $j$  are said to be *parallel* if either  $[s_i, t_i] \subseteq [s_j, t_j]$  or  $[s_j, t_j] \subseteq [s_i, t_i]$ ; otherwise, they are said to be *crossing*. Regarding a link also as a directed chord, a request is parallel to a link just when the request can be routed through that link. Thus, any link partitions the requests into two groups: those which are parallel to it and those which are parallel to its reverse. A routing is said to be *parallel* if no two parallel requests are both split. In any parallel routing, any two split requests must be crossing and, therefore, cannot share a source. This implies that the number of split requests is at most the ring size  $n$ . Moreover, the targets of any two split requests must follow the same clockwise order of their sources. Thus, we have the following lemma:

**Lemma 6.** *Any parallel routing splits at most  $n$  requests where  $n$  is the ring size. Moreover, the targets of those split requests follow the same clockwise order of their sources.*

Now, we describe how to obtain a parallel routing from any given routing by unsplitting some requests without increasing the ring load. The following lemma generalizes the transforming technique in [10]:

**Lemma 7.** Any routing  $x$  can be transformed to a parallel routing  $y$  in polynomial time satisfying that  $\|y\| = \|x\|$  and every link load is not increased.

**Proof.** Suppose that requests  $i$  and  $j$  are a pair of parallel requests that are both split by  $x$ . Without loss of generality, we assume that  $[s_i, t_i] \subseteq [s_j, t_j]$ , as shown in Figure 2(a).

If  $x_i + x_j > d_i$ , then define a new routing  $y$  by setting

$$\begin{aligned} y_i &= d_i, \\ y_j &= x_i + x_j - d_i, \\ y_k &= x_k, \quad \forall k \neq i, j, \end{aligned}$$

as illustrated in Figure 2(b). If  $x_i + x_j \leq d_i$ , we define a new routing  $y$  by setting

$$\begin{aligned} y_i &= x_i + x_j, \\ y_j &= 0, \\ y_k &= x_k, \quad \forall k \neq i, j, \end{aligned}$$

as illustrated in Figure 2(c). In both cases, one of two requests is no longer split, every link load is either maintained or reduced, and the total demand routed clockwise remains unchanged. This procedure can be performed repeatedly until no two parallel requests are both split. Since each time we reduce by one (or two) the total number of split demands, at most  $m$  such procedures will produce the desired routing.  $\blacksquare$

By applying the unsplitting technique in Lemma 7 to any optimal fractional routing, we can get an optimal fractional routing which is also parallel, referred to as an *optimal parallel fractional routing*. Similarly, by applying the same unsplitting technique to any optimal semi-integral routing, we get an optimal semi-integral routing which is also parallel, referred to as an *optimal parallel semi-integral routing*.

## 5. INTEGRAL ROUTING

In this section, we first describe how to round parallel semi-integral routing to an integral routing. The following lemma generalizes the rounding techniques used in [8] and [10]. For the sake of completeness, we provide the generalized proof as well.

**Lemma 8.** Any parallel semi-integral routing  $x$  can be rounded in polynomial time into an integral routing  $y$  satisfying that  $\|y\| = \|x\|$  and the increase of every link load is less than one.

**Proof.** From Lemma 6, the number of requests split by  $x$  is at most the ring size  $n$ . In particular, the number of fractionally split requests is at most  $n$ . Let  $\{f_1, f_2, \dots, f_q\}$  be the set of fractionally split requests with  $f_1 < f_2 < \dots < f_q$ , where  $q \leq n$ . We define an

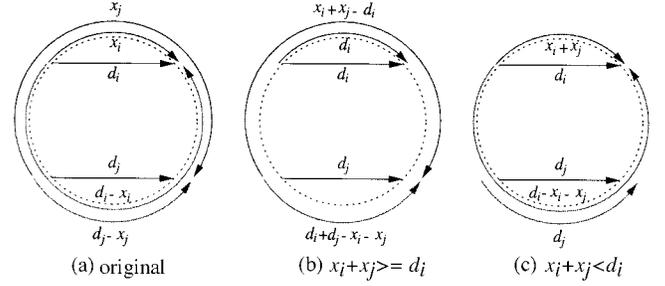


FIG. 2. Unsplit one request in a parallel pair.

integral routing  $y$  by rounding fractionally split requests of  $x$  as follows: For each request  $i \notin \{f_1, f_2, \dots, f_q\}$ , set  $y_i = x_i$ . We then define integers  $y_{f_1}, y_{f_2}, \dots, y_{f_q}$  sequentially by ensuring every partial sum

$$\sum_{j=1}^i (y_{f_j} - x_{f_j}) \in \left[-\frac{1}{2}, \frac{1}{2}\right)$$

for all  $1 \leq i \leq q$ . Since  $x$  is semi-integral and  $y$  is integral, the sum  $\sum_{i=1}^m (y_i - x_i) = \sum_{i=1}^m y_i - \sum_{i=1}^m x_i$  is an integer. Thus,

$$\sum_{i=1}^m (y_i - x_i) = \sum_{j=1}^q (y_{f_j} - x_{f_j}) \in \left[-\frac{1}{2}, \frac{1}{2}\right)$$

implies that

$$\sum_{i=1}^m (y_i - x_i) = \sum_{j=1}^q (y_{f_j} - x_{f_j}) = 0,$$

that is,  $\|y\| = \|x\|$ .

Now, we give the second part of the lemma: By Lemma 6, the sources and targets of requests  $\{f_1, f_2, \dots, f_q\}$  are positioned in the ring in the same clockwise order. Thus, for any  $0 \leq k \leq n - 1$ , there is an interval  $[a_k, b_k] \subseteq \{1, \dots, q\}$ , interpreted if necessary “around the corner” modulo  $q$ , such that

- The set  $\{f_j : j \in [a_k, b_k]\}$  contains exactly the indices of those fractionally split requests which are parallel to the link  $k \rightarrow k + 1$ ;
- The set  $\{f_j : j \notin [a_k, b_k]\}$  contains exactly the indices of those fractionally split requests which are parallel to the link  $k + 1 \rightarrow k$ .

Therefore, the load increment of the clockwise link  $k \rightarrow k + 1$  is

$$l_k^+(y) - l_k^+(x) = \sum_{j \in [a_k, b_k]} (y_{f_j} - x_{f_j}).$$

If  $a_k \leq b_k$ ,

$$\begin{aligned} l_k^+(y) - l_k^+(x) &= \sum_{j \in [a_k, b_k]} (y_{f_j} - x_{f_j}) \\ &= \sum_{j=1}^{b_k} (y_{f_j} - x_{f_j}) - \sum_{j=1}^{a_k-1} (y_{f_j} - x_{f_j}) \\ &< \frac{1}{2} - \left(-\frac{1}{2}\right) = 1. \end{aligned}$$

If  $a_k > b_k$ ,

$$\begin{aligned}
l_k^+(y) - l_k^+(x) &= \sum_{j \in [a_k, b_k]} (y_{f_j} - x_{f_j}) \\
&= \sum_{j=1}^q (y_{f_j} - x_{f_j}) + \sum_{j=1}^{b_k} (y_{f_j} - x_{f_j}) - \sum_{j=1}^{a_k-1} (y_{f_j} - x_{f_j}) \\
&= \sum_{j=1}^{b_k} (y_{f_j} - x_{f_j}) - \sum_{j=1}^{a_k-1} (y_{f_j} - x_{f_j}) \\
&< \frac{1}{2} - \left(-\frac{1}{2}\right) = 1
\end{aligned}$$

as well. A symmetric argument for the counterclockwise links shows that the link load increment is also less than one. ■

Let  $x$  be any optimal parallel semi-integral routing. If it is integral,  $L_I^* = L_{SI}^* = \lceil L_{SI}^* \rceil$ ; otherwise, we apply the rounding technique in Lemma 8 to  $x$  and let  $y$  be the resulting integral routing. Then,  $l(y) < l(x) + 1 = L_{SI}^* + 1$ , which implies that  $l(y) \leq \lceil L_{SI}^* \rceil$ . As  $l(y) \geq L_I^* \geq \lceil L_{SI}^* \rceil$ ,  $y$  is an optimal integral routing and  $L_I^* = \lceil L_{SI}^* \rceil$ . Therefore, we have the following theorem:

**Theorem 9.** *The optimal integral routing problem can be solved in polynomial time. Furthermore,  $\lceil L_F^* \rceil \leq \lceil L_{SI}^* \rceil = L_I^* \leq \lceil L_F^* + \frac{1}{2} \rceil$ ; in particular, if  $L_{SI}^*$  is an integer,  $L_I^* = L_{SI}^*$ .*

## 6. UNSPLIT ROUTING

Unlike the optimal integral routing which can be solved in polynomial time, the optimal unsplit routing is NP-complete. Therefore, we shift our attention to polynomial-time approximation algorithms. In this section, we first show the NP-completeness of the optimal unsplit routing. We then present several polynomial-time approximation algorithms from simple to complex.

**Lemma 10.** *The optimal unsplit routing problem is NP-complete even with the following patterns of requests:*

1. All requests share the same source;
2. Any pair of requests is crossing;
3. Any pair of requests is parallel but shares no source or target.

**Proof.** A simple reduction is available from the PARTITION problem [3], in which positive integers  $d_1, d_2, \dots, d_m$  are given and the question is whether one can divide them into two groups of equal sum.

1. Set  $n = m + 1$  and construct the  $m$  requests as follows: All these  $m$  requests have the node 0 as their source, the target of the request  $i$  is node  $i$ , and the demand of request  $i$  is  $d_i$  [see Fig. 3(a)]. Then, any request routed clockwise must pass through the link  $0 \rightarrow 1$

and any request routed counterclockwise must pass through the link  $0 \rightarrow m$ . Therefore, for any routing  $x$ ,  $l^+(x) = \|x\|$  and  $l^-(x) = \sum_{i=1}^m d_i - \|x\|$ . This implies that the optimal unsplit ring load is at least  $\sum_{i=1}^m d_i / 2$  and the ring load is optimal if and only if the set of integers  $d_1, d_2, \dots, d_m$  can be divided into two groups of equal sum.

2. Set  $n = 2m$  and construct the  $m$  requests as follows: The request  $i$  is from node  $i - 1$  to node  $m + i - 1$  with demand  $d_i$  [see Fig. 3(b)]. Then, any request routed clockwise must pass through the link  $m - 1 \rightarrow m$  and any request routed counterclockwise must pass through the link  $0 \rightarrow 2m - 1$ . Therefore, for any routing  $x$ ,  $l^+(x) = \|x\|$  and  $l^-(x) = \sum_{i=1}^m d_i - \|x\|$ . This implies that the optimal unsplit ring load is at least  $\sum_{i=1}^m d_i / 2$  and the ring load is optimal if and only if the set of integers  $d_1, d_2, \dots, d_m$  can be divided into two groups of equal sum.
3. Set  $n = 2m$  and construct the  $m$  requests as follows: The request  $i$  is from node  $i - 1$  to node  $2m - i$  with demand  $d_i$  [see Fig. 3(c)]. Then, any request routed clockwise must pass through the link  $m - 1 \rightarrow m$  and any request routed counterclockwise must pass through the link  $0 \rightarrow 2m - 1$ . Therefore, for any routing  $x$ ,  $l^+(x) = \|x\|$  and  $l^-(x) = \sum_{i=1}^m d_i - \|x\|$ . This implies that the optimal unsplit ring load is at least  $\sum_{i=1}^m d_i / 2$  and the ring load is optimal if and only if the set of integers  $d_1, d_2, \dots, d_m$  can be divided into two groups of equal sum. ■

Next, we will present several polynomial-time approximation algorithms for unsplit routing.

### 6.1. Edge-Avoidance Routing

For any  $0 \leq i \leq n - 1$ , the  $(i, i + 1)$ -avoidance routing routes each request along the unique path which avoids the link  $i \rightarrow i + 1$  and the link  $i + 1 \rightarrow i$ . An alternative interpretation is to cut these two links, turning the counter rotated ring into two unidirectional chains, on which each request has a unique path. Despite the simplicity of the edge-avoidance routing, the next theorem states that the ring load of any edge-avoidance routing is within twice the optimal unsplit ring load.

**Theorem 11.** *Edge-avoidance routing is a 2-approximation.*

**Proof.** Let  $x$  be any optimal unsplit routing and  $i$  be any node. Then, the reversing of all requests routed through the link  $i \rightarrow i + 1$  in  $x$  induces a load of no more than  $L_U^*$  on any other counterclockwise link. Similarly, the reversing of all requests routed through the link  $i + 1 \rightarrow i$  in  $x$  also induces a load of no more than  $L_U^*$  on any other clockwise link. Therefore, the ring load of the  $(i, i + 1)$ -avoidance routing is at most  $2L_U^*$ . ■

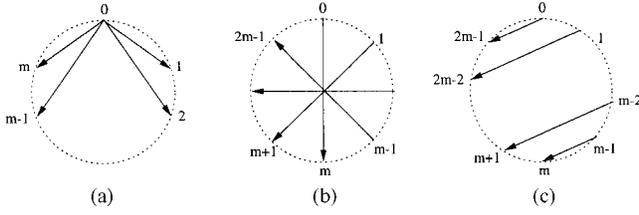


FIG. 3. Reductions from the PARTITION problem.

## 6.2. Short-way Routing

A request is said to be routed in the short (or long) way if it is routed along the shorter (or longer) of the two paths connecting the request. If the two paths connecting the request have the same length, one of them is chosen arbitrarily as the short way and the other as the long way. The *short-way routing* is a routing in which each request is routed the short way. The next theorem states that the ring load of any short-way routing is within twice the optimal unsplit ring load.

**Theorem 12.** *Short-way routing is a 2-approximation.*

**Proof.** Let  $x$  be any optimal unsplit routing and  $y$  be any short-way routing. We first assume that  $l(y) = l_k^+(y)$  for some  $0 \leq k \leq n-1$ . Let  $S$  be the set of requests whose routes pass through the link  $k \rightarrow k+1$  in  $y$ . Then,  $l(y) = \sum_{i \in S} d_i$ . Let  $S'$  be the set of requests in  $S$  which are also routed through the link  $k \rightarrow k+1$  in  $x$ . Then,  $l_k^+(x) \geq \sum_{i \in S'} d_i$ . But, on the other hand, all requests in  $S - S'$  must be routed through the link  $\lfloor n/2 \rfloor + k + 1 \rightarrow \lfloor n/2 \rfloor + k$ . This implies that  $l_{\lfloor n/2 \rfloor + k}^-(x) \geq \sum_{i \in S - S'} d_i = l(y) - \sum_{i \in S'} d_i$ . Therefore,

$$\begin{aligned} L_U^* &= l(x) \geq \max\{l_k^+(x), l_{\lfloor n/2 \rfloor + k}^-(x)\} \\ &\geq \max\left\{\sum_{i \in S'} d_i, l(y) - \sum_{i \in S'} d_i\right\} \geq l(y)/2. \end{aligned}$$

Similarly, we can prove that  $L_U^* \geq l(y)/2$  if  $l(y) = l_k^-(y)$  for some  $0 \leq k \leq n-1$ .  $\blacksquare$

## 6.3. $(1 + \epsilon)$ -Approximation Algorithm

In this subsection, we will show that for any fixed  $\epsilon > 0$  we can find in polynomial time an unsplit routing whose ring load is within  $(1 + \epsilon)$  times the optimum. The approximation algorithm uses a rounding technique which obtains a suboptimal unsplit routing from a parallel optimal fractional routing.

**Lemma 13.** *Any parallel routing  $x$  can be rounded in polynomial time into an unsplit routing  $y$  satisfying that  $l(y) - l(x) < 3d_{\max}/2$ , where*

$$d_{\max} = \max\{d_i | 0 < x_i < d_i\},$$

that is, the maximum of the demands of split requests by  $x$ .

**Proof.** The proof given here is similar to that in Lemma 8 with minor modifications: Let  $\{f_1, f_2, \dots, f_q\}$  be the set of split requests with  $f_1 < f_2 < \dots < f_q$ , where  $q \leq n$ . Define an unsplit routing  $y$  by unsplitting the requests split by  $x$  as follows: For all requests  $i \notin \{f_1, f_2, \dots, f_q\}$ , set  $y_i = x_i$ . Then, define  $y_{f_1}, y_{f_2}, \dots, y_{f_q}$  sequentially by ensuring every partial sum  $\sum_{j=1}^i (y_{f_j} - x_{f_j}) \in [-d_{\max}/2, d_{\max}/2]$  for all  $1 \leq i \leq q$ . Then, we can prove that  $l(y) - l(x) < 3d_{\max}/2$  in a similar way as we did in the proof of Lemma 8. The detail is omitted here.  $\blacksquare$

By applying the transform described in Lemma 13 to an optimal parallel fractional routing, we get an unsplit routing that does not exceed  $L_F^*$  by  $3/2$  times the maximum demand. In addition, Lemma 13 implies that if a routing  $x$  with  $l(x) \leq L_U^*$  splits no request with demand more than  $2\epsilon L_U^*/3$  and  $l(x) \leq L_U^*$  then it can be transformed in polynomial time to an unsplit routing with load no more than  $(1 + \epsilon)L_U^*$ . Next, we will study how to find such a routing  $x$  in polynomial time: Let  $L^*$  denote the ring load of any short-way routing. A request is said to be *light* if its demand is at most  $\epsilon L^*/3$  ( $\leq 2\epsilon L_U^*/3$ ) and *heavy* otherwise. The next lemma bounds the total number of heavy requests.

**Lemma 14.** *There are fewer than  $6n/\epsilon$  heavy requests.*

**Proof.** The proof uses the classical double-counting technique. For any  $0 \leq i \neq j \leq n-1$ , the pair of links  $i \rightarrow i+1$  and  $j+1 \rightarrow j$  are called a *cut*. A request is said to be *across* a cut if each of its two paths of the request passes through one link in the cut. Obviously, the total demand of all requests across any cut cannot exceed  $2L_F^*$ . Let  $h$  be the number of heavy requests. Since every request participates in at least  $n-1$  cuts and each heavy request has demand at least  $\epsilon L^*/3$ ,

$$(n-1) \frac{\epsilon L^*}{3} h < n(n-1) \cdot 2L_F^*.$$

The left-hand side gives a (loose) lower bound on the total contribution to the requests across all possible cuts, while the right-hand side specifies an upper bound on the aggregated demand across all possible cuts. As  $L_F^* \leq L^*$ ,  $h < 6n/\epsilon$ .  $\blacksquare$

The next lemma bounds the total number of heavy requests that could be routed in the long way in any optimal unsplit routing:

**Lemma 15.** *In any optimal unsplit routing, less than  $12/\epsilon$  heavy requests are routed in the long way.*

**Proof.** Suppose to the contrary that at least  $12/\epsilon$  heavy requests were routed in the long way in some optimal unsplit routing  $x$ . As a request that is routed the long way must traverse at least  $\lfloor n/2 \rfloor$  links, the total load induced by those heavy requests routed in the long way

is more than  $\lceil n/2 \rceil \cdot (\varepsilon L^*)/3 \cdot 12/\varepsilon \geq 2nL^*$ . By the pigeonhole principle, some link must have load more than  $((2nL^*)/(2n)) = L^* \geq L_U^*$ , which contradicts the optimality of  $x$ . ■

Let  $\Pi$  denote the collection of all sets consisting of less than  $12/\varepsilon$  heavy requests in an unsplit routing. From Lemma 14 and the well-known inequality  $\binom{n}{k} \leq ((en)/k)^k$ ,

$$\begin{aligned} |\Pi| &< \sum_{i=0}^{12/\varepsilon} \binom{6n/\varepsilon}{i} \leq \frac{12}{\varepsilon} \left( \frac{e \cdot 6n/\varepsilon}{12/\varepsilon} \right)^{12/\varepsilon} \\ &\leq \frac{12}{\varepsilon} \left( \frac{en}{2} \right)^{12/\varepsilon}. \end{aligned}$$

This means that the size of  $\Pi$  is a polynomial function of  $n$ .

For any set  $S \in \Pi$ , let  $\Gamma_S$  denote the set of all feasible (possibly fractional) routings, in which all requests in  $S$  are routed in the long way; the remaining heavy requests are routed in the short way. From Lemma 15,

$$\min_{S \in \Pi} \min_{x \in \Gamma_S} l(x) \leq L_U^*.$$

As the size of  $\Pi$  is a polynomial function of  $n$  and for each  $S \in \Pi$  a routing in  $\Gamma_S$  with ring load  $\min_{x \in \Gamma_S} l(x)$  can be obtained by solving a linear program in polynomial time, we can find a set  $S \in \Pi$  and a routing  $x \in \Gamma_S$  which has ring load

$$l(x) = \min_{S \in \Pi} \min_{z \in \Gamma_S} l(z).$$

Notice that  $l(x) \leq L_U^*$  and any light request has demand at most  $2\varepsilon L_U^*/3$ . Thus, once  $x$  is obtained, we can transform  $x$  in polynomial time into an unsplit routing  $y$  with  $l(y) < (1 + \varepsilon)L_U^*$ . Therefore, we have the following result:

**Theorem 16.** *For any  $\varepsilon > 0$ , we can find an unsplit routing whose ring load is within  $(1 + \varepsilon)$  of the optimal unsplit ring load.*

A remark should be made here on the differences between our  $(1 + \varepsilon)$ -approximation algorithm and  $(1 + \varepsilon)$ -approximation algorithm given in [5] for the optimal unsplit routing problem in undirected rings. The first difference is that the algorithm in [5] first assumes that the optimal unsplit load is known and later use some standard techniques to bypass such a clairvoyance assumption. Our algorithm does not make such an assumption at all and therefore is simpler. Secondly, the algorithm in [5]

transforms each fractional routing into an unsplit routing and then picks the best unsplit routing. We observe that it is unnecessary. In fact, we can first choose the best fractional routing and then transform *only* the best fractional routing into an unsplit routing just *once*. The resulting routing is still an  $(1 + \varepsilon)$ -approximation, but the time-complexity is reduced by a factor of  $\Theta(n^3/\log n)$ .

## 7. CONCLUSIONS

In this paper, we studied several variants of load balancing in counter rotated directed SONET rings. The optimal fractional routing can be obtained by solving a linear program. The optimal semi-integral routing can be obtained by solving at most three linear programs. The optimal integral routing can be obtained by rounding any optimal parallel semi-integral routing. The optimal unsplit routing is NP-complete. Both the edge-avoidance routing and the short-way routing are 2-approximations. Finally, a polynomial-time approximation algorithm is presented for any fixed  $\varepsilon > 0$  to find an unsplit routing whose ring load is within  $(1 + \varepsilon)$  times the optimum.

## REFERENCES

- [1] D. Clark, Heavy traffic drives networks to IP over SONET, IEEE Comput Dec (1998), 17–20.
- [2] S. Cosares and I. Saniee, An optimization problem related to balancing loads on SONET rings, Telecommun Syst 3 (1994), 165–181.
- [3] M.R. Garey and D.S. Johnson, Computers and intractability: A guide to the theory of NP-completeness, W.H. Freeman, San Francisco, CA, 1979.
- [4] I. Haque, W. Kremer, and K. Raychauduri, “Self-healing rings in a synchronous environment,” SONET/SDH: A sourcebook of synchronous networking, C.A. Siller and M. Shafi, (Editors), IEEE Press, New York, 1996, pp. 131–139.
- [5] S. Khanna, A polynomial-time approximation scheme for the SONET ring loading problem, Bell Lab Tech J 2(2) (1997).
- [6] H. Okamura and P. Seymour, Multicommodity flows in planar graphs, J Combin Theory Ser B 31 (1981), 75–81.
- [7] I. Saniee, Optimal routing in self-healing communications networks, Int Trans Oper Res 3 (1996), 187–195.
- [8] A. Schrijver, P.D. Seymour, and P. Winkler, The ring loading problem, SIAM J Discr Math 11 (1998), 1–14.
- [9] R. Vachani, A. Shulman, P. Kubat, and J. Ward, Multicommodity flows in ring networks, INFORMS J Comput 8 (1996), 235–242.
- [10] P. Wilfong and P. Winkler, Ring routing and wavelength translation, Proc Ninth Annual ACM–SIAM Symp on Discrete Algorithms, 1998, pp. 333–341.