Reprinted from INFORMATION AND COMPUTATION
Vol. 72, No. 3, March 1987
Printed in Belgium

A New Look at Fault-Tolerant Network Routing*

DANNY DOLEV

Hebrew University, Givat Ram, 91904 Jerusalem, Israel

AND

JOSEPH Y. HALPERN, BARBARA SIMONS, AND H. RAYMOND STRONG

IBM Almaden Research Center, San Jose, California 95120

We model a communication network as a graph in which a processor is a node and a communication link is an edge. A routing for such a network is a fixed path, or route, between each pair of nodes. Given a network with a predefined routing, we study the effects of faulty components on the routing. Of particular interest is the number of routes along which a message must travel between any two non-faulty nodes. This problem is analyzed for specific families of graphs and for classes of routings. We also give some bounds for general versions of the problem. Finally, we conclude with one of the most important contributions of this paper, a list of interesting and apparently difficult open problems. © 1987 Academic Press, Inc.

1. INTRODUCTION

We consider the problem of obtaining efficient, reliable, fault-tolerant routings in a network. As usual, a network is modeled as a graph, with nodes representing processors and edges representing communication links. A *routing* is a partial function that assigns to pairs of nodes in the network a fixed path between them. We assume that the network communication protocol has no information about the topology of the network, and thus all communication between nodes must go on this fixed routing (and only nodes that have a route between them can communicate directly).

In local area networks, the time required to send a message along a route is often dominated by the message processing time at either end; intermediate nodes on a fixed route relay messages without doing any extensive processing. Metaphorically speaking, the intermediate nodes pass on the message without having to open its envelope. Thus, to a first approximation, the time required to send a message along a fixed route is independent of the length of the route.

* A preliminary version of this paper appeared in the "Proceedings, Sixteenth Annual ACM Symposium on Theory of Computing, Washington, D.C., 1984.

0890-5401/87 $3.00

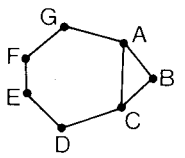

FIGURE 1

Consider the network shown in Fig. 1. Suppose we choose a *minimal length routing* on this network; i.e., one for which the route between any pair of nodes is a minimal length path between them. Where they exist, we break ties by always taking the route that goes through the edge CD.

If in this example the edge CD becomes faulty, then many routes become unavailable. Figure 2 is the *surviving route graph*, where two nodes are joined by an edge exactly if the route between them is still up (i.e., it did not go through the edge CD).

Suppose processor C wants to broadcast a message to all processors. Since C can send messages only along the fixed routes, the message will not reach D, E, or F. If G rebroadcasts the message, it will reach E and F, but not D, since the route from G to D is also down. One more rebroadcast by E or F is necessary to ensure that D gets the message.

Note that the worst case number of rebroadcasts needed to ensure that all processors get a message will be the diameter of the induced graph of Fig. 2. This observation generalizes. Given a set of faults, the diameter of the surviving route graph induced by these faults is exactly the number of rebroadcasts required to ensure that all processors get a message. In general, which nodes and edges in a graph will become faulty is not known in advance. If we can calculate an upper bound d on the diameter of the surviving route graph in the presence of f faults for some fixed f, then by rebroadcasting a message d times we are guaranteed that all processors get the message (provided that indeed there are no more than f faults). Thus such a bound can be used to determine the number of phases required for each round of certain distributed protocols (such as the Byzantine agreement protocols of (Dolev and Strong, 1983; Strong and Dolev, 1982). Given the assumption that the time to send a message along a fixed route is independent of its length, the diameter of the surviving route graph also

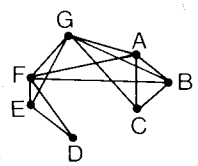

FIGURE 2

gives a good estimate of the time required to complete a broadcast in the presence of faults.

These observations motivate the problem we consider in this paper: analyzing the number of routes along which a message must travel between any two nonfaulty processors. In particular, we want to find good fault-tolerant routings, i.e., routings that keep the diameter of the surviving route graph small for any set of faults of a given cardinality. Of course, the analysis depends on both the types of faults that are considered (node faults, edge faults, or both) and details of the topology of the original network (for example, its connectivity). This problem has given rise to many interesting questions in graph theory, some of them still open.

Roughly speaking, the problem can be formalized as follows (detailed definitions are given in Section 2). Given a graph G, a routing ρ, and a set of faults F, we consider the *surviving route graph* $R(G, \rho)/F$ with the same nodes as $G - F$, and an edge joining two nodes whenever the route between them avoids F. We want to choose a routing ρ such that the diameter of $R(G, \rho)/F$ is minimized for any set of faults F of a given cardinality.

We first note that minimal length routings are not always optimal. Consider the spoke graph shown in Fig. 3. In this case, for any points on the circumference that are not joined by an edge, there exists a minimal length route that goes through the center node. If, however, the center node fails, then it is easy to see that with a minimal length routing the diameter of the surviving route graph grows to $(n-1)/2$ (where n is the total number of nodes). The problem with a minimal length routing in this case is that the center node is overworked. Consider instead the routing ρ on S_n (the spoke graph with n nodes) in which the route between two nodes on the circumference is a minimal length path around the circumference (so that, for example, the route from A to D in Fig. 3 would be $ABCD$, rather than AMD). In this case, the diameter of $R(S_n, \rho)/F$ is easily seen to be $\leqslant 2$ if $|F| \leqslant 2$.

This leads us to ask if we can always find good routings. We show (Theorem 3) that for any $(t+1)$ node-connected graph G, we can efficiently find a routing ρ such that the diameter of $R(G, \rho)/F$ is no greater than $\max(2t, 4)$ if $|F| \leqslant t$.

Although minimal length routings are not always optimal, they are

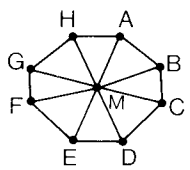

FIGURE 3

useful and easy to generate. Indeed, a common routing algorithm (used, for example, in the Highly Available Systems project at IBM (Aghili *et al.*, 1983) produces random minimal length routings. Thus, it becomes important to find networks for which *all* minimal length routings are fault tolerant.

As an example, consider K_n, the completely connected network on n nodes. If ρ is the unique minimal length routing on K_n, then the diameter of $R(K_n, \rho)/F$ is 2 if $|F| \leqslant n-2$. (To see this, suppose F is fixed and that a and b are any two nonfaulty nodes in K_n. Then either the link between a and b is nonfaulty, or, since $|F| \leqslant n-2$, there must exist a nonfaulty node c such that both the link between a and c and the link between c and b are nonfaulty.)

Unfortunately, because of high fan-in and fan-out, completely connected networks are often impractical. As in several other contexts (e.g., Valiant, 1982) networks laid out as an n-dimensional cube (C_n) achieve surprisingly good results. In Theorem 1 we show that for *any* minimal length routing ρ on C_n and *any* set of faults F with $|F| < n-1$, the diameter of $R(C_n, \rho)/F \leqslant 3$, independent of n. The proof of Theorem 1 is short but nontrivial. The result generalizes to n-dimensional rectangular grids and is easily seen to be optimal.

We also show (Theorem 2) that there exists a minimal length routing λ_n on C_n such that $R(C_n, \lambda_n)/F \leqslant 2$ if $|F| < n$. This in fact is a corollary to a more general result of (Broder *et al.*, 1984) (although the proof for this special case is much simpler than that of op. cit.).

We can also obtain bounds on the diameter of the surviving route graph for arbitrary graphs, provided minimal length routings are used. If we restrict our attention to edge faults, then the diameter of the surviving route graph grows at worse linearly with the number of faults. In an earlier version of this paper (Dolev *et al.*, 1984), we showed that if F consists only of edge faults, G/F (i.e., G with all the elements of F removed) is connected, and ρ is any minimal length routing on G, then the diameter of $R(G, \rho)/F$ is $\leqslant 3|F| + 1$. We conjectured that this result could be improved to $2|F| + 1$, a conjecture that was recently proved by (Feldman, 1985). This result is optimal, since we can also exhibit graphs where this bound is attained. The spoke example shows that with node faults things may get much worse. Even a single node fault can force the diameter of the surviving route graph to grow to $O(n)$. However, a closer look at this example suggests that the diameter can only grow in this way if there are nodes of high degree. In (Dolev *et al.*, 1984) we substantiated this intuition by showing that if F consists only of node faults, G/F is connected, and ρ is a minimal length routing on G, then the diameter of $R(G, \rho)/F$ is bounded above by $\|F\|$, the sum of the degrees of the faulty nodes in F. We conjectured that this bound could be improved to $\|F\| - |F| + 1$; this conjecture

was also proved in (Feldman, 1985). We can also exhibit graphs to show that the latter bound is tight.

Chung and Garey (1983) were able to obtain analogous results for surviving graphs G/F (as opposed to surviving route graphs). This can be viewed as dealing with the important special case where the routing just consists of the edges in the original graph. (More precisely, $G/F = R(G, \rho)/F$, where ρ is that routing such that $\rho(x, y) = xy$ if (x, y) is an edge in the graph, and is undefined otherwise.) Again the spoke example shows that one node fault can cause the diameter of the surviving graph to be $O(n)$. However, Chung and Garey show that if F consists of only edge faults and G/F is connected, then the diameter of G/F is $\leqslant (1 + |F|)$(the diameter of $G) + O(|F|)$. In the case of node faults, they compute a bound on the diameter of G/F in terms of the degree of the faulty nodes. They also give examples in both cases where their bounds are essentially achieved.

The rest of the paper is organized as follows. In Section 2 the necessary definitions are given. Section 3 contains the results on the n-dimensional cube. In Section 4 good routings for general graphs are discussed. Section 5 gives general results for minimal length routings. There are still many open questions in this area; we list a few of them in Section 6.

2. Surviving Route Graphs

Unless otherwise noted, we deal with an undirected graph $G = (V, E)$ that corresponds to a communication network. A *node routing* ρ on V is a partial function $\rho: V \times V \to V^*$ such that $\rho(x, y)$, if it is defined, is a sequence of nodes in V starting with x and ending with y; i.e., a word of the form xuy with $u \in V^*$. A node routing ρ on V is a *routing* on $G = (V, E)$ if $\rho(x, y)$ (when defined) corresponds to a simple path (one with no loops) in G from x to y; i.e., every consecutive pair of nodes in $\rho(x, y)$ is an edge in E. A routing ρ on V determines an edge-labelled, directed *route graph* $R = (V, \mathrm{dom}(\rho))$, where two nodes x and y are joined by an edge exactly if $\rho(x, y)$ is defined. In this case the edge is labelled by $\rho(x, y)$. If ρ is a routing on G, we use the notation $R(G, \rho)$ for the route graph determined by ρ. (We occasionally omit the G and ρ if they are clear from context.)

A routing ρ is a *partial routing* if $\rho(x, y)$ is undefined for some nodes $x \neq y$; otherwise ρ is a *total routing*. Note that if ρ is a total routing then $R(G, \rho)$ is a complete graph on the nodes of V.

Let F be a set of nodes and edges called the set of *faults*. F can be partitioned into the set of node faults, F_v, and the set of edge faults, F_E. We define V/F to be $V - F_V$, E/F to be $E - F_E - \{(a, b) \in E \mid a \in F_V \text{ or } b \in F_V\}$, and $G/F = (V/F, E/F)$. G/F is called the *surviving* graph.

An object (path, subgraph, etc.) *avoids* F if no element of F is contained in that object. Thus, a path avoids F if no node or edge on the path is in F. A routing avoids F if each of its routes does. An edge of a route graph avoids F if the sequence (path) which is its label does.

For a given set of faults F, let ρ/F be the subrouting of ρ consisting of those routes that avoid F; i.e., $(\rho/F)(x, y) = \rho(x, y)$ if $\rho(x, y)$ avoids F, otherwise $(\rho/F)(x, y)$ is undefined. If $R = (V, \mathrm{dom}(\rho))$ is a route graph and F is a set of faults, the *surviving route graph* is $R/F = (V/F, \mathrm{dom}(\rho/F))$. Thus, two nodes are joined by an edge in the surviving route graph exactly if the route between them avoids F.

We now briefly review some standard definitions from graph theory. We refer the reader to (Berge, 1976) for more details. A graph G is *connected* if there exists a path in G between any pair of nodes in G; a graph G is $(t+1)$ *node connected* if there are $t+1$ node disjoint paths between any pair of nodes in G. Given nodes u and v in G, the *distance* between u and v in G, denoted $d_G(u, v)$, is the shortest path in G between u and v. The *diameter* of G, written $\mathrm{DIAM}(G)$, is the maximum of $d_G(u, v)$ for every pair of nodes u, v in G.

3. The Diameter of the Surviving Route Cube

Let $C_n = (V_n, E_n)$ be the n-dimensional cube. We represent nodes of C_n as words of length n on the alphabet $\{0, 1\}$. If x is a node, its ith coordinate is denoted x_i. Edges exist only between nodes that differ on exactly one coordinate. Thus we represent edges as words of length n on the alphabet $\{0, 1, *\}$ with exactly one occurrence of * in the coordinate where the two nodes joined by this edge differ.

Networks in the form of n-dimensional cubes display surprisingly good performance. Theorem 1 states that the surviving route graph that results from any minimal length routing on C_n and fewer than n faults has diameter at most 3. Theorem 2 defines a specific minimal length routing and asserts that the diameter of the n-dimensional cube with this routing is 2.

THEOREM 1. *Let ρ be a minimal length routing on C_n. If $|F| < n$, then* $\mathrm{DIAM}(R(C_n, \rho)/F) \leqslant 3$.

THEOREM 2. *Let $\lambda_n(x, y)$ be the (minimal length) routing on the n-dimensional cube that proceeds from x to y by moving along the coordinates on which they differ one at a time from left to right. Then if $|F| < n$,* $\mathrm{DIAM}(R(C_n, \lambda_n)/F) \leqslant 2$.

For example, $\lambda_3(011, 110) = (011, 111, 110)$ and $\lambda_3(110, 011) = (110, 010, 011)$. Note that $\lambda_n(x, y) \neq \lambda_n(y, x)$, in general.

We first develop some machinery to prove these theorems. Define the *weight* of a node or an edge to be the sum of its coordinates where * carries the value $\frac{1}{2}$. Let $|x|$ denote the weight of x. Thus $|11101| = 4$ and $|1*101| = 3.5$. By dropping the ith coordinate, any n-dimensional object can be *projected* along the ith coordinate onto an $(n-1)$-dimensional object. Let P_i be the operator for projecting along the ith coordinate. Note that an edge may project to a node. Thus $P_2(11101) = 1101 = P_2(1*101)$. We write $x \leqslant y$ when $\leqslant$ holds on each coordinate. We write $x < y$ when $x \leqslant y$ and $<$ hold on some coordinate. We say x and y are *maximally far apart* when $\neq$ holds on each coordinate. If x and y are nodes, let $C(x, y)$ be the subgraph consisting of nodes and edges z satisfying the condition, if $x_i = y_i$ then $z_i = y_i$. We call $C(x, y)$ the *subcube generated by x and y*. Informally it consists of the graph induced by all nodes in minimal length paths between x and y.

We define a pair of nodes x and y to be *safe with respect to a set of faults F* iff every minimal length path from x to y avoids F. A sequence of nodes $x_1, ..., x_k$ is safe with respect to F if each consecutive pair of nodes in the sequence is safe with respect to F.

LEMMA 1. *$C(x, y)$ avoids F iff the sequence x, y is safe with respect to F.*

Proof. No minimal length path from x to y can leave $C(x, y)$. ∎

Proof of Theorem 1. By Lemma 1, it follows that if x, y is safe with respect to F, then there will be an edge from x to y in $R(C_n, \rho)/F$ for *every* minimal length routing ρ. Lemma 2 below says that if $|F| < n$, then for any pair of nodes x, y in C_n/F, there are nodes u, v such that x, u, v, y is safe with respect to F. This means that x, u, v, y forms a path of length 3 in $R(C_n, \rho)/F$ for every minimal length routing ρ. Thus Theorem 1 follows immediately from Lemma 2, which we now state and prove.

LEMMA 2. *If $|F| < n$, then for any pair of nodes x and y in C_n/F there are nodes u and v such that the sequence x, u, v, y is safe with respect to F.*

Proof. We proceed by induction on n, carrying along the extra induction hypothesis that if $n > 1$ and if nodes x and y are maximally far apart, then nodes u and v, with $x \neq u$ and $u \neq v$, can be chosen such that x, u, v, y is safe with respect to F, u is in $C(x, v)$ and v is in $C(u, y)$. Note that if $x = 0^n$ and $y = 1^n$, then the last condition is equivalent to $x < u < v \leqslant y$.

The arguments for $n = 1$ and $n = 2$ are straightforward and left to the reader. Assume the induction hypothesis for dimension $n-1$ with $n > 2$. Let x and y be nodes in C_n/F. There are two cases.

Case (a). The nodes x and y have the same value on some coordinate. Without loss of generality $x_1 = y_1 = 1$. If every element of F has a 1 in its first coordinate, then the sequence x, $0P_1(x)$, $0P_1(y)$, y is safe. Otherwise, the safe sequence can be constructed entirely in $C(10^{n-1}, 1^n)$ (the subgraph consisting of the nodes and edges with a 1 in the first coordinate) by the induction hypothesis, since at least one element of F is avoided by this subgraph.

Case (b). The nodes x and y are maximally far apart. Without loss of generality $x = 0^n$ and $y = 1^n$. Case (b) has two subcases.

Case (b1). There is an i and an element f of F such that $P_i(f)$ is in $\{0^{n-1}, 1^{n-1}\}$. Without loss of generality $i = 1$. Let $F' = P_1(F) - \{0^{n-1}, 1^{n-1}\}$. Then $|F'| < n-1$. Thus, by the induction hypothesis there is a sequence $0^{n-1} < u < v \leqslant 1^{n-1}$ that is safe with respect to F'. Suppose $v < 1^{n-1}$. Since $C(0^{n-1}, u)$ (resp. $C(u, v)$, $C(v, 1^{n-1})$) avoids F' by Lemma 1, it is easy to check that $C(0^n, 0u)$ (resp. $C(0u, 1v)$, $C(1v, 1^n)$) avoids F. Thus $0^n < 0u < 1v < 1^n$ is safe with respect to F. And if $v = 1^{n-1}$, then it is again easy to see that $0^n < 0u < 1u < 1^n$ is safe with respect to F.

Case (b2). For each i, $P_i(F)$ does not include either 0^{n-1} or 1^{n-1}. Let f be a minimal weight element of F. Without loss of generality assume $f_1 = 1$ so that $P_1(f)$ has minimal weight in $P_1(F)$. Let $F' = P_1(F - \{f\})$. If F' is empty, then (since the projection of a nonempty set is nonempty) $F = \{f\}$. Consequently, since $f_1 = 1$, $0^n < 01^{n-1} < 1^n$ is safe with respect to F. Suppose that F' is not empty. Then $|F'| < n-1$, so by the induction hypothesis there exists at least one sequence safe with respect to F' of the form $0^{n-1} < a < b \leqslant 1^{n-1}$. Among all such sequences there must be one $0^{n-1} < u < v \leqslant 1^{n-1}$ with $|u|$ maximal. We claim that $0^n < 0u < 0v < 1^n$ is safe with respect to F. It is clearly safe with respect to $F - \{f\}$: since $C(0^{n-1}, u)$ (resp. $C(u, v)$, $C(v, 1^{n-1})$) avoids F', then $C(0^n, 0u)$ (resp. $C(0u, 0v)$, $C(0v, 1^n)$) must avoid $F - \{f\}$. Thus it suffices to show that 0^n, $0u$, $0v$, 1^n is safe with respect to $\{f\}$. Since $f_1 = 1$, clearly 0^n, $0u$, $0v$ is safe with respect to $\{f\}$. Thus it suffices to show that $f \notin C(0v, 1^n)$. But if $f \in C(0v, 1^n)$, we must have $|P_1(f)| \geqslant |v|$ (and, in particular, we have that $v < 1^{n-1}$). Since f was chosen with minimal weight and $f_1 = 1$, it follows that $|P_1(f')| \geqslant |P_1(f)| \geqslant |v|$ for all $f' \in F$. Thus $C(0^{n-1}, v)$ avoids F', so $0^{n-1} < v < 1^{n-1}$ ($\leqslant 1^{n-1}$) must be safe with respect to F'. Since $|v| > |u|$, this contradicts the choice of u. (Recall we chose u with maximal weight.) ∎

Proof of Theorem 2. We proceed by induction on n. The case $n = 1$ is trivial. For $n > 1$ there are two cases.

Case (a). The nodes x and y agree on coordinate i. Without loss of generality $x_i = y_i = 1$. If every element of F has 1 in the ith coordinate, then $x,\ x_1 \cdots x_{i-1} 0 y_{i+1} \cdots y_n,\ y$ is a path in $R(C_n, \lambda_n)/F$. Otherwise, let $F' = P_i(\{f \in F | f_i = 1\})$. Since $|F'| < n-1$, we can apply our induction hypothesis to $P_i(C_n)$. Thus, there is a path of length one or two from $P_i(x)$ to $P_i(y)$ in $R(P_i(C_n), \lambda_{n-1})/F'$. If the path is of length one, then (x, y) is an edge in $R(C_n, \lambda_n)/F$, since all faults not in F' have either 0 or * in the ith coordinate. And if $P_i(x),\ u,\ P_i(y)$ is a path of length two in $R(P_i(C_n), \lambda_{n-1})/F'$, then it is easy to see that $x,\ u_1 \cdots u_{i-1}\, 1 u_{i+1} \cdots u_{n-1},\ y$ is a path in $R(C_n, \lambda_n)/F$.

Case (b). The nodes x and y are maximally far apart. Without loss of generality, $x = 0^n$ and $y = 1^n$. The paths in C_n formed by concatenating $\lambda_n(0^n, 0^i 1^{n-i})$ and $\lambda_n(0^i 1^{n-i}, 1^n)$ for $1 \leqslant i \leqslant n$ are node disjoint so one of them must avoid F because $|F| < n$. ▮

Remarks. 1. We have shown that when $|F| < n$ and ρ is a minimal length routing on C_n, the diameter of $R(C_n, \rho)/F$ is no greater than 3. However it does not require $|F| = n-1$ to force the diameter to be 3. If we choose ρ so that $\rho(0^n, 1x)$ always goes through 10^{n-1} and $\rho(0y, 1^n)$, $y \neq 0^{n-1}$, always goes through 01^{n-1}, and choose $F = \{10^{n-1}, 01^{n-1}\}$, it is easy to check that the diameter of $R(C_n, \rho)/F$ is 3. A similar example can be obtained by placing * in the first coordinates of either or both elements of F.

2. We call a routing *bidirectional* if the route from x to y is the same as the route from y to x (i.e., $\rho(x, y) = \rho(y, x)$) for all x and y; otherwise, it is called *unidirectional.* We have allowed routings that are not bidirectional. Theorem 1 clearly still holds if we restrict to bidirectional routings, but there is no bidirectional analog of Theorem 2. To see this, consider any minimal length bidirectional routing ρ on the square C_2. (There are not very many.) Note that $\rho(00, 11)$ and $\rho(01, 10)$, the routes to opposite corners of the square, must have an edge in common. If F consists of this single faulty edge, then the distance between its endpoints in $R(C_2, \rho)/F$ must be 3. For $n \geqslant 3$, it is still an open question if there exists a bidirectional analog of Theorem 2. It would also be interesting to know whether there is a bidirectional analog to Theorem 2 if F consists only of node faults. (Note that the counterexample given above for C_2 does not hold for node faults.) Again this remains an open question.

3. For any pair of nodes x, y in C_n, we can find n midpoints $z_1, \ldots, z_n$ with $z_1 = y$ such that the n routes from x to y formed by concatenating $\lambda_n(x, z_i)$ and $\lambda_n(z_i, y)$, $i = 1, \ldots, n$, are node disjoint. A proof of the existence of these midpoints may be obtained by carrying it along as an induction hypothesis in the proof of Theorem 2. These node disjoint routes can be

useful in certain applications. For example, if processor x wants to guarantee that a message gets through to y quickly, it computes $z_1,..., z_n$ and sends the message to $z_1,..., z_n$ with instructions to forward it to y. One message must get through so long as $|F| < n$.

4. Theorems 1 and 2 also hold for any n-dimensional rectangular grid (i.e., a product of n intervals of the form $I_1 \times \cdots \times I_n$, where I_j is of the form $\{0,..., k_j\}$); the techniques of the proof generalize immediately.

4. Routings in a General Network

As we showed in the Introduction, if S_n is a spoke graph with n nodes and ρ is a minimal length routing on S_n, then the diameter of $R(S_n, \rho)/F$ can be $O(n)$, even if F consists of a single node. However, there does exist a non-minimal length routing on the spoke for which the diameter of the surviving route graph is 2 as long as $|F| \leqslant 2$. In this section we show that this result generalizes.

Theorem 3. *If G is $t+1$ node connected, then there is a bidirectional routing ρ such that if $|F| \leqslant t$, then* $\mathrm{DIAM}(R(G, \rho)/F) \leqslant \max(2t, 4)$.

Proof. In order to prove the theorem, we will first need the following lemma.

Lemma 3. *Let $G = (V, E)$ be $t+1$ node connected but not $t+2$ node connected, with $|V| \geqslant t+3$. Then there exists a set of nodes $M \subset V$ with $|M| = t+1$ such that the removal of the nodes in M and all of their adjacent edges partitions G into non-empty disconnected subgraphs, G_1, G_2,..., G_k, with $k \geqslant 2$. Moreover, if $x \in G_i$, $i = 1, 2,..., k$, then there exists $t+1$ node disjoint paths in G_i from x to the nodes in M. If $(x, m) \in E$ for some $m \in M$, we can take xm to be the path from x to m.*

Proof of Lemma 3. The fact that we can find M follows immediately from the fact that G is $t+1$ node connected. Without loss of generality, let $x \in G_1$ and choose some $y \in G_2$. Then by the definition of connectivity, there exist $t+1$ node disjoint paths from x to y in G. Since $|M| = t+1$, and the removal of the nodes in M and all of their adjacent edges leaves x and y in disjoint subgraphs, each of these paths must include exactly one node of M, with the path from x to each such node staying completely in G_1. If $(x, m) \in E$ for some $m \in M$ and if the path from x to m in G_1 which is obtained by the above construction is not xm, then replacing that path with xm does not contradict the node disjoint requirement for the paths from x to M. ∎

Returning now to the proof of Theorem 3, given G, we can assume without loss of generality that G is not $t+2$ node connected (otherwise we find a routing on G', which is the result of removing enough edges from G so that it is not $t+2$ node connected). We must have $|V| \geq t+2$ (otherwise G could not be $t+1$ node connected). If $|V| = t+2$, then G is completely connected, and we just take $\rho(x, y)$ to be the edge xy in this case. It is easy to see that $\text{DIAM}(R(G, \rho)/F) \leq 2$ for any set of faults F with $|F| \leq t+1$ in this case. If $|V| \geq t+3$, we choose M and node disjoint paths from each node $x \notin M$ to each node $m \in M$ as in Lemma 3. We now define a partial routing ρ on G by two rules:

1. If $(u, v) \in E$, then $\rho(u, v) = uv$, i.e., the route from u to v is the edge between them.

2. If $x \notin M$ and $m \in M$, then $\rho(x, m)$ is the path described above.

We note that by using standard techniques from network flow (Even, 1979) such a routing can even be found efficiently, in time $O(|V|^{1/2}\,|E|^2)$.

Rule 1 guarantees that if $|F| \leq 1$, then $R(G, \rho)/F$ is connected and $\text{DIAM}(R(G, \rho)/F) \leq \text{DIAM}(G/F)$. Note that although $\text{DIAM}(G/F)$ could be $O(|V|)$, Theorem 3 gives a bound on $\text{DIAM}(R(G, \rho)/F)$ which is independent of $|V|$.

If $f \in F$ is either a faulty node in G_i (resp. M) or a faulty edge with both endpoints in G_i (resp. M), then f is said to be in G_i (resp. M). If $f \in F$ is a faulty edge which has one end point in M and the other in G_i, then f is said to be in G_i. Let F_i be the set of faults in G_i, $i = 1,\ldots, k$, and F_M be the set of faults in M. Note $|F_1| + \cdots + |F_k| + |F_M| \leq t$.

We now complete the proof that $\text{DIAM}(R(G, \rho)/F) \leq \max(2t, 4)$ by a case analysis.

Case 1. For some $i \in \{1, 2,\ldots, k\}$, $|F_i| = 0$. Without loss of generality, assume that $|F_1| = 0$. Since G_1 is not empty, there exists a node $z \in G_1$ such that there is an edge in R/F from z to every non-faulty node in M. Therefore, there exists a path of length 2 in R/F between any two non-faulty nodes of M via z. Any $x \notin M$ must be adjacent in R/F to some non-faulty $m \in M$ since $|F| < |M|$. This immediately gives a bound of 4 between any two nodes which are neither in M nor in G_1.

Case 2. $|F_i| \neq 0$, for all $i \in \{1,\ldots, k\}$. Let $P = x_0 \cdots x_h$ be some minimum length path in R/F between $x = x_0$ and $y = x_h$. We bound the length of P by counting nodes in M which either appear on P or are adjacent to internal nodes of P. Thus, for $x_i \in P$, let $(x_i) = \{\text{nonfaulty nodes in } M \text{ to which } x_i \text{ has an edge in } R/F\} \cup (\{x_i\} \cap M)$.

Let x_i be a node of P which is not in M, and assume that $x_i \in G_j$. There are paths in R/F from x_i to at least $t+1-(|F_j| + |F_M|)$ non-faulty nodes of M. Since $k \geq 2$ by hypothesis, $|F_1|$, $|F_2| \geq 1$, and

$|F_1| + \cdots + |F_k| + |F_M| = t$, we must have $(|F_j| + |F_M|) \leqslant t-1$, and so $|(x_i)| \geqslant 2$. Let $P_{i..j}$ be the partial path $x_i x_{i+1} \cdots x_j$. Let $S(P_{i..j}) = (x_i) \cup \cdots \cup (x_j)$. We prove the $2t$ bound by showing that $|S(P_{0..i})| \geqslant \lceil i/2 \rceil + 1$ by induction on i.

Since $|(x_0)| \geqslant 2$, the claim holds for $i = 0, 1, 2$. Assume the claim holds up to $i-1$ for $i > 2$. The bound is obtained by the following counting argument. There are two cases, $x_i \in M$ and $x_i \notin M$. If $x_i \in M$, then $x_i \notin (x_j)$ for $j \leqslant i-2$, for otherwise $P_{0..j} x_i P_{(i+1)..h}$ is a shorter path from x to y than is P. Thus,

$$|S(P_{0..i})| \geqslant |S(P_{0..i-2})| + 1 \geqslant \lceil (i-2)/2 \rceil + 2 = \lceil i/2 \rceil + 1.$$

If $x_i \notin M$, then $(x_i) \cap (x_j) = \varnothing$ for $j \leqslant i-3$. Otherwise, the existence of some m with $m \in (x_i) \cap (x_j)$ implies that $P_{0..j} m P_{i+1..h}$ is shorter than P. Since for $x_i \notin M$ we have $|(x_i)| \geqslant 2$, then

$$|S(P_{0..i})| \geqslant |S(P_{0..i-3})| + 2 \geqslant \lceil (i-3)/2 \rceil + 3 \geqslant \lceil i/2 \rceil + 1.$$

Since $P = x_0 x_1 \cdots x_h$, it follows that $|S(P)| \geqslant \lceil h/2 \rceil + 1$. Since $|M| \leqslant t+1$, we must have $\lceil h/2 \rceil \leqslant t$. Consequently, $h \leqslant 2t$ and $|P| \leqslant 2t$. ∎

5. Missing Nodes and Missing Links

In this section we return to minimal length routings and obtain bounds for the diameter of a surviving route graph in terms of the number of faulty edges and the degrees of the faulty nodes. We first consider the case where there are only edge faults. In an earlier version of this paper, we showed that if F consists only of edge faults and G/F is connected, then for any minimal length routing ρ, the diameter of $R(G, \rho)/F$ is $\leqslant 3|F| + 1$. We conjectured that this bound could be improved to $2|F| + 1$, a result which was recently proved by Feldman (1985).

Theorem 4 (Feldman, 1985). *If F consists only of edges, G/F is connected, and ρ is any minimal length routing of G, then* $\mathrm{DIAM}(R(G, \rho)/F) \leqslant 2|F| + 1$.

This result is essentially optimal, as the following theorem shows.

Theorem 5. *For each t there is a graph G_t, a minimal length routing ρ_t of G_t, and a set F_t of t edges that does not disconnect G_t such that* $\mathrm{DIAM}(R(G_t, \rho_t)/F_t) = 2t + 1$.

Proof. The required graph G_t is obtained by the obvious generalization from the graph G_4 shown in Fig. 4, where the edges marked with an $\times$

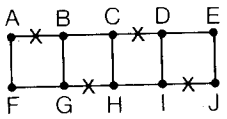

FIGURE 4

through them are in F, and ρ goes through a faulty edge whenever possible (for example, $\rho(A, G) = ABG$). ▮

The spoke example of the Introduction shows that we cannot expect such good behavior from node faults, since even one node fault f in a graph G can cause the diameter of $R(G, \rho)/\{f\}$ to be $O(|V|)$, for every minimal length routing ρ. But this bad behavior can only come about if the node f has high out-degree.

DEFINITION. For a node a, define $\|a\|$ to be the degree of a; i.e., the number of edges with endpoint a. For an edge e, define $\|e\| = 2$. Finally, define $\|F\| = \sum_{f \in F} \|f\|$.

In (Dolev *et al.*, 1984) we showed that if F consists only of node faults and G/F is connected, then the diameter of $R(G, \rho)/F \leqslant \|F\|$ for every minimal length routing ρ. We conjectured that in fact $R(G, \rho)/F \leqslant \|F\| - |F| + 1$. This conjecture was also proved by Feldman.

THEOREM 6 (Feldman, 1985). *If F consists only of nodes, G/F is connected, and ρ is any minimal length routing of G, then* $\mathrm{DIAM}(R(G, \rho)/F) \leqslant \|F\| - |F| + 1$.

Again this result is essentially optimal, since we have

THEOREM 7. *For all $d_1, \ldots, d_k$, there exists a graph G, a minimal length routing ρ on G, and a set of node faults $F = \{f_1, \ldots, f_k\}$ which does not disconnect G such that the degree of f_i is d_i, $i = 1, \ldots, k$, and* $\mathrm{DIAM}(R(G, \rho)/F) = \|F\| - |F| + 1$.

Proof. Given $d_1, \ldots, d_k$, we first construct graphs $G_1, \ldots, G_k$ such that G_i has a central node f_i of degree d_i, and $2d_i$ nodes on the "circumference," $x_{i1}, \ldots, x_{i(2d_i)}$. We then obtain G by joining $x_{i(2d_i)}$ and $x_{(i+1)1}$ by an edge, for $1 \leqslant i < k$, as shown in Fig. 5. We choose the minimal length routing ρ that takes a path through $f_1, \ldots, f_k$ whenever possible. We leave it to the reader to check that G and ρ have the required properties. ▮

Not surprisingly, the bounds of Theorems 4 and 6 can be combined to get

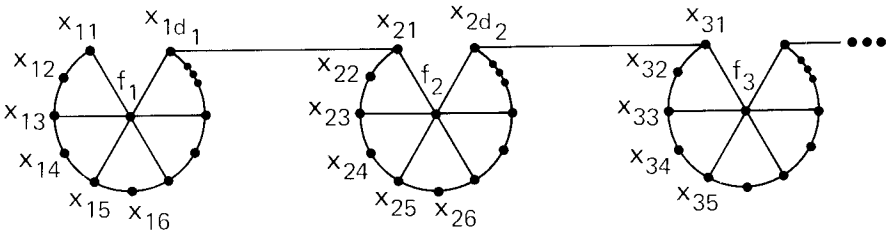

FIGURE 5

THEOREM 8. (Feldman, 1985). *If G/F is connected, and ρ is any minimal length routing of G, then* $\mathrm{DIAM}(R(G, \rho)/F) = \|F\| - |F_V| + 1$.

We can also combine the constructions of Theorems 5 and 7 to show that this result is optimal. We leave details to the reader.

The last issue we consider in this section is connectivity. The examples of Theorems 5 and 7 of graphs with a given diameter were graphs of low connectivity. The reader may wonder if we could have also constructed similar examples with high connectivity. The answer is yes, as the following theorem shows. Once we have an example of a graph where a certain number of edge faults and vertex faults cause the resulting surviving route graph to have a given diameter, we can construct a graph with arbitrarily high connectivity with the same property.

THEOREM 9. *Given a minimal length routing ρ on a graph G, a set F of faults that does not disconnect G, and any desired node connectivity k, there is a graph $G^* = (V^*, E^*)$ containing G as a subgraph and a minimal length routing ρ^* on G^* containing ρ as a subrouting such that G^* is at least k connected and* $\mathrm{DIAM}(R(G^*, \rho^*)/F)$ *is at least as large as* $\mathrm{DIAM}(R(G, \rho)/F)$.

Proof. Let $G = (V, E)$, ρ, and F be as in the statement of the theorem. Roughly speaking, G^* consists of G together with k copies of G/F, with corresponding nodes on G and each of the copies joined to form complete graphs on $k+1$ nodes. However, in each of the copies, we place two extra nodes on each of the edges of G/F. The result is that the distance between x and y in the copy is three times that between x and y in G. This means that it is always "faster" to travel in G than to travel in a copy.

More formally, let $G^* = (V^*, E^*)$, where $V^* = V \cup \{xyi \mid x = y \in V/F$ or $(x, y) \in E/F, 0 < i \leqslant k\}$ and E^* consists of all the edges of E as well as:

1. If $xyi, xyj \in V^*$ and $i \neq j$, then $(xyi, xyj) \in E^*$.
2. If $xxi, xyi \in V^*$ and $x \neq y$, then $(xxi, xyi) \in E^*$.
3. If $xyi, yxi \in V^*$ and $x \neq y$, then $(xyi, yxi) \in E^*$.
4. If $x, xyi \in V^*$, then $(x, xyi) \in E^*$.

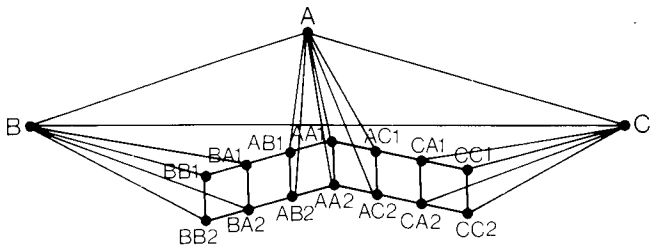

FIGURE 6

Note that corresponding to the edge (x, y) in G/F, we have the path xxi, xyi, yxi, yyi in G^*. Using this observation, we can show that given any two nodes of the form xyi and uvi in V^*, there is a path in G^* between these nodes that stays on the "ith level"; i.e., all nodes on the path are of the form wzi. It easily follows that G^* is $k+1$ node connected. Figure 6 is an example of the construction of G^* when G is the triangle ABC, F consists of the edge BC, and $k=2$.

Let ρ^* be a minimal length routing on G^* which extends ρ. For x in V define $\text{New}(x) = \{xyi \mid xyi \in V^*\}$. If x and y are nodes in G/F such that $d_{G/F}(x, y) > 1$, $a \in \text{New}(x)$ and $b \in \text{New}(y)$, then it is easy to check that

(i) $d_{G^*/F}(a, b) = d_{G/F}(x, y) + 2$.

(ii) $d_{G^*/F}(x, b) = d_{G/F}(x, y) + 1$.

(iii) $d_{G^*/F}(x, y) = d_{G/F}(x, y)$.

(This is where we need the extra nodes added in the copies of G/F to ensure that it is always faster to travel through G.)

Using the properties of ρ^* described above, it now follows by a straightforward induction on i that if $a \in \text{New}(x) \cup \{x\}$ and $b \in \text{New}(y) \cup \{y\}$, and $d_{R(G^*, \rho^*)/F}(a, b) = i$, then $d_{R(G, \rho)/F}(x, y) \leqslant i$. The theorem immediately follows. ∎

6. Open Problems

Although we have obtained a number of results, many open problems remain in this area. We list a few of them here:

1. We have obtained much better bounds than the general bounds for completely connected graphs and for the n-dimensional cube. Are there other classes of networks with equally good bounds? (Some results along these lines have been proved in Broder *et al.*, 1984.)

2. Can the upper bound of Theorem 3 be improved? In case $t \leqslant 2$ we can show that any graph G has a routing ρ with $\mathrm{DIAM}(R(G, \rho)/F) \leqslant 3$. For case $t = 1$, the square shows that this result is best possible for bidirectional routing. We conjecture that the results for the n-dimensional cube generalize; that is, for any graph G there is a bidirectional routing ρ such that if $|F|$ is less than the connectivity of G, we have $\mathrm{DIAM}(R(G, \rho)/F) \leqslant 3$ and a unidirectional routing ρ' such that (again if $|F|$ is less than the connectivity of G) $\mathrm{DIAM}(R(G, \rho')/F) \leqslant 2$.

3. The proofs of Theorems 4 and 6 do not use the connectivity of G but only the fact that G/F is connected. However, the connectivity of G is heavily used in Theorem 3. Can results along the lines of Theorem 3 be proved for graphs whose connectivity is less than $t+1$?

4. A routing ρ is *consistent (prefix consistent, suffix consistent)* if every subpath (resp. prefix, suffix) of a route is also a route. Consistent routings are of interest, since they are the ones that arise in practice (for example, the routings constructed by the algorithm used in the Highly Available Systems project are consistent). The routings λ_n of Theorem 2 are consistent, but the routing constructed in the proof of Theorem 3 is not necessarily even suffix consistent. What are the corresponding bounds for consistent, prefix consistent, and suffix consistent routings?

5. What happens to the diameter of the surviving route graph if the routing is a random routing?

6. What, if anything, can one say about routings that are almost minimal length?

7. We have assumed that the graphs representing communication networks have undirected edges. We can also consider what happens if we have directed communication networks. This corresponds to having one-way communication links. What are the analogs of our results for directed graphs? We remark that we can construct an example of a directed graph G and a minimal length routing ρ on G such that the diameter of $R(G, \rho)/F$ is $O(n)$ even if F consists of only one faulty edge, so that Theorem 6 does *not* hold if G is a directed graph. (The example has much the same flavor of the spoke example given in the Introduction.)

Define $\|e\|$ for a directed edge e with source node a to be $\|a\|$ and define $\|e\| = 2$ for an undirected edge e. As before, $\|F\| = \sum_{f \in F} \|f\|$. We conjecture that if ρ is a minimal length routing and G/F is connected, then

$$\mathrm{DIAM}(R(G, \rho)/F) \leqslant \|F\| - |F_V| + 1.$$

Note that this is a generalization of Theorem 8.

In practice graphs where every node has degree $\leqslant 3$ frequently arise. If

this conjecture is true, then if G is such a graph and ρ is a minimal length routing, then $\mathrm{DIAM}(R(G, \rho)/F) \leqslant 2|F| + 1$ for any collection F of node and edge faults that do not disconnect G.

ACKNOWLEDGMENTS

We thank Fan Chung, Larry Stockmeyer, and Dexter Kozen for many interesting discussions on the subject of proving diameter bounds. The referees of this paper also made a number of useful suggestions for improving the presentation.

RECEIVED July 2, 1984; ACCEPTED July 15, 1986

REFERENCES

AGHILI, H., ASTRAHAN, M., FINKELSTEIN, S., KIM, W., MCPHERSON, J., SCHKOLNICK, M., AND STRONG, R. (1983), "A Prototype for a Highly Available Database System," IBM Research Report RJ3755, January.

BERGE, C. (1976), "Graphs and Hypergraphs," North-Holland, New York.

BRODER, A., DOLEV, D., FISCHER, M., AND SIMONS, B., (1984), Efficient fault-tolerant routings in networks, *in* "Proceedings, 16th Annual ACM Symposium on Theory of Computing," pp. 536–541.

CHUNG, F. R. K., AND GAREY, M. R. (1984), Diameter bounds for altered graphs, *Journal of Graph Theory* **8**, 511–534.

DOLEV, D., HALPERN, J. Y., SIMONS, B., AND STRONG, R. (1984), A new look at fault-tolerant network Routing, *in* "Proceedings, 16th Annual ACM Symposium on Theory of Computing," pp. 526–535.

DOLEV, D., AND STRONG, R. (1983), Authenticated algorithms for Byzantine agreement, *SIAM J. Comput.* **12**, No. 4, 656–666.

EVEN, S. (1979), "Graph Algorithms," Comput. Sci. Press, Potomac, MD.

FELDMAN, P. (1985), Fault tolerance of minimal routings in a network, *in* "Proceedings, 17th Annual ACM Symposium on Theory of Computing," pp. 327–334.

STRONG, R., AND DOLEV, D. (1982), Byzantine agreement, *in* "COMPCON83: Digest of Papers," pp. 77–82; IBM Research Report RJ3714, December.

VALIANT, L. G. (1982), A scheme for fast parallel communications, *SIAM J. Comput.* **11**, May, 350–361.

Printed by the St. Catherine Press Ltd., Tempelhof 41, Bruges, Belgium