Fault Tolerance in Networks of Bounded Degree

Will Rosenbaum
Updated: March 15, 2015

Department of Mathematics
University of California, Los Angeles

This note serves as an executive summary of the results and techniques described in *Fault Tolerance in Networks of Bounded Degree* by Dwork, Peleg, Pippenger, and Upfal [2].

1 Overview

In *Fault Tolerance in Networks of Bounded Degree* [2], the authors consider a distributed computational model where some processors may be faulty. They model a distributed network as a graph $G = (V, E)$, where V is the (vertex) set of processors and E is the (edge) set of connections in the network. In this work, the authors explore the setting where the graph G has bounded degree. In particular, they are interested in the classical problem of Byzantine agreement (BA), where all processors wish to simply agree on a common value in $\{0, 1\}$.

If some processors $v \in V$ are faulty, it is natural to ask to what extent the non-faulty processors can still achieve Byzantine agreement. Unfortunately, for bounded degree networks, total agreement is impossible for even tolerating a modest number of faults in the network:

Theorem 1 (Dolev, 1982 [1])**.** Let t denote the number of faulty processors in a network $G = (V, E)$. Then Byzantine agreement cannot be achieved if G has connectivity less than $t + 1$.

In particular, Dolev's result implies that bounded degree networks can only tolerate a constant number of faults (the degree of the graph) for Byzantine agreement. Motivated by Dolev's result, the authors of [2] consider the following relaxation of Byzantine agreement, which we refer to as ***almost everywhere agreement*** (AEA). In AEA, we only require that a large fraction of processors agree on a common value in $\{0, 1\}$. Thus the author's goals are twofold:

1. design (bounded degree) networks for which AEA can be achieved;
2. design protocols which achieve AEA on the specified networks.

The primary strategy employed by the authors is to create sparse networks which can simulate an arbitrary protocol on the complete network. This allows one to lift *any* protocol for complete networks to a corresponding protocol for the specified bounded degree networks (possibly with some degradation in the performance of the protocol).

1.1 Results

For parameters t and X, the authors say that a protocol achieves t-***resilient*** X ***agreement*** if all but X processors in V achieve agreement in the presence of t failures. Using this notion of AEA, the authors prove the following results.

Theorem 2 (Main results)**.**

1. For all $r \geq 5$, almost all r-regular graphs admit t-resilient $O(t)$ agreement protocols for $t \leq n^{1-\varepsilon}$, where $\varepsilon(r) \to 0$ as $r \to n$.
2. The butterfly network* admits t-resilient $O(t \log t)$ agreement for $t \leq cn/\log n$.
3. There exists a network of degree 9 that admits t-resilient $O(t)$ agreement for $t \leq cn/\log n$.

* See Section 2.2

4. For every ε with $0 < \varepsilon < 1$, there exists $c = c(\varepsilon)$ and a graph of degree $O(n^c)$ and t resilient $O(t)$ agreement for $t \leq cn$.

5. If failures in the network can be authenticated, then for every $r \geq 5$ there exists $\varepsilon = \varepsilon(r)$ such that for all $t < \varepsilon n$, almost all r-regular graphs admit t-resilient $O(t)$ agreement.

In this note, we focus on the first three results, as the last two are of a decidedly different nature.

2 Simulation of the Complete Network

In this section, we describe the main technique employed in [2] to achieve AEA: simulation of the complete network. First, we set up notation. As before, $G = (V, E)$ is the network. Let P be a fixed protocol which simulates message passing on the complete network. That is, for any $u, v \in V$ and any message m, P allows u to send message m to v over the network G (in the absence of faults).

- T is the set of faulty processors, and $t = |T|$.
- Call $u, v \in V$ ***successful*** (for P) if whenever all $w \notin T$ follow the protocol P, any message sent from u to v succeeds.
- Let $\mathrm{Poor}(G, T)$ denote the minimum set of correct nodes such that every pair $u, v \notin T \cup \mathrm{Poor}(G, T)$ is successful.
- $p(G, t) = \max\{|\mathrm{Poor}(G, T)| \,|\, |T| = t\}$.

Theorem 3. Suppose a protocol P' on a complete graph tolerates $t + P(G, t)$ faults. Then the simulation of P' using P can tolerate t faults.

Proof sketch. Since all messages between $u, v \notin T \cup \mathrm{Poor}(G, T)$ are successful, only simulated messages involving $w \in T \cup \mathrm{Poor}(G, T)$ could be corrupted. The result follows since $|T \cup \mathrm{Poor}(G, T)| = t + p(G, t)$. □

2.1 **Transmission scheme strategy** The authors devise a class of protocols P (or ***transmission schemes***) that simulate communication on the complete network which they refer to as "three phase transmission schemes." The idea idea is that each vertex $v \in V$ specifies sets $\Gamma_{out}(v)$ and $\Gamma_{in}(v)$ along with specified paths from v to each w in $\Gamma_{out}(v)$ and $\Gamma_{in}(v)$. Further, for each pair $v, u \in V$, vertex disjoint paths from $\Gamma_{out}(v)$ to $\Gamma_{in}(u)$ are specified. Messages are sent from v to u in the following manner: the message is first broadcast from v to $\Gamma_{out}(v)$ (along the specified paths), then sent from $\Gamma_{out}(v)$ to $\Gamma_{in}(u)$, and finally from $\Gamma_{in}(u)$ to u. Since some messages may be corrupted, u takes the majority opinion of the messages received. Intuitively, this scheme will be successful in networks for which $\Gamma_{out}(v)$ and $\Gamma_{in}(v)$ are large (so that there is a lot of redundancy in the messages sent), but such that the specified paths are all small (to minimize the chances of a faulty transmission). Thus, we expect they strategy to succeed for expander graphs.

Let $s = |\Gamma_{out}(v)| = |\Gamma_{in}(v)|$. Call a vertex $v \in V$ ***out bad*** if at least $1/8$-th fraction of paths from v to $\Gamma_{out}(v)$ contain faulty processors, and similarly for ***in bad***. Let $\mathrm{Bad}(G, T)$ be the set of vertices which are out bad or in bad, and

$$b(G, t) = \max_{|T|=t} \{|\mathrm{Bad}(G, T|\}\,.$$

The following claim (whose proof is a simple counting argument) reveals that controlling the number of bad vertices is sufficient to bound poor vertices in a transmission scheme.

Claim 4. If $t < s/4$, then $p(G,t) \leq b(G,t)$.

2.2 **The butterfly network** The authors of [2] show that the following ***butterfly network*** admits a good three phase transmission scheme. Let $G_m = (V_m, E_m)$ be the network on $m2^m$ vertices where

$$V_m = \{(a,i) \,|\, 0 \leq a \leq m-1,\ 0 \leq i \leq 2^m - 1\}$$

and

$$((a,i),(b,j)) \in E_m \iff b \equiv a+1 \mod m \text{ and } i = j \text{ or } i \text{ and } j \text{ differ only in } a\text{th bit.}$$

For each $(a,i) \in V_m$, we take $\Gamma_{out}((a,i)) = \Gamma_{in}((a,i)) = \{(a,j) \,|\, j = 0,1,\ldots,2^m-1\}$. The paths from (a,i) to $\Gamma_{out}((a,i))$ are all "downward" edges from (a,i), as are edges from $\Gamma_{out}((a,i))$ to each $\Gamma_{in}((b,j))$.

Claim 5. For the butterfly network, $b(G_m,t) = O(t \log t)$.

Part 2 of Theorem 2 follows from Theorem 3 and Claims 4 and 5.

2.3 **Random regular graphs** Let $H(r,n)$ denote the set of r-regular graphs on n vertices. Part 1 of Theorem 2 follows from the following theorem about $H(r,n)$.

Theorem 6. Let $r \geq 5$ and choose $G \in H(r,n)$ uniformly at random. Then with high probability, G admits a three phase transmission scheme with $p(G,t) \leq O(t^{1+\delta} \log t)$ for $t \in O(n^{1-\varepsilon})$, where $0 < \delta < \varepsilon < 1$ and $\varepsilon = \varepsilon(r) \to 0$ as $r \to \infty$.

The proof of Theorem 6 comes from two lemmas concerning expansion and super-concentration properties of random regular graphs.

Lemma 7 (Expansion)**.** For all $r \geq 5$, there exists $\alpha = \alpha(r)$, $0 < \alpha < 1$ such that with high probability G satisfies the following property: every $U \subseteq V$ with $|U| \leq \alpha n$ has at least $|U|(r-3)$ neighbors outside of U.

Lemma 8 (Super-concentration)**.** Suppose G has the property that all $U \subseteq V$ with $|U| \leq \alpha n$ has at least $2|U|$ neighbors outside U. Then every two subsets U and W of V with $|U| = |V| \leq \alpha n$ are connected by $|U|$ vertex disjoint paths.

The proof of the expansion lemma uses the probabilistic method. The super-concentration lemma appeals to a variant of Menger's theorem, which characterizes the number of vertex disjoint paths between sets in a graph.

2.4 **Byzantine agreement on compressor graphs** A graph G is called a ***θ-compressor*** if for every subset $U \subseteq V$ of vertices with $|U| \leq \theta n$, there are at most $|U|/2$ vertices which have at least half their neighbors in U.

For compressor graphs, the authors of [2] consider the following protocol for Byzantine agreement. In each round, all processors:

1. send their current value to their neighbors;
2. receive values from all neighbors;
3. choose new values based on the majority message received.

Lemma 9. Suppose $|T| = t \leq \theta n/2$. If there are $(1-\theta)n$ correct processors sharing the same initial value x, then after $\log n$ iterations, at most $t-1$ correct processors will have a value different from x.

To prove Lemma 9, the authors show that for $A_k = ((2^k-1)t + |V_0|)/2^k$, $A_k + t \leq \theta n$ and $|V_k| \leq A_k$. Here V_k is the set of vertices with values other than x after k rounds. The lemma follows by taking $k = \log n$.

Lemma 10. For $r \geq 5$, there exists θ with $0 < \theta < 1$ such that a random graph $G \in H(r, n)$ is a θ-compressor with high probability.

Part 3 of Theorem 2 follows from Lemmas 9 and 10 and the following construction. Let G be the 9-regular graph on $m2^m$ vertices whose edge set is a disjoint union of the butterfly network and a compressor graph. The butterfly network can be used to achieve $O(t \log t)$ agreement, while the compressor network can be applied to sharpen the agreement to $O(t)$.

3 Open(?) Questions

The authors leave the following open questions:

1. Optimize the communication cost of protocols for almost every agreement.
2. Construct protocols which do not require processors to know the global topology of the network.
3. Find protocols for t-resilient $O(t)$ agreement for $t \in \omega(n/\log n)$ (and in particular $t = \Omega(n)$), or prove that no such protocol exists.

Sources

[1] Danny Dolev. The byzantine generals strike again. *Journal of algorithms*, 3(1):14–30, 1982.

[2] Cynthia Dwork, David Peleg, Nicholas Pippenger, and Eli Upfal. Fault tolerance in networks of bounded degree. *SIAM Journal on Computing*, 17(5):975–988, 1988.