1983

Optimal Routing in the Cube-Connected Cycles Interconnection Network

Daniel A. Reed

Report Number:

81-412

Optimal Routing in the Cube-Connected Cycles Interconnection Network

Daniel A. Reed †

Department of Computer Science
University of North Carolina
Chapel Hill, North Carolina 27514

CSD-TR-412 (Revised)

ABSTRACT

An optimal message routing algorithm for the cube-connected cycles processor interconnection network is described, and the average message path length is derived assuming a uniform message routing distribution. The optimal algorithm is compared to one previously proposed and is shown to have significantly shorter average path length.

Key Words and Phrases:

Parallel processing, message routing, multicomputer networks, performance evaluation

† CSNET address: dar.unc@UDel-Relay
This work was performed at the Department of Computer Sciences, Purdue University.

June 27, 1983

Introduction

Several researchers have recently proposed large networks of computation nodes whose communication paradigm is message passing [1, 3, 4]. Each network node, implemented as one or two VLSI chips, would contain a processing element with some local memory, a communication processor capable of routing messages without delaying the processing element, and a few connections to other network nodes.

Selecting an appropriate interconnection network is a particularly perplexing design problem. Among the proposed interconnections, the binary multidimensional cube has been shown to efficiently support the communication patterns of several important algorithms. Unfortunately, each of the 2^D nodes in such a cube must be connected to D other nodes. As D increases this violates the fanout limitations imposed by the VLSI implementation of the nodes.

However, Preparata and Vuillemin [2] have suggested a variation of the cube, called the cube-connected cycles, that eliminates the fanout problem. As the name suggests, each node at a vertex of a D-dimensional cube is replaced with a ring of D nodes, numbered from 0 through $D-1$. Node addresses then take the form (i,j) where $0 \le i < D$ and $0 \le j < 2^D$. Each node is connected to the two neighboring ring nodes at its vertex and the node with the same ring number i and vertex number given by toggling the i-th bit of the current vertex number j. (See Figure I.) Consequently, each node has three neighbors regardless of the cube dimension.

Although this approach solves the fanout problem, it complicates the algorithm for routing messages from source to destination nodes. In this paper, we analyze a simple routing algorithm proposed by Wittie [4], outline an optimal routing algorithm for the cube-connected cycles interconnection, and discuss the relative message intensities on the communication links of the ordinary cube and the cube-connected cycles.

Definitions and Assumptions

We define a *cube* link as a link connecting two nodes with different vertex addresses and a *ring* link as one connecting two nodes with different ring addresses. Because traversing a cube link moves a message to another node with the same ring address at another vertex, finding a shortest path from a source node to a destination node can be reduced to the following optimization problem.

(1) Consider a ring of D nodes.

(2) Distinguish a source node (s), a destination node (d), and k intermediate nodes ($0 \le k < D-1$).

(3) Find a shortest path from the source node to the destination node that passes through all intermediate nodes.

When analyzing routing algorithms for solving this problem, we shall assume a uniform message routing distribution. That is, all source and destination addresses are presumed to be equally likely. Under this assumption, each ring node will be a distinguished node with probability 0.5 since each bit of the source and destination vertex addresses differ with this probability.

Simple Routing Algorithm

The routing algorithm proposed by Wittie [4] begins at the source node and involves two steps:

(A) Cross ring links in the clockwise direction until all distinguished intermediate nodes have been visited.

(B) Find the shortest path, clockwise or counterclockwise, from the current position to the final destination node.

For step A, a message traverses i ring links if and only if the node l links away from the source were distinguished and the last $D - l - 1$ nodes were not distinguished. Under the uniform routing distribution, the probability of traversing l ring links is just a geometric random variate with value $\frac{1}{2^{D-l}}$. Thus, the average number of ring links traversed in performing step A is

$$\sum_{l=0}^{D-1} \frac{l}{2^{D-l}} = D + \frac{1}{2^{D-1}} - 2 .$$

For step B, the number of links traversed is at most $\left\lfloor \frac{D}{2} \right\rfloor$. On the average

$$\frac{\sum_{l=0}^{D-1} \min \left\{ l, D - l \right\}}{D} = \begin{cases} \dfrac{2\sum_{l=0}^{\left\lfloor \frac{D}{2} \right\rfloor} l - \dfrac{D}{2}}{D} = \dfrac{D}{4} & D \text{ even} \\[2ex] \dfrac{2\sum_{l=0}^{\left\lfloor \frac{D}{2} \right\rfloor} l}{D} = \dfrac{D^2 - 1}{4D} & D \text{ odd} \end{cases}$$

links must be traversed. Hence, the average number of ring links traversed by this algorithm is

$$\frac{5D}{4} + \frac{1}{2^{D-1}} - 2 - \begin{cases} 0 & D \text{ even} \\ \frac{1}{4D} & D \text{ odd.} \end{cases} \quad (1)$$

Optimal Routing Algorithm

Wittie's simple routing algorithm does not use the location of source and destination address differences to reduce the number of ring link traversals. One would expect any algorithm employing this information to perform significantly better. Finkel and Solomon presented an optimal routing algorithm for the lens [1], a shared bus interconnection, that can be applied to the cube-connected cycles interconnection, and it does exhibit shorter mean path length than the simple algorithm.

We begin by observing that the source and destination nodes divide the ring into two arcs. Let A be the longer of these, and let a be its length. Let B be a maximal contiguous sequence of non-distinguished nodes of length $b - 1$ in the interior of A. Finally, let C be a maximal contiguous sequence of non-distinguished nodes of length $c - 1$ outside A. Figure II shows one possible arrangement of arcs for a ring of 11 nodes with distinguished nodes indicated by •. Note that $\left\lceil \frac{D}{2} \right\rceil \le a \le D$, $1 \le b \le a$, and $1 \le c \le D - a$. A shortest path from source to destination is given by the minimum of cases I and II below.

Case I (See Figure III)

Move from the source (s) to one end of B and back to the other end of B. Finally, move to the final destination (d). Inspection shows the length of this path to be $D + a - 2b$.

Case II (See Figure IV)

Move from the source (s) to one end of C and back to the other end of C. Finally, move to the final destination (d). The length of this path is $2D - a - 2c$.

The optimality of this algorithm rests on the observation that any path must include either the links in A or those outside A. Having included these links, one need only visit the distinguished nodes not encountered along this path. Clearly, these are best visited by short excursions from the source or destination node. Thus, the minimum length ring path from source to destination is

$$\min\left\{D + a - 2b \;,\; 2D - a - 2c\right\}.$$

Again, we would like to find the average number of ring link traversals required to reach a destination under the uniform routing assumption. To do this, we must first establish two apparently unrelated lemmas.

Lemma 1

The probability of arc A having length a is $P(a,D)$ where

$$P(a,D) = \begin{cases} \dfrac{1}{D} & D = a \quad or \quad D = 2a \\ \\ \dfrac{2}{D} & otherwise. \end{cases}$$

Proof:

Since all sources and destinations are equally likely and A is, by definition, the longer of the two arcs between the source and destination, there are two complementary positions on the ring, each occurring with probability $\frac{1}{D}$, such that A has length a. The only exceptions occur if source and destination are coincident, or a is half the ring circumference. In these cases, there is only one possible arc of this length. ◻

Lemma 2

The probability of a run of at least m successes in n trials is $R(m,n)$ where

$$R(m,n) = p^m + q\sum_{j=0}^{m-1} p^j R(m,n-j-1)$$

and

p	*probability of a success*
q	*probability of a failure*
$R(m,n) = 0$	$m > n.$

Proof:

Let $P(E)$ denote the probability of the desired run of m successes (i.e., $P(E) = R(m,n)$), and let $P(S)$ denote the probability of an individual success (i.e., $P(S) = p$). Applying the laws of conditional probability, we have

$$P(E) = P(S)P(E|S) + P(\bar{S})P(E|\bar{S})$$
$$= pP(E|S) + qP(E|\bar{S}).$$

Now $P(E|\bar{S})$ is just $R(m,n-1)$. Employing conditional probabilities again, we have

$$P(E|S) = P(S|S)P(E|S^2) + P(\bar{S}|S)P(E|\bar{S}S)$$
$$= pP(E|S^2) + qR(m,n-2)$$

so

$$P(E) = p^2P(E|S^2) + pqR(m,n-2) + qR(m,n-1).$$

By induction, one can show that

$$P(E) = p^iP(E|S^i) + q\sum_{j=0}^{i-1}p^jR(m,n-j-1).$$

Note, however, that $P(E|S^m) = 1$. Hence,

$$P(E) = R(m,n) = p^m + q\sum_{j=0}^{m-1}p^jR(m,n-j-1). \qquad \blacksquare$$

Finally, $Q(m,n) = R(m,n) - R(m+1,n)$ is the probability of a run of *exactly* m successes in n trials. In the context of our discussion, $Q(m,n)$ corresponds to the probability of a contiguous group of non-distinguished ring nodes of length m in an arc of length n. Under the uniform message routing assumption, the probability of an individual success, p, is 0.5. We can now state the following theorem.

Theorem

The average number of ring link traversals required to route a message to its destination using the optimal routing algorithm is

$$\sum_{a=\left\lceil \frac{D}{2} \right\rceil}^{D-1} P(a,D) \sum_{b=1}^{a} Q(b-1,a-1) \sum_{c=1}^{D-a} Q(c-1,D-a-1) \min\left\{ D+a-2b,\ 2D-a-2c \right\}$$

$$+ P(D,D) \sum_{b=1}^{D} Q(b-1,D-1) \min\left\{ 2(D-b),\ D \right\} . \qquad (2)$$

Proof:

Recall that $P(a,D)$ is the probability of arc A having length a, $Q(b-1,a-1)$ is the probability of a group of contiguous non-distinguished nodes of length $b-1$ occurring in the interior of A, and $Q(c-1,D-a-1)$ is the probability of a similar group of non-distinguished nodes occurring exterior to A. Then the sum represents all possible values of a, b, and c weighted by their probability of occurrence multiplied by the minimum path length for those values. □

Although (2) is unwieldy, is seems unlikely that a closed form solution can be found. Fortunately, two related factors make such a solution unnecessary. First, the computational complexity of (2) is only $O(D^3)$ if recurrence values are precalculated. This contrasts with the $O(D2^D)$ operations needed to determine the mean number of ring link traversals by exhaustive enumeration. Second, the number network nodes rises exponentially with D. Thus, we need only consider relatively small values of D to obtain networks with thousands of nodes (e.g., $D = 10$ gives 10,240 nodes).

Comparisons

Figure V shows the average number of ring link traversals required for both the simple routing algorithm and the optimal one. Over the range of D shown, the optimal routing algorithm shows a 20.5 percent reduction in the mean number of ring link traversals. As we shall see, only marginal performance improvements result if a reduction of greater than 25 percent is obtained.

The expected number of link traversals alone fails to capture the performance of an interconnection network. Indeed, a single bus and a completely connected network have the same performance by this metric. Another, perhaps more important, metric of interconnection network performance is the link message intensity, the average number of link traversals required by a message divided by the number of links. For a network such as the cube-connected cycles, there are two message intensities, one for the ring links and one for the cube links.

Under uniform routing, the average number of cube link traversals required by a message is $\frac{D}{2}$ because each of the D bits of the vertex addresses differ with probability 0.5. Since there are $D2^{D-1}$ cube links, the message intensity for these links is $\frac{1}{2^D}$.

For the $D2^D$ ring links, equation (1) shows that the message intensity for the simple routing algorithm is approximately $\frac{1.25}{2^D}$. Since this is greater than the message intensity of the cube links, the ring links will be the dominant factor in communication delays if this routing algorithm is used. By way of contrast, the message intensity for the ring links using the optimal routing algorithm is only slightly greater than $\frac{1}{2^D}$. This implies that the performance of the cube-connected cycles interconnection is near that of an ordinary cube, confirming the cube-connected cycles interconnection as a feasible alternative to an ordinary cube.

Conclusions

We have presented an optimal routing algorithm for the cube-connected cycles interconnection and analyzed its performance. The algorithm is only slightly more complicated than the simple algorithm and significantly reduces the message intensity of the ring links. Because of this, the cube-connected cycles interconnection performance should approach the performance of an ordinary cube interconnection.

References

[1] R. A. Finkel and M. H. Solomon, "The Lens Interconnection Strategy," *IEEE Trans. Comput*, Vol. C-30, pp. 960-965, Dec. 1981.

[2] F. P. Preparata and J. Vuillemin, "The Cube-Connected Cycles: A Versatile Network for Parallel Computation," *Commun. Ass. Comput. Mach*, vol. 24, pp. 300-309, May 1981.

[3] D. A. Reed and H. D. Schwetman, "Cost-Performance Bounds for Multi-Microcomputer Networks," *IEEE Trans. Comput*, Vol. C-32, pp. 83-95, Jan. 1983.

[4] L. D. Wittie, "Communication Structures for Large Multimicrocomputer Systems," *IEEE Trans. Comput*, Vol. C-30, pp. 264-273, Apr. 1981.

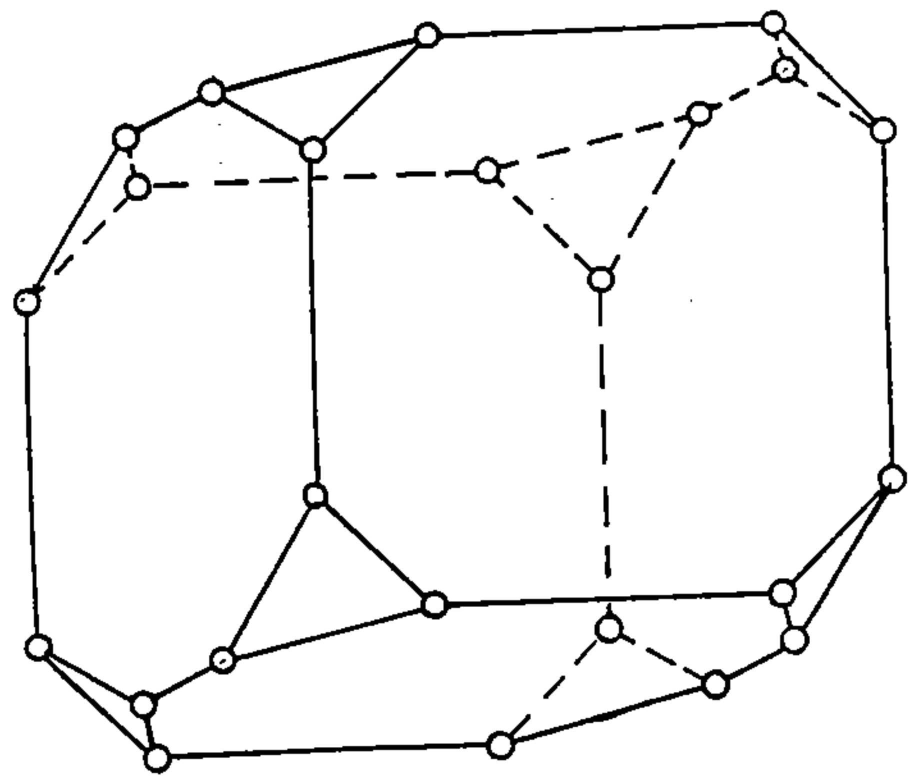

Figure 1

3-dimensional cube-connected cycles

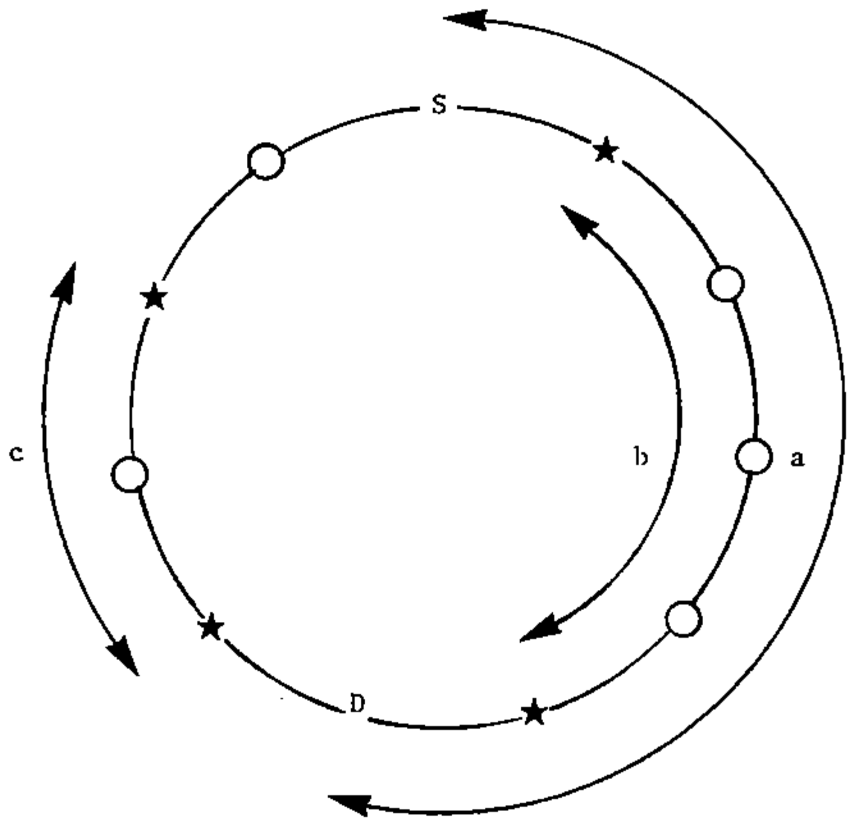

Figure II

Arc lengths for a ring of eleven nodes

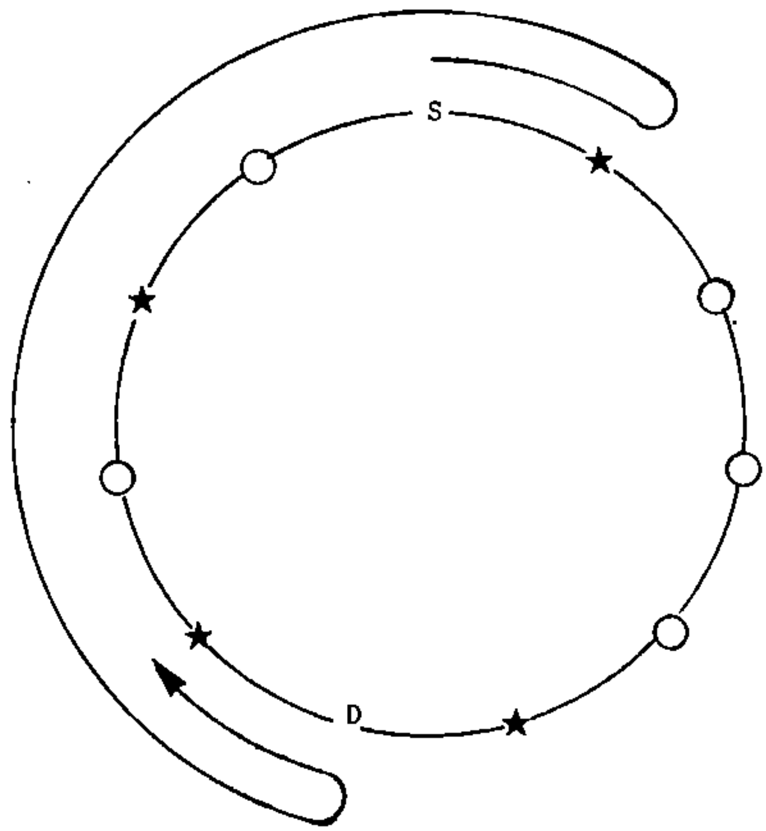

Figure III

Optimal routing algorithm, case 1
Path length $= D + a + 2b = 9$

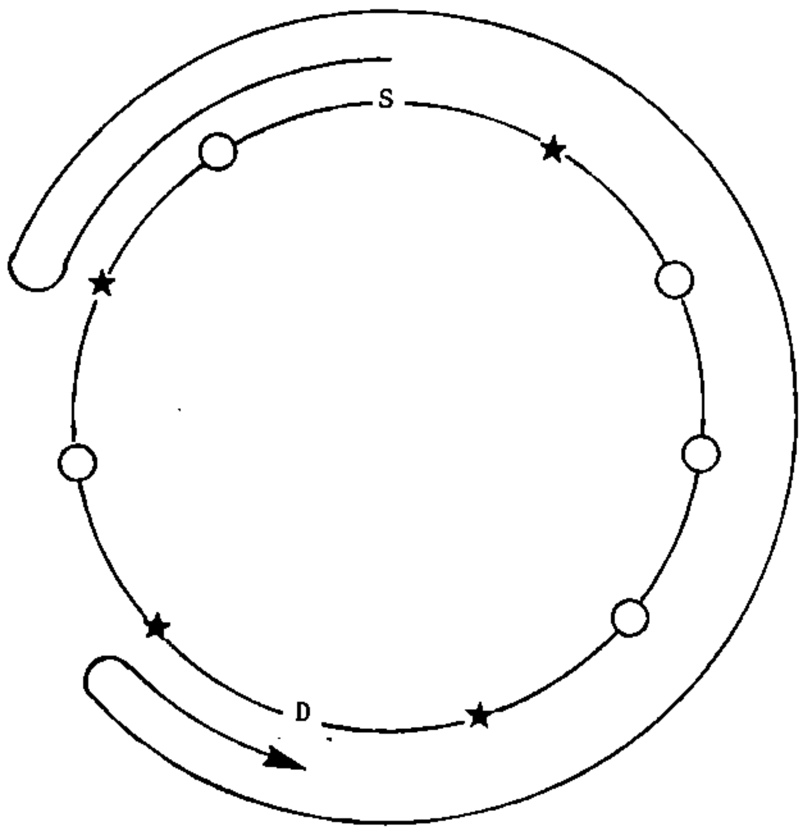

Figure IV

Optimal routing algorithm, case II
Path length $= 2D - a - 2c = 12$

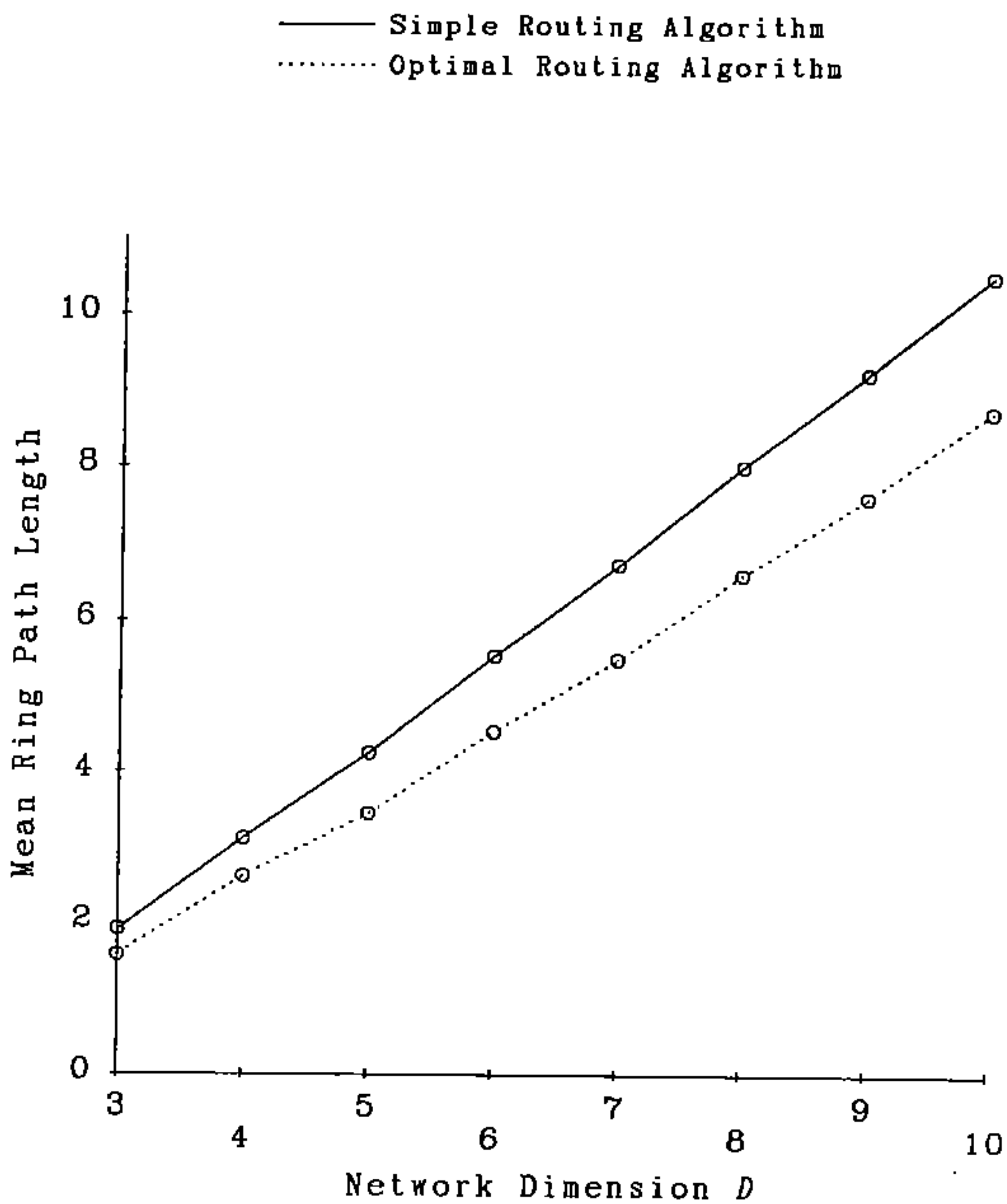

Figure V
Mean ring path length
for the cube-connected cycles