Communication in Networks with Random Dependent Faults

Evangelos Kranakis[1], Michel Paquette[1], and Andrzej Pelc[2]

[1] School of Computer Science, Carleton University,
Ottawa, Ontario, K1S 5B6, Canada
kranakis@scs.carleton.ca, michel.paquette@polymtl.ca
[2] Département d'informatique et d'ingénierie, Université du Québec en Outaouais.
Gatineau, Québec, J8X 3X7, Canada
pelc@uqo.ca

Abstract. The aim of this paper is to study communication in networks where nodes fail in a random dependent way. In order to capture fault dependencies, we introduce the *neighborhood fault* model, where damaging events, called *spots*, occur randomly and *independently* with probability p at nodes of a network, and cause faults in the given node and all of its neighbors. Faults at distance at most 2 become dependent in this model and are positively correlated. We investigate the impact of spot probability on feasibility and time of communication in the fault-free part of the network. We show a network which supports fast communication with high probability, if $p \leq 1/c \log n$. We also show that communication is not feasible with high probability in most classes of networks, for constant spot probabilities. For smaller spot probabilities, high probability communication is supported even by bounded degree networks. It is shown that the torus supports communication with high probability when p decreases faster than $1/n^{1/2}$, and does not when $p \in 1/O(n^{1/2})$. Furthermore, a network built of tori is designed, with the same fault-tolerance properties and additionally supporting fast communication. We show, however, that networks of degree bounded by a constant d do not support communication with high probability, if $p \in 1/O(n^{1/d})$. While communication in networks with independent faults was widely studied, this is the first analytic paper which investigates network communication for random dependent faults.

Keywords: Fault-tolerance, dependent faults, communication, crash faults, network connectivity.

1 Introduction

As interconnection networks grow in size and complexity, they become increasingly vulnerable to component failures. Links and nodes of a network may fail, and these failures often result in delaying, blocking, or even distorting transmitted messages. It becomes important to design networks in such a way that the desired communication task be accomplished efficiently in spite of these faults,

L. Kučera and A. Kučera (Eds.): MFCS 2007, LNCS 4708, pp. 418–429, 2007.

usually without knowing their location ahead of time. Such networks are called fault-tolerant.

The fundamental questions of network reliability have received much attention in past research under the assumption that components fail randomly and independently (cf., e.g. [1,2,3,4] and the survey [5]). On the other hand, empirical work has shown that positive correlation of faults is a more reasonable assumption for networks [6,7,8]. In [8], the authors provide empirical evidence that data packets losses are spatially correlated in networks, and in [7], the authors use the assumption of failure spatial correlation to enhance network traffic management. Furthermore, in [6], the authors simulate failures in a sensor network using a model much like that of the present paper; according to these authors, the environment provides many spatially correlated phenomena resulting in such fault patterns. Physical and logical phenomena generally affect physical components, causing failures in a positively correlated way. E.g., on August 14, 2003, faults cascaded on the power distribution network and deprived part of North America of electricity. Logical phenomena, like computer viruses and worms, also cause dependent faults. Lightning strikes hitting one node of an electric network cause power outages in entire city blocks.

As our society is increasingly dependent on information networks, it becomes essential to study questions relating to tolerance of dependent positively correlated faults. However, no analytic work has been done for communication networks under this assumption about faults.

In this paper, we consider the problem of feasibility and time of communication in networks with dependent positively correlated faults. To the best of our knowledge, this is the first analytic paper which provides this type of results for network communication.

1.1 Model and Problem Definition

A communication network is modeled as an undirected graph $G = (V, E)$ with a set of nodes V connected by a set of undirected links E. We say that two nodes are *adjacent* (or *neighbors*) if they share a link. The distance between nodes $u, v \in V$ is the minimum number of links which must be traversed from u to reach v; it is denoted by $dist(u, v)$. In a network, $\Gamma(u)$ is the set of nodes adjacent to u; $\Gamma_i(u)$ is the set of nodes $v \in V$ whose distance from u is i; we also denote by $\Gamma_{\leq i}(u)$ the set of nodes $v \in V$ whose distance from u is at most i. A node is said to be *functional*, or *fault-free*, when it executes only its predefined algorithm without any deviation, and doing so, transmits all messages correctly, in a timely manner and without any loss; a node which is not functional is said to be *faulty*. Faults can be of different types: at opposite ends of the spectrum are *crash* and *Byzantine* faults. Faults of the crash type cause faulty components to stop all communication; these components can neither send, receive nor relay any message. Faulty nodes of the Byzantine type may behave arbitrarily (even maliciously) as transmitters. We say that faults are permanent when they affect the nodes for the entire duration of a communication process; otherwise, the faults are said to be transient. In this paper, we assume that faults are permanent and of crash type. Throughout the paper, log means logarithm with base 2 and ln means the natural logarithm.

We consider communication in the fault-free part of the network, where all nodes exchange messages with each other. Communication among functional nodes is feasible if the fault-free part of the network is connected and contains at least two nodes. We measure communication time under the all-port message passing model, where nodes can communicate with all their neighbors in one round, and under the 1-port model, in which every node can send a message to at most one neighbor in one round. Under the all-port model, communication can be completed in time D if the fault-free part of the network has diameter D. Hence, we study the connectivity and diameter of the fault-free part of the network. Moreover, we seek networks of low maximum degree Δ. Since in the 1-port model communication can be completed in time $D\Delta$, networks of low maximum degree and low diameter of the fault-free part support fast communication also in the 1-port model.

In order to capture fault dependencies, we introduce the *neighborhood fault* model, where damaging events, called *spots*, occur randomly and *independently* at nodes of a network, with probability p, and cause permanent crash faults in the given node and all of its neighbors. Faults at distance at most 2 become dependent in this model and are positively correlated. We investigate the impact of spot probability on feasibility and time of communication in the fault-free part of the network.

We design general networks and bounded degree networks which support fast and highly reliable communication despite relatively high spot probabilities. We also prove bounds on spot probability such that highly reliable communication is not supported.

We focus attention on the problem of feasibility and time of communication, guaranteed with *high probability*, i.e., with probability converging to 1 as the size of the network grows. Under the all-port model, in which time of communication is proportional to the diameter, this problem reduces to the question for what spot probability the fault-free part of the network is connected and when it has diameter at most D, with high probability. Under the 1-port model, the same reduction is valid for networks of a given degree.

1.2 Related Work

Dependent fault models were introduced in the study of integrated circuit manufacturing yields. This research models defects as the result of impurities, positioned randomly and independently, affecting nearby circuit components in a dependent way. Results were proposed mainly according to the *quadrat-based* and *center-satellite* approaches. In [9], the author proposed a coarse approach to analyzing production yields based on the assumption that faults occurred in clusters inside a defined grid pattern on Very Large Scale Integration (VLSI) wafers; this quadrat-based model offered provably good results and ease of use required by the industry. Then, in [10], the authors introduced a detailed model of manufacturing defects in VLSI wafers based on the *center-satellite concept* for ecological sampling [11]. Later on, in [12], the authors proposed a simplified center-satellite model of manufacturing defects on VLSI wafers for the study of the memory array reconfiguration problem. In fact, both the center-satellite and quadrat-based approaches are still in use for System on Chip (SoC) (cf., e.g., [13]) and VLSI (cf., e.g., [14,15]) applications. Throughout this field of literature, the consensus is that results originating from the center-satellite approach, as opposed to quadrat-based approaches, are more difficult to apply but provide better prediction quality.

The above approach should be contrasted with the literature on fault-tolerant communication in networks. Many results concerned random link and/or node failures (cf., e.g. [1,2,3,4] and the survey [5]) but, to the best of our knowledge, in all cases faults were assumed to be independent. In [1], the author shows the existence of networks in which $O(\log n)$-time broadcast can be done, under the 1-port model, with high probability, despite links which fail randomly and independently with positive constant probability. In [2], the authors design a network of logarithmic degree which can support high probability communication in time $O(\log n)$ when faults occur randomly and independently on links and nodes with any constant probabilities smaller than 1. In [4], the authors design a similar network which can support communication with high probability in time $O(\log^2 n)$ with Byzantine faults.

Our present research focuses on communication network failures which occur in a dependent way. We consider networks modeled by arbitrary graphs, hence the geometry-dependent, quadrat-based approach to fault dependencies is not appropriate. Our neighborhood fault model, more appropriate for general graphs, is a simplified version of the center-satellite approach.

1.3 Our Results

All our results address the general problem for which spot probabilities p there exist networks supporting communication with high probability, and if so, if this communication is fast in the all-port and 1-port models. Hence we ask for which spot probabilities the fault-free part of the network is connected of size larger than 1, and if so, does it have a small diameter. Moreover, in our positive results we seek networks of low maximum degree.

In Section 2, we address the questions regarding general networks. We first show that there exists a constant c, such that for spot probability $p \leq 1/c \log n$, there exists an n-node graph whose fault-free part has logarithmic diameter and logarithmic degree, with high probability. Hence it supports high probability communication in time $O(\log n)$ in the all-port model and in time $O(\log^2 n)$ in the 1-port model. On the negative side, we show that for constant spot probability p, there exist constants c_1 and c_2 such that: if all degrees in a graph are at most $c_1 \log n$ then the graph is disconnected with high probability; if all degrees in a graph are at least $c_2 \log n$ then the graph has all nodes faulty with high probability. In either case, highly reliable communication is not possible. This leaves some very particular networks undecided. For example, this negative result does not cover the important case of the hypercube, for some constant spot probabilities. Therefore, we study the hypercube separately and prove that, for any constant spot probability $0 < p \leq 1$, this network does not support high probability communication. The above should be contrasted with the results from [2,3] showing that, for *independent* faults, fast highly reliable communication is possible for arbitrary constant fault probabilities in some graphs and for small constant fault probability, even in the hypercube.

In Section 3, we investigate communication in bounded degree networks. We show that the torus supports communication with high probability when $p \in 1/\omega(n^{1/2})$. (As usual, $\omega(f)$ denotes the set of functions g such that $g/f \to \infty$.) However, the diameter of an n-node torus is at least $\Theta(\sqrt{n})$ and the fault-free part has the same large diameter. Hence we seek networks with the same fault-tolerance properties, but with small diameter. We construct a bounded degree network built of tori, whose fault-free part has diameter $O(\log n)$ whenever $p \in 1/\omega(n^{1/2})$. Hence this network supports high probability communication in logarithmic time, both in the all-port and in the 1-port models. On the negative side, we show that neither the torus nor the above network can support highly reliable communication when $p \in 1/O(n^{1/2})$. Finally, we prove that networks of degree bounded by a constant d cannot support communication with high probability when $p \in 1/O(n^{1/d})$. Due to lack of space, many proofs are deferred to the journal version of this paper.

2 General Networks

In this section, we focus on general networks. We first design a network which supports communication with high probability when the spot probability is at most $1/c \log n$, for some positive constant c. We then establish two bounds on node degrees showing that a large class of networks cannot support communication with high probability when spot probability is a positive constant.

2.1 Upper Bounds

This section is dedicated to proving the following result.

Theorem 1. *There exists an n-node graph whose fault-free part has diameter $O(\log n)$ and logarithmic degree, with high probability, for spot probability $p \leq 1/c \log n$, where c is some positive constant.*

The network construction is based on a binary tree structure where each node of the tree represents a group of nodes and each link of the tree represents a random set of links between nodes in adjacent groups. To be more precise, for a fixed m, we define a random n-node graph $G(n, m)$. Let $x = \lceil n/m \rceil$. Partition the set of all nodes into subsets $S_1, \ldots, S_x$, of size m, (S_x of size at most m) called *supernodes*. Let $\mathcal{S} = \{S_1, \ldots, S_x\}$ be the set of all supernodes.

Let $L = \lfloor \log x \rfloor$. Arrange all supernodes into a binary tree T with $L+1$ levels $0, 1, \ldots, L$, placing each supernode S_i on level $\lfloor \log i \rfloor$. Level 0 contains the root and levels $L-1$ and L contain leaves of T. The supernode S_1, is the root of T.

For every $1 \leq i \leq \lfloor x/2 \rfloor$, S_{2i} is the left child of S_i and S_{2i+1} is the right child of S_i in T (S_{2i+1} exists if $x \geq 2i+1$). For every $1 < i \leq x$, supernode $S_{\lfloor i/2 \rfloor}$ is the parent of S_i. If a supernode is a parent or a child of another supernode, we say that these supernodes are adjacent in T.

The set of edges of $G(n,m)$ is defined as follows. If supernodes S_i and S_j are adjacent in T, then there is an edge in $G(n,m)$ between any node in S_i and any node in S_j with probability p_l. Moreover, supernodes have no interior links. The graph $G(n,m)$ is called a Random Binary Thick Tree (*RBTT*).

In the remainder of this section, we analyze *RBTT* and show that, if $p \leq 1/c \log n$, for some constant $c > 0$ to be defined below, then it supports communication with high probability in time $O(\log n)$. We consider the n-node *RBTT* with link probability $p_l = 1/18 \ln n$ and $m = \lceil 1152 \ln^2 n \rceil$ nodes per supernode, when spot probability is $p \leq 1/(768 \ln n)$. Hence, we take $c = 768/\ln 2$.

Let C_1 be the event that all supernodes in *RBTT* contain less than $6 \ln n + 1$ spots.

Lemma 1. *The event C_1 occurs with probability at least $1 - 1/n$.*

For a given constant $0 < \epsilon \leq 1$, let C_2 be the event that all supernodes in *RBTT* have more than $288(1-\epsilon)\ln^2 n$ functional nodes.

Lemma 2. *The event C_2 occurs with probability at least $1 - 1/n^{d \log n}$, for some positive constant d.*

Using the previous results, we now present two connectivity lemmas in preparation for the proof of the main theorem of this section.

Lemma 3. *All functional nodes are connected to at least one functional node in each supernode adjacent to their own, with probability exceeding $1 - 1/n^{13}$.*

Proof. Fix a node u. Let $N(u)$ denote the set of supernodes adjacent to the supernode containing u. Consider the event γ_{u,S_k} that u has a link to at least one functional node in a given supernode $S_k \in N(u)$. The event γ_{u,S_k} occurs unless all links from u to functional nodes in S_k do not exist. From Lemma 2, we get for any constants $0 < \epsilon', \epsilon'' \leq 1$

$$\begin{aligned}
\Pr[\gamma_{u,S_k}] &\geq \Pr[\gamma_{u,S_k} \wedge C_2] = \Pr[C_2] \Pr[\gamma_{u,S_k} \mid C_2] \\
&\geq \Pr[C_2] \left(1 - (1 - 1/18 \ln n)^{288(1-\epsilon')(1-\epsilon'') \ln^2 n}\right) \\
&\geq \left(1 - n^{-d \ln n}\right) \left(1 - n^{-16(1-\epsilon')(1-\epsilon'')}\right),
\end{aligned}$$

and hence $\Pr[\gamma_{u,S_k}] \geq 1 - n^{-15}$. Furthermore, since the graph contains at most n functional nodes, which should be connected to at least one functional node in at most 3 supernodes, the estimated probability is at least

$$\begin{aligned}
\Pr[(\forall u \in V \ (\forall S_k \in N(u)))\ \gamma_{u,S_k}] &\geq 1 - \sum_{u \in V} \sum_{S_k \in N(u)} \Pr[\neg \gamma_{u,S_k}] \\
&\geq 1 - 3nn^{-15} > 1 - n^{-13}.
\end{aligned}$$

□

Lemma 4. *All functional node pairs in supernodes at distance* 3 *are connected by a fault-free path with probability at least* $1 - 1/n^{1.9}$.

Proof. This lemma is proven in steps, defining connection probabilities and lower bounds on the number of connected nodes at distances 1, 2, and 3.

Fix 4 supernodes, S_u, S_i, S_j, S_k, which form a simple path in RBTT. I.e., S_u is adjacent to S_i, which is adjacent to S_j, which is adjacent to S_k.

Fix a node u in S_u. Let X_i be the random variable which counts the number of functional nodes $i \in \Gamma(u)$ located in S_i. From Lemma 2, each supernode contains more than $288(1-\epsilon)\ln^2 n$ fault-free nodes, for any $0 < \epsilon \leq 1$ with probability $1 - 1/n^{d \log n}$, for some positive constant d. Since the link probability is $p_l = 1/18 \ln n$,

$$E[X_i] \geq \Pr[C_2] \frac{288(1-\epsilon)\ln^2 n}{18 \ln n} = (1 - n^{-d \log n}) 16(1-\epsilon)\ln n \geq 16(1-\epsilon')\ln n,$$

with some $1 > \epsilon' > \epsilon$. We also have that a fixed functional node has at most $16(1 - \sqrt{3/8(1-\epsilon')})(1-\epsilon')\ln n$ such neighbors with probability

$$\Pr\left[X_i \leq 16\left(1 - \sqrt{\frac{3}{8(1-\epsilon')}}\right)(1-\epsilon')\ln n\right] \leq e^{-\left(\sqrt{\frac{3}{8(1-\epsilon')}}\right)^2 \frac{16(1-\epsilon')\ln n}{2}} = n^{-3}.$$

Let A be the event that node u has at least $16(1 - \sqrt{3/(8(1-\epsilon'))})(1-\epsilon')\ln n$ functional neighbors in S_i.

Assume event A occurs. Now, fix a node x in S_j. Fix a subset $S \subseteq \Gamma(u) \cap S_i$ of functional nodes, with size $16(1 - \sqrt{3/(8(1-\epsilon'))})(1-\epsilon')\ln n$. Denote by P_{Sx} the event that there exists a link between the node x and any node from S. This event occurs unless x has no link to some node in S. Hence,

$$\begin{aligned}\Pr[P_{Sx}|A] &= 1 - (1 - 1/18\ln n)^{16\left(1-\sqrt{3/(8(1-\epsilon'))}\right)(1-\epsilon')\ln n} \\ &\geq 1 - e^{-8\left(1-\sqrt{3/(8(1-\epsilon'))}\right)(1-\epsilon')/9} \geq 1/4\end{aligned}$$

for some small ϵ'.

Let X_j be the random variable which counts the number of functional nodes $j \in S_j$ which are adjacent to some node in S. We have that $E[X_j] \geq (1/4) \cdot 288(1-\epsilon')\ln^2 n$, assuming that A holds. Let B be the event that $X_j \geq 72(1-\epsilon'')\ln^2 n$, for some small $\epsilon'' > \epsilon'$. Since all events P_{Sx}, for fixed S and varying x, are independent, we use a Chernoff bound to show that, if event A occurs, event B occurs with probability $1 - 1/n^{k' \log n}$, for some positive constant k'.

Assume event $A \cap B$. Fix a functional node k in S_k. Fix a subset $S' \subseteq S_j$ of functional nodes, each of which is a neighbor of some element of S, with size $72(1-\epsilon'')\ln^2 n$. Denote by $P_{S'k}$ the event that there exists a link between node k and some node in S'. This event occurs unless k has no link to any node in S'. Hence,

$$\begin{aligned}\Pr[P_{S'k}|B \cap A] &= \left(1 - (1 - 1/18\ln n)^{72(1-\epsilon'')\ln^2 n}\right) \\ &\geq 1 - e^{-72(1-\epsilon'')\ln^2 n/(18\ln n)} \geq 1 - n^{-4(1-\epsilon'')}.\end{aligned}$$

Consider the event P_{uijk} that there exists a fault-free path of the form $uijk$ from a fixed node u to a fixed node k. Clearly, P_{uijk} is a subset of the event detailed in the above argument. Hence,

$$\begin{aligned}\Pr[P_{uijk}] &\geq \Pr[P_{S'k} \cap B \cap A] = \Pr[A]\Pr[B|A]\Pr[P_{S'k}|B \cap A] \\ &\geq \left(1 - n^{-3}\right)\left(1 - 1/n^{k' \log n}\right)\left(1 - n^{-4(1-\epsilon'')}\right) \geq \left(1 - n^{-3+\epsilon'''}\right),\end{aligned}$$

for some $0 < \epsilon''' < 0.1$.

There are at most n functional nodes in $RBTT$, each with $O(\log^2 n)$ other functional nodes in supernodes at distance 3. Hence, there are $O(n \log^2 n)$ functional node pairs in supernodes at distance 3. It follows that all node pairs in supernodes at distance 3 are connected with probability at least $1 - n^{-1.9}$. □

Combining the previous lemmas, we are now ready to prove Theorem 1.

Proof of Theorem 1. The $RBTT$ contains $O(n/\log^2 n)$ supernodes connected in a binary-tree structure of diameter $D \in O(\log n)$. It follows from the construction that the maximum degree of the $RBTT$ is $O(\log n)$, with high probability. By Lemma 4, all functional node pairs in supernodes at distance 3 are connected by at least one fault-free path of length 3 with probability greater than $1 - 1/n^{1.9}$. Therefore, all functional nodes in the subgraph $RBTT'$ composed of the root supernode S_1 and of all supernodes at distances multiple of 3 from S_1 are connected with this probability. Clearly, functional nodes not in $RBTT'$ are in supernodes adjacent to supernodes in $RBTT'$. Thus, by Lemma 3, all these functional nodes are also connected to at least one functional node in $RBTT'$ with probability exceeding $1 - 1/n^{13}$. Hence, with probability exceeding $1 - 1/n^{1.8}$, the fault-free part of $RBTT$ is connected.

We now investigate the diameter of the fault-free part of $RBTT$. From the above argument, we observe that 1) nodes in supernodes at distances multiple of 3 are connected with high probability by a path of length equal to the distance of the supernodes; 2) functional nodes in all other supernodes are connected with high probability by a path of length at most 2 longer than the distance of the supernodes. This leads to the conclusion that the diameter of the fault-free part of $RBTT$ is also in $O(\log n)$, with high probability. □

2.2 Lower Bounds

We have shown that it is possible to build a logarithmic-degree graph which supports communication with high probability in spite of spot probabilities $p \leq 1/c \log n$, for some positive constant c. The natural question then is whether it is possible to build arbitrarily large networks which can support communication with high probability despite larger spot probabilities. In what follows, we show that for constant spot probabilities, most networks do not have this property. More formally, the following theorem holds.

Theorem 2. *For any constant spot probability $p > 0$, there exist constants c_1 and c_2 such that: if all degrees in a graph are at most $c_1 \log n$ then the fault-free* *part of the graph is disconnected with high probability; if all degrees in a graph are at least $c_2 \log n$ then the graph has all nodes faulty with high probability. In either case, highly reliable communication is not possible.*

The preceding theorem leads to the conclusion that high probability communication is not possible, for a large class of graphs, when spot probability is a positive constant. However, the bounds $c_1 \log n$ and $c_2 \log n$ do not coincide. Since $c_1 < \frac{1}{\log(1/(p(1-p)))}$ and $c_2 > \frac{1}{\log(1/(1-p))}$, we have $c_1 < c_2$ for all positive values of p. It remains open whether or not there exists an arbitrarily large graph which supports reliable communication despite constant spot probabilities.

We will now attempt to provide insight into the question of what happens when node degrees lie between these bounds. For example, when $p = 1/2$, we have $c_1 < 1/2$ and $c_2 > 1$. Thus, with degree $\log n$, the important case of the n-node hypercube is not covered by Theorem 2. We will investigate this case in the following section.

2.3 Communication in the Hypercube

The hypercube H_k of dimension k is a 2^k-node graph with the set of nodes with identifiers from $\{0, 1\}^k$ and the set of links between nodes whose identifiers have a Hamming distance of 1. Hence the n-node hypercube H_k has dimension $\log n$.

Theorem 3. *The n-node hypercube H_k does not support high probability communication for any constant spot probability $0 < p \leq 1$.*

We first show that for constant $0 < p < 1/2$, the fault-free part of the graph is disconnected with high probability. We then show that for $1/2 < p \leq 1$, the graph has all nodes faulty with high probability, and that for $p = 1/2$, the graph has all nodes faulty with constant probability. This will prove Theorem 3.

3 Bounded Degree Networks

The RBTT presented in Section 2 remained connected despite relatively high spot probabilities. However, its degree is unbounded. For certain applications, smaller-degree networks may be preferred as they are easier to implement and give shorter communication time in the 1-port model. Therefore, it is natural to ask if bounded-degree networks can also support high-probability communication with comparable spot probabilities.

In this section we construct bounded-degree networks which tolerate inverse polynomial spot probabilities and which support high-probability communication with optimal time complexity. Furthermore, we prove that bounded-degree networks can tolerate at most inverse polynomial spot probabilities.

3.1 Upper Bounds

We now study the properties of two networks: the torus and a torus-based tree-like network that we call the *toroidal tree*. We show that the torus supports high-probability communication for spot probability in $1/\omega(n^{1/2})$. However, the diameter of the torus is quite large, which prohibits fast communication. Thus we design a tree-like structure based on the torus which provides the same fault-tolerance properties and supports communication in time $O(\log n)$, even in the 1-port model.

The Torus. In this section, we show an upper bound on the spot probability such that the fault-free part of the torus remains connected. Denote by $\mathcal{T}_{m\times k}$ the $m \times k$ torus with $m, k \geq 4$. The torus has the set of nodes $\{u = (u_x, u_y) : u_x \in \{0, 1, \ldots, m-1\}, u_y \in \{0, 1, \ldots, k-1\}\}$ and the set of links $\{(u, v) : |u_x - v_x| \bmod m + |u_y - v_y| \bmod k = 1\}$.

Theorem 4. *The fault-free part of the n-node torus $\mathcal{T}_{m\times k}$ is connected with high probability for $p \in 1/\omega(n^{1/2})$.*

The Toroidal Tree. We now design a network which provides the same fault-tolerance as the torus, while also providing optimal-order communication time for bounded-degree graphs. Since the diameter of a bounded-degree graph is at least logarithmic, our aim is to construct a network whose fault-free part has logarithmic diameter. Such a network supports highly reliable communication in optimal time $O(\log n)$, even in the 1-port model. The network construction is based on two binary trees, T and T', connected by a link between their root nodes. Each node of T, T' represents a group of nodes, and groups adjacent in T, T' have a subset of nodes in common. More precisely, for constant $k \geq 4$, we define a n-node graph $G(n, k)$. Assume that the set of nodes can be partitioned exactly as described below; this is easy to modify in the general case, by adding nodes to a leaf group.

Let the sets $\mathcal{T}_1, \ldots, \mathcal{T}_x$ and $\mathcal{T}'_1 \ldots, \mathcal{T}'_{x'}$ be tori with $2k$ rows $\{0, 1, \ldots 2k-1\}$ and k columns $\{0, 1, \ldots k-1\}$; $|x - x'| \leq 1$. We describe the construction for the tree T; the same construction is applied for the tree T'. Arrange all $\mathcal{T}_i$ as the nodes of T, with $L+1$ levels $0, 1, 2, \ldots, L$, placing each $\mathcal{T}_i$ on level $\lfloor \log i \rfloor$ of T. Level 0 contains the root of T and levels $L-1$ and L contain the leaves. For every $1 \leq i \leq \lfloor x/2 \rfloor$, $\mathcal{T}_{2i}$ is the left child of $\mathcal{T}_i$ in T and $\mathcal{T}_{2i+1}$ is the right child of $\mathcal{T}_i$ in T ($\mathcal{T}_{2i+1}$ exists if $x \geq 2i+1$). For every $1 < i \leq x$, $\mathcal{T}_{\lfloor i/2 \rfloor}$ is the parent of $\mathcal{T}_i$. Use row 0 of each child torus to connect it to its parent in T. Use row k of each parent torus to connect it to both its children in T. Use row 0 of both roots in T, T' to connect them together. Connections between tori adjacent in T, T' are done by identifying the respective rows.

It follows from the above description that $x + x' = \lfloor (n - 2k^2)/((2k-1)k) \rfloor + 1$ tori are located on $L = \lfloor \log(x + x' + 1) \rfloor$ levels in $G(n, k)$. The graph $G(n, k)$ is called a Toroidal Tree. It has bounded maximal degree.

Theorem 5. *For $p \in 1/\omega(n^{1/2})$, the n-node Toroidal Tree supports high probability communication in time $O(\log n)$.*

3.2 Lower Bounds

In this section, we show that bounded-degree graphs do not support high probability communication even for relatively small spot probabilities. We first show that the bounds on spot probability provided in Theorem 4 and Theorem 5 are tight for tori and toroidal trees. We then show that for general bounded-degree networks, if spot probability is the inverse of some polynomial, then high probability communication is not supported.

The Torus and Toroidal Tree. The following lower bounds match the upper bounds from Theorem 4 and Theorem 5, thus showing that the results are tight.

Theorem 6. *For spot probability $p \in 1/O(n^{1/2})$, the n-node torus $\mathcal{T}_{m\times k}$ does not support high probability communication.*

Theorem 7. *For spot probability $p \in 1/O(n^{1/2})$, the n-node Toroidal Tree does not support high probability communication.*

General Bounded Degree Graphs. We showed in the preceding section that in the case of the torus and the Toroidal Tree, spot probabilities at most $1/\omega(n^{1/2})$ can be tolerated if these graphs support high-probability communication. In the following theorem, we show that a similar phenomenon occurs for all graphs whose degree is bounded by a constant.

Theorem 8. *For spot probability $p \in 1/O(n^{1/d})$, no n-node graph with degree bounded above by $d \in \Theta(1)$ supports high probability communication.*

4 Conclusion

We provided what is, to the best of our knowledge, the first analytic results on fault-tolerance of networks in the presence of dependent, positively correlated faults. To do so, we introduced the *neighborhood fault* model where damaging events, called *spots*, occur randomly and independently at nodes of a network with probability p, and cause faults in the affected node and its neighbors.

We addressed questions regarding the connectivity and diameter of the fault-free part of networks in this fault model, as these characteristics of the network are responsible for the feasibility of communication and for its time. Our results show clear differences between the assumption of independent faults and that of the neighborhood fault model. For example, while under independent faults with small constant fault probability $p > 0$ the fault-free part of the hypercube remains connected with high probability [3], this is not the case under the neighborhood fault model with any positive constant spot probability. Likewise, the fault-free part of the torus is connected with high probability for fault probability $p \in 1/\Omega(n^{1/4})$ when faults are independent, but this is not the case for such spot probabilities under the neighborhood fault model.

It remains open whether or not there exists a network, which, under the neighborhood fault model, has the fault-free part connected with high probability despite constant spot probabilities. We conjecture that this is not the case.

The neighborhood fault model is the first step in modeling dependent positively correlated faults in networks. It would be interesting to analyze more precise center-satellite based models in which independent spots yield faults in nodes with probability decreasing with the distance of the node from the spot.

Acknowledgements. Evangelos Kranakis and Michel Paquette were supported by NSERC and MITACS. Andrzej Pelc was supported by NSERC and the Research Chair in Distributed Computing of the Université du Québec en Outaouais.

References

1. Bienstock, D.: Broadcasting with random faults. Discr. Appl. Math. 20, 1–7 (1988)
2. Chlebus, B.S., Diks, K., Pelc, A.: Sparse networks supporting efficient reliable broadcasting. Nordic Journal of Computing 1, 332–345 (1994)
3. Chlebus, B.S., Diks, K., Pelc, A.: Reliable broadcasting in hypercubes with random link and node failures. Comb., Prob. and Computing 5, 337–350 (1996)
4. Paquette, M., Pelc, A.: Fast broadcasting with byzantine faults. International Journal of Foundations of Computer Science 17(6), 1423–1439 (2006)
5. Pelc, A.: Fault-tolerant broadcasting and gossiping in communication networks. Networks 28(6), 143–156 (1996)
6. Ganesan, D., Govindan, R., Shenker, S., Estrin, D.: Highly-resilient, energy-efficient multipath routing in wireless sensor networks. ACM SIGMOBILE Mobile Computing and Communications Review 5(4), 11–25 (2001)
7. Thottan, M., Ji, C.: Using network fault predictions to enable IP traffic management. J. Network Syst. Manage 9(3), 327–346 (2001)
8. Yajnik, M., Kurose, J., Towsley, D.: Packet loss correlation in the MBone multicast network. Proceedings of IEEE Global Internet (May 27, 1996)
9. Stapper, C.H.: On yield, fault distributions and clustering of particles. IBM Journal of Research and Development 30(3), 326–338 (1986)
10. Meyer, F.J., Pradhan, D.K.: Modeling defect spatial distribution. IEEE Trans. Computers 38(4), 538–546 (1989)
11. Warren, W.: The center-satellite concept as a basis for ecological sampling. Stat. Ecol. 2, 87–118 (1971)
12. Blough, D.M., Pelc, A.: A clustered failure model for the memory array reconfiguration problem. IEEE Trans. Computers 42(5), 518–528 (1993)
13. Meyer, F.J., Park, N.: Predicting defect-tolerant yield in the embedded core context. IEEE Trans. Computers 52(11), 1470–1479 (2003)
14. Choi, A., Park, N., Meyer, F.J., Lombardi, F., Piuri, V.: Reliability measurement of fault-tolerant onboard memory system under fault clustering. In: Proceedings of 19th Instrumentation and Measurement Technology Conference, 2002. IMTC, vol. 2, pp. 1161–1166. IEEE Computer Society Press, Los Alamitos (2002)
15. Yu, F., Tsai, C.H., Huang, Y.W., Lee, D.T., Lin, H.Y., Kuo, S.Y.: Efficient exact spare allocation via boolean satisfiability. In: 20th IEEE Int. Symp. on Defect and Fault Tolerance in VLSI Systems (DFT'05), pp. 361–370. IEEE Computer Society Press, Los Alamitos (2005)