A NEW LOOK AT FAULT TOLERANT NETWORK ROUTING

Danny Dolev

Hebrew University, Givat Ram
91904 Jerusalem, Israel

Joe Halpern
Barbara Simons
Ray Strong

IBM Research Laboratory
San Jose, California 95193

ABSTRACT: Consider a communication network G in which a limited number of link and/or node faults F might occur. A routing ρ for the network (a fixed path between each pair of nodes) must be chosen without any knowledge of which components might become faulty. Choosing a good routing corresponds to bounding the diameter of the *surviving route graph* $R(G,\rho)/F$, where two nonfaulty nodes are joined by an edge if there are no faults on the route between them. We prove a number of results concerning the diameter of surviving route graphs. We show that if ρ is a minimal length routing, then the diameter of $R(G,\rho)/F$ can be on the order of the number of nodes of G, even if F consists of only a single node. However, if G is the n-dimensional cube, the diameter of $R(G,\rho)/F \leq 3$ for any minimal length routing ρ and any set of faults F with $|F|<n$. We also show that if F consists only of edges and does not disconnect G, then the diameter of $R(G,\rho)/F$ is $\leq 3|F|+1$, while if F consists only of nodes and does not disconnect G, then the diameter of $R(G,\rho)/F$ is $\leq$ the sum of the degrees of the nodes in F, where in both cases ρ is an arbitrary minimal length routing. We conclude with one of the most important contributions of this paper: a list of interesting and apparently difficult open problems.

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* assigns to any pair 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.

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.

Consider the network shown in Figure 1.

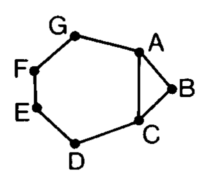

Figure 1.

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).

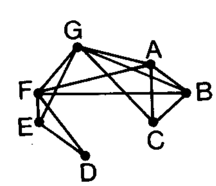

Figure 2.

Suppose processor C wants to broadcast a message to all processors. Since C can only send messages 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 Figure 2.

In general, 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. As we noted above, the diameter of $R(G,\rho)/F$ measures the number of rebroadcasts necessary to ensure that all processors get a message. This number determines the number of phases for which it is necessary to run certain distributed protocols (such as the Byzantine agreement protocols of [DS1, DS2]). 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 gives a good estimate on the time required to complete a broadcast in the presence of faults. Thus, our problem will be to choose a routing ρ on G that is fault tolerant because the diameter of $R(G,\rho)/F$ remains small for any set of faults F of a given cardinality. This problem has given rise to many interesting questions in graph theory, some of them still open.

We first note that minimal length routings are not always optimal. Consider the spoke graph shown in figure 3.

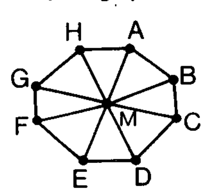

Figure 3.

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. But now suppose the center node becomes faulty. Then with this routing it is easy to see that 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 figure 3 would be ABCD, rather than AMD). In this case, the diameter of $R(S_n,\rho)/F$ is easily seen to be ≤ 2 if $|F| \leq 2$.

This leads us to ask if we can always find good routings. We show (Theorem 3) that for any $(t+1)$-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 $|F| \leq t$.

Although minimal length routings are not always optimal, they are useful and easy to generate. A common routing algorithm (used for example in the Highly Available Systems project at IBM [AA]) 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 it is easy to check that the diameter of $R(K_n,\rho)/F$ is 2 if $|F| \leq n-2$. (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 $|F| \leq n-2$, 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 (eg. [Va]) 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 $R(C_n,\rho)/F \leq 3$, 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 $R(C_n,\lambda_n)/F \leq 2$ if $|F| < n$. This in fact is a corollary to a more general result of [BD] (although the proof for this special case is much simpler than that of [BD]).

The spoke example shows that if we use minimal length routings, even a single node fault can force the diameter of the surviving route graph to grow to O(n). A closer look at this example suggests that the diameter can grow this way only if there are nodes of high degree. Indeed, we show (Theorem 4) that if F consists only of node faults, G/F (i.e.

G with all the elements of F removed) 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. The situation is quite different for edge faults. If F consists only of edge faults, we can show (Theorem 6) that if G/F is connected and ρ is any minimal length routing on G, then the diameter of $R(G,\rho)/F$ is $\leq 3|F|+1$. We conjecture that both of these results can be somewhat improved: to $\|F\|-|F|+1$ in the case of node faults, and to $2|F|+1$ in the case of edge faults. We show in Theorems 5 and 7 that there exist graphs in which these conjectured bounds are attained.

Chung and Garey [CG] were able to obtain analogous results to Theorems 4, 5, 6, and 7 for surviving graphs G/F (as opposed to surviving route graphs). 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 $\leq (1+|F|)(1 +$ the diameter of G). 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 an 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 $\rho: V\times V \rightarrow V^*$ 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 routed 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 [Be] 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 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 i^{th} 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 *.

Networks in the form of n-dimensional cubes display surprisingly good performance. Theorem 1 states that the surviving route graph produced 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. Then if $|F|<n$, $DIAM(R(C_n,\rho)/F)\leq 3$.

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 $|F|<n$, $DIAM(R(C_n,\lambda_n)/F)\leq 2$.

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 1/2. Let $|x|$ denote the weight of x. Thus $|11101|=4$ and $|1{*}101|=3.5$. By dropping the i^{th} coordinate, any n-dimensional object can be *projected* along the i^{th} coordinate onto an (n-1)-dimensional object. Let P_i be the operator for projecting along the i^{th} coordinate. Note that an edge may project to a node. Thus $P_2(11101)=1101=P_2(1{*}101)$. We write $x\leq y$ when $\leq$ holds on each coordinate. We write $x<y$ when $x\leq y$ and $<$ holds 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 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 $x_1,\ldots,x_k$ 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 x,y is safe with respect to F.

Proof: A minimal length path from x to y must stay in $C(x,y)$. □

By Lemma 1, if a sequence is safe with respect to F, then it will be a path in $R(C_n,\rho)/F$ for every minimal length routing ρ. Thus, Theorem 1 follows immediately from Lemma 2 below.

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 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 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 $x<u<v\leq y$.

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 $0^{n-1}<u<v\leq 1^{n-1}$ that is safe with respect to F'. If $v<1^{n-1}$, then it is easy to check that $0^n<0u<1v<1^n$ 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 $0^{n-1}<a<b\leq 1^{n-1}$. Among all such sequences there must be one $0^{n-1}<u<v\leq 1^{n-1}$ with $|u|$ maximal. We claim that $0^n<0u<0v<1^n$ is safe for F. It is clearly safe for F - $\{f\}$, so we must show only that it is safe for f. Since $f_1 = 1$, it suffices to show that $f \notin C(0v, 1^n)$. But if $f \in C(0v, 1^n)$, we must have $|P_1(f)| \geq |v|$. Since f was chosen with minimal weight and $f_1 = 1$, it follows that $|P_1(f')| \geq |P_1(f)| \geq |v|$ for all $f' \in F'$. But then $0^{n-1} < v < 1^{n-1}$ ($\leq 1^{n-1}$) is safe for F', contradicting 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, $x_1 \ldots x_{i-1} 0 y_{i+1} \ldots y_n$, y is a path in $R(C_n, \lambda_n)/F$. Otherwise, let $F' = P_i(\{f \in F \mid f_i = 1\})$. 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 i^{th} 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, $u_1 \ldots u_{i-1} 1 u_{i+1} \ldots u_n$, 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 $1 \leq i \leq n$ 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 $\rho(0y, 1^n)$ 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 analogue 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 $n \geq 3$, it is still an open question if there exists a bidirectional analogue of Theorem 2. It would also be interesting to know whether there is a bidirectional analogue 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 $z_1, \ldots, z_n$ and sends the message to $z_1, \ldots, z_n$ with instructions to forward it to y. One message must get through so long as $|F| < n$.

4. The techniques of proof of Theorems 1 and 2 easily generalize to any n-dimensional rectangular grid (product of n intervals).

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 $|F| \leq 2$. 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 $|F| \leq t$, then $DIAM(R(G, \rho)/F) \leq max(2t, 4)$.

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. 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 subgraphs, $G_1, G_2, \ldots, G_k$, with $k \geq 2$. Moreover, if $x \in G_i$, $i = 1, 2, \ldots, k$, 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 is a separating set of G, 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, 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 the following 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 just the path chosen above.

We note that standard techniques from network flow ([Ev]) can be used to obtain the routes in ρ in time $O(|V|^{1/2}|E|^2)$.

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

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 $|F_1| + \ldots + |F_k| + |F_M| = t$.

We now complete the proof that $DIAM(R(G,\rho)/F) \leq \max(2t,4)$ 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 z has a path in R/F to every node in M. Therefore, there exists a route of length 2 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. No F_i is empty.
Let $P = x_0 \ldots x_k$ be some minimum length path in R/F between $x = x_0$ and $y = x_k$. We bound the length of P by counting nodes in M which either appear on P or are adjacent to those internal nodes of P which are themselves not in M. Thus, for $x_i \in P$, let $(x_i) = \{$nonfaulty nodes in M to which x_i 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 is a path in R/F from x_i to at least $t+1-|F_j|$ non-faulty nodes of M. Since $|F_1|, |F_2| \geq 1$ and $|F_1| + \ldots + |F_k| + |F_m| = t$, we must have $|F_j| \leq t-1$, and so $|(x_i)| \geq 2$.

Let $P_{i..j}$ be the partial path $x_i x_{i+1} \ldots x_j$. Let $S(P_{i..j}) = (x_i) \cup \ldots \cup (x_j)$. We prove the 2t bound by showing that $|S(P_{0..i})| \geq \lceil i/2 \rceil + 1$ by induction on i.

Since $|(x_0)| \geq 2$, 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 $j \leq i-2$, for otherwise $P_{0..j}x_iP_{(i+1)..k}$ is a shorter path from x to y than is P. Thus,

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

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

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

Since $P = x_0 x_1 \ldots x_k$, it follows that $S(P) \geq \lceil k/2 \rceil + 1$. Since $|M| \leq t+1$, we must have $\lceil k/2 \rceil \leq t$. Consequently, $k \leq 2t$ and $|P| \leq 2t$. □

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.

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 $\|F\|=\Sigma_{f \in F}\|f\|$.

We first consider the case where F consists only of nodes.

Theorem 4: If F consists only of nodes, G/F is connected, and ρ is any minimal length routing of G, then $DIAM(R(G,\rho)/F) \le \|F\|$.

Proof: First note that if the result holds for $|F|=1$, then it can be extended to all sets of node faults by the following argument. Given a graph $G=(V,E)$, a set of node faults F such that G/F is connected, and a minimal length routing ρ, construct a graph G' in which all the faulty nodes are combined into one node. More precisely, let $G'=(V',E')$, where $V'=(V-F)\cup\{x\}$, where x is a distinguished node, and $E'=E/F\cup\{(a,x) \mid a\in V/F \text{ and for some } f\in F, (a,f)\in E\}$. Note that since the neighbors of x in G' are exactly the non-faulty neighbors in G of the nodes in F, we have $\|x\| \le \Sigma_{f\in F}\|f\|$. Let ρ' be any minimal length routing on G' such that if there is a minimal length path from v to w which goes through x, then $\rho'(v,w)$ goes through x. It is now easy to check that if $\rho(v,w)$ goes through a node in F, then $\rho'(v,w)$ goes through x. Let $F'=\{x\}$. By the observation above, it follows that $R(G',\rho')/F'$ is a subgraph of $R(G,\rho)/F$: it has the same nodes, but possibly fewer edges. Thus $DIAM(R(G,\rho)/F)\le DIAM(R(G',\rho')/F')$. But by hypothesis, $DIAM(R(G',\rho')/F')\le\|x\|$, giving the desired result.

We now prove the result for $|F|=1$. So suppose that F consists of only one faulty node f. Let $N(f)=\{w \mid \text{there exists an edge from f to w in G}\}$ (these are the neighbors of f in G). Note that by definition $|N(f)|=\|f\|=\|F\|$. Fix a minimal length route ρ, and let $H=R(G,\rho)/F$. We want to show $DIAM(H)\le\|f\|$.

Since ρ is a minimal length routing, if two nodes u and v are neighbors in G, then $\rho(u,v)$ is just the edge between them. It follows that G/F is a subgraph of H, and so H is connected. Choose any two nodes a and b in G/F such that $d_H(a,b)=DIAM(H)$. Let L_i, $i=0,\ldots,DIAM(H)$, consist of $\{c \mid d_H(a,c)=i\}$. For any node $c\in G/F$, let $Near(c)=\{w\in N(f) \mid \text{for all } w'\in N(f),\ d_G(c,w)\le d_G(c,w')\}$. Thus Near(c) consists of the nodes in N(f) that are closest (in G) to c. Finally, let $Near(L_i)=\cup_{c\in L_i}Near(c)$. Note that since $L_i\neq\phi$ for $i=0,1,\ldots,DIAM(H)$, it follows that $Near(L_i)\neq\phi$. We will show:

Claim 1: $|\cup_{i<k}Near(L_i)|\ge k$.

From the claim it follows that

$$\|f\|=|N(f)|\ge|\cup_{i<DIAM(H)}Near(L_i)|\ge DIAM(H),$$

which establishes the result for $|F|=1$. To prove the claim, we proceed by induction. The claim is trivial if $k=0$ or $k=1$. Suppose we have proved the claim for $k'<k$. We now prove it for $k\ge2$. We first need to show:

Claim 2: If $|i-j|\ge2$, then $Near(L_i)$ and $Near(L_j)$ are disjoint.

To prove Claim 2, suppose otherwise. Then there exists $b\in N(f)$ such that $b\in Near(u)$ and $b\in Near(v)$ for some $u\in L_i$ and some $v\in L_j$. We will now show that no minimal length path in G from u to v can contain the node f. This then shows that u and v must have an edge between them in H. Consequently $|i-j|\le1$, which contradicts the assumption that $|i-j|\ge2$.

So suppose by way of contradiction that some minimal length path in G from u to v contains f. Thus the path is of the form $\pi=u\ldots wfw'\ldots v$, where $w,w'\in N(f)$. Since $b\in Near(u)$, we must have $d_G(u,b)\le d_G(u,w)$. Similarly, since $b\in Near(v)$, we must have $d_G(b,v)\le d_G(b,w')$. One way of getting from u to v in G is to go from u to b and then from b to w. By the observations above, this gives a shorter path than π, contradicting the minimality of π and proving Claim 2.

Returning to the proof of Claim 1, suppose by way of contradiction that Claim 1 does not hold and consider the minimum $k\ge2$ for which it does not hold. Note that $|Near(L_{k-1})|=1$, for if $|Near(L_{k-1})|\ge2$, Claim 2 and the fact that the result holds for k-2 implies that Claim 1 also holds for k. Moreover, $Near(L_{k-1})$ cannot be disjoint from $Near(L_{k-2})$, for then from the fact that Claim 1 holds for k-1 we could again immediately prove it for k. So assume without loss of generality that $Near(L_{k-1})=\{c\}$, and that $c\in Near(L_{k-2})$.

Choose $u\in L_{k-1}$ and $v\in L_k$ such that u and v are neighbors in G. (There must be such u and v. Choose any path in G/F from a to b. Let v be the first node in L_k on that path, and u be the node just before v. We leave it to the reader to check that there exist such u and v and that they have the desired property.) Let $d\in Near(v)$. We must have $c\neq d$, otherwise we would have $c\in Near(L_{k-2})\cap Near(L_k)$, contradicting Claim 2. We now show that $d\in Near(u)$, contradicting the hypothesis that $Near(L_{k-1})=\{c\}$. Choose $t\in L_{k-2}$ such that $c\in Near(t)$. (Such a t exists since $c\in Near(L_{k-2})$ by hypothesis.) There must be some minimal length path in G from t to v that goes through f (otherwise v would be in L_j for some $j\le k-1$). Suppose this path is of the form $t\ldots wfw'\ldots v$, where $w,w'\in N(f)$. Another way of getting from t to v is to take a minimal length path from t to c, followed by a minimal length path from c to u, followed by the edge from u to v. Thus we must have $d_G(t,w)+2+d_G(w',v) \le d_G(t,c)+d_G(c,u)+1$. Since

$c \in Near(t)$ and $d \in Near(v)$, we must have $d_G(t,c) \leq d_G(t,w)$ and $d_G(d,v) \leq d_G(w',v)$. Hence $d_G(t,c)+2+d_G(d,v) \leq d_G(t,c)+d_G(c,u)+1$, which implies $d_G(d,v)+1 \leq d_G(c,u)$. One way to get from d to u is to go from d to v and then to take the edge from v to u. Hence $d_G(d,u) \leq d_G(d,v)+1$. Combining these inequalities, we get $d_G(d,u) \leq d_G(c,u)$. But since $c \in Near(u)$, it follows that we must also have $d \in Near(u)$, and thus $d \in Near(L_{k-1})$, contradicting the assumption that $Near(L_{k-1})=\{c\}$. □

The previous theorem is close to optimal, as we now show.

Theorem 5: For all $d_1,...,d_k$, there exists a graph G, a minimal length routing ρ on G, and a set of node faults $F=\{f_1,...,f_k\}$ which does not disconnect G such that the degree of f_i is d_i, $i=1,...,k$, and $DIAM(R(G,\rho)/F)=\|F\|-|F|+1$.

Proof: The graph G is a simple generalization of the spoke example of the introduction. We give G below in Figure 4 in the case that $k=3$ and $d_1=d_2=3$ and $d_3=4$. We choose the minimal length routing that takes a path through f_1, f_2, or f_3 whenever possible:

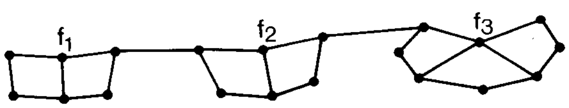

Figure 4.

We believe that the upper bound of Theorem 4 can be improved to the lower bound shown in Theorem 5. We formalize this as:

Conjecture 1: For any graph G, any minimal length routing ρ, and any set of node faults F which does not disconnect G,

$$DIAM(R(G,\rho)/F) \leq \|F\|-|F|+1.$$

The situation for edge faults is somewhat different.

Theorem 6: If F consists only of edges, G/F is connected, and ρ is any minimal length routing of G, then $DIAM(R(G,\rho)/F) \leq 3|F|+1$.

We use the following lemmas to prove Theorem 6. Recall that the pair of nodes u, v is *safe* with respect to F if all the minimum length paths from u to v avoid F.

Lemma 4: Assume F has only edges. If the pair of nodes x,y is not safe with respect to F, then there is an edge (u,v) in F and a minimal length path from x to y containing (u,v) such that v,y is safe with respect to F.

The proof of Lemma 4 is a straightforward induction on the distance between x and y in G.

Proof of Theorem 6: Let x and y be a pair of nodes of G and let $x_0x_1...x_k$ be a minimal length path from $x=x_0$ to $y=x_k$ in $H=R(G,\rho)/F$ (such a path must exist by the same arguments as in Theorem 4 above; since ρ is a minimal length routing, it follows that G/F is a subgraph of H so H is connected). Note that by construction $d_H(x_i,x_j) = |i-j|$.

For each x_i with $i \geq 2$, we select an endpoint v of an edge in F such that v,x_i is safe with respect to F and v lies on a minimal length path in G from x to x_i (Lemma 4). We say that v is *associated with* x_i. Intuitively, v is the endpoint of a faulty edge closest to x_i on a path from x. Note that if $(u,v) \in F$ and v is associated with some x_i, then $d_G(x,u) < d_G(x,v)$, and so u cannot be associated with any x_i. Thus, at most one node of each faulty edge can be associated with some x_i. Moreover, a particular node v can be associated with at most three x_i's. For if v were associated with x_i and x_j, and $|i-j|>2$, then since $d_H(x_i,v) = d_H(x_j,v) = 1$, we have $d_H(x_i,x_j) \leq 2$, a contradiction. Now, by a simple counting argument, we get that $k \leq 3|F|+1$. □

Again we have a lower bound that almost matches the upper bound of Theorem 6.

Theorem 7: 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 $DIAM(R(G_t,\rho_t)/F_t) = \|F_t\|+1$ $(= 2t+1)$.

Proof: The required graph G_t is obtained by the obvious generalization from the graph G_4 shown in figure 5, where the edges marked with an x through them are in F, and ρ goes through a faulty edge whenever possible (for example, $\rho(A,G)=ABG$).

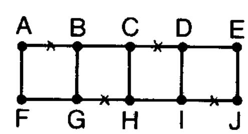

Figure 5.

Conjecture 2: For any graph G, any minimal length routing ρ on G, and any set of edge faults F which does not disconnect G, $DIAM(R(G,\rho)/F) \leq \|F\|+1 (=2|F|+1)$.

Combining the graphs of Theorems 5 and 7 leads us to the following:

Corollary: For all k, m, d_1, ..., d_m, we can find a graph G, a minimal length routing ρ on G and set of k edge faults and m node faults F with degrees $d_1,...,d_m$ such that G/F is connected and $DIAM(R(G,\rho)/F)=\|F\|-|F_V|+1$.

Of course, we also have:

Conjecture 3: For any graph G, minimal length routing ρ, and set of faults F which does not disconnect G, $DIAM(R(G,\rho)/F)=\|F\|-|F_V|+1$.

Conjectures 1, 2, and 3 seem to be quite difficult to prove. L. Stockmeyer has recently shown that Conjecture 1 implies Conjecture 2. We (and independently F. Chung) have also shown that the bound of Theorem 6 can be improved to $2.5|F|+1$. We can show that the bound of $2|F|+1$ of Conjecture 2 holds for $|F|=1,2,3$. The proof for $|F|=1$ is of independent interest, and we present it below. Our proofs for $|F|=2$ and $|F|=3$ grow exponentially more difficult, and do not seem to lead to any generally applicable technique.

Theorem 8: If ρ is a minimal length routing on G, and F consists of one edge which does not disconnect G, then $DIAM(R(G,\rho)/F)\leq 3$.

Proof: Let $H=R(G,\rho)/F$. Choose $a,b \in G/F$ such that $d_H(a,b)\neq 1$. Since G/F is connected, there must be some path $a_1...a_k$ from $a=a_1$ to $b=a_k$ in G/F. Let a_j be the last point on this path such that $d_H(a,a_j)=1$. Note $j\geq 2$, since the edge between $a=a_1$ and a_2 in G is nonfaulty by hypothesis. Since the edge in G between a_j and a_{j+1} is also nonfaulty by hypothesis, $d_H(a_j,a_{j+1})=1$. We now show $d_H(a_{j+1},b)=1$, which will complete our argument (since $d_H(a,b) \leq d_H(a,a_j)+d_H(a_j,a_{j+1})+d_H(a_{j+1},b) = 3$). Let the faulty edge in F be (x,y). Since $d_H(a,b)\neq 1$, $\rho(a,b)$ must be of the form a...xy...b. Because ρ is a minimal length routing, we must have $d_G(a,x)\leq d_G(a,y)$, and $d_G(b,x)\geq d_G(b,y)$. By choice of a_j, $d_H(a,a_{j+1})\neq 1$, so the edge (x,y) must also be on $\rho(a,a_{j+1})$. But since $d_G(a,x)\leq d_G(a,y)$, $\rho(a,a_{j+1})$ must be of the form $a...xy...a_{j+1}$. Thus $d_G(a_{j+1},y)\leq d_H(a_{j+1},x)$. The above inequalities tell us that y is closer to both a_{j+1} and b than x is, so no minimal length path from a_{j+1} to b could have the edge (x,y) on it. Thus $\rho(a_{j+1},b)$ is fault-free, and $d_H(a_{j+1},b)=1$ as desired. □

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. However, 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 $DIAM(R(G^*,\rho^*)/F)$ is at least as large as $DIAM(R(G,\rho)/F)$.

Proof: Left to full paper. □

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?

2. Prove Conjectures 1, 2, and 3.

3. Can the upper bound of Theorem 3 be improved? In case $t\leq 2$ we can show that any graph G has a routing ρ with $DIAM(R(G,\rho)/F)\leq 3$. 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 with $DIAM(R(G,\rho)/F)\leq 3$ and a unidirectional routing ρ' such that $DIAM(R(G,\rho')/F)\leq 2$.

4. 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. What happens when this assumption is relaxed?

5. A routing ρ is *consistent (prefix consistent, suffix consistent)* if every subpath (respectively, prefix, suffix) of a route is also a route. 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?

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

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

8. 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 analogues 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, $\|F\|=\Sigma_{f \in F}\|f\|$. We conjecture that if ρ is a minimal length routing and G/F is connected, then

$$DIAM(R(G,\rho)/F) \le \|F\|-|F_V|+1.$$

Note that this is a generalization of Conjecture 3.

In practice graphs where every node has degree ≤ 3 frequently arise. If these conjectures are true, then if G is such a graph and ρ is a minimal length routing, then $DIAM(R(G,\rho)/F) \le 2|F|+1$ for any collection F of node and edge faults that do not disconnect G.

Acknowledgements

We would like to thank Fan Chung, Larry Stockmeyer, and Dexter Kozen for many interesting discussions on the subject of proving diameter bounds. Fan was instrumental in pointing out many of the ideas that led to the proof of Theorem 4.

References

[AA] H. Aghili, M. Astrahan, S. Finkelstein, W. Kim, J. McPherson, M. Schkolnick, R. Strong, "A Prototype for a Highly Available Database System," *IBM Research Report RJ3755*, January, 1983.

[Be] C. Berge, *Graphs and Hypergraphs*, North Holland, New York, 1976

[BD] A. Broder, D. Dolev, M. Fischer, B. Simons, Efficient fault tolerant routings in networks, unpublished manuscript, 1983.

[CG] F. R. K. Chung and M. R. Garey, "Diameter Bounds for Altered Graphs," unpublished manuscript, 1983.

[DS1] D. Dolev and R. Strong, "Authenticated Algorithms for Byzantine Agreement," to appear in *SIAM J. Comput*, see also *IBM Research Report RJ3416*, March, 1982.

[DS2] D. Dolev and R. Strong, "Byzantine Agreement," *IBM Research Report RJ3714*, December, 1982.

[Ev] S. Even, *Graph Algorithms*, Computer Science Press, Potomac, Maryland, 1979.

[Va] L. G. Valiant, "A Scheme for Fast Parallel Communications," *SIAM J. Comput.*, 11, May 1982, pp. 350-361.