Spatial Gossip and Resource Location Protocols

David Kempe* Jon Kleinberg † Alan Demers‡

Abstract

The dynamic behavior of a network in which information is changing continuously over time requires robust and efficient mechanisms for keeping nodes updated about new information. *Gossip protocols* are mechanisms for this task in which nodes communicate with one another according to some underlying deterministic or randomized algorithm, exchanging information in each communication step. In a variety of contexts, the use of randomization to propagate information has been found to provide better reliability and scalability than more regimented deterministic approaches.

In many settings — consider a network of sensors, or a cluster of distributed computing hosts — new information is generated at individual nodes, and is most "interesting" to nodes that are nearby. Thus, we propose distance-based *propagation bounds* as a performance measure for gossip algorithms: a node at distance d from the origin of a new piece of information should be able to learn about this information with a delay that grows slowly with d, and is *independent* of the size of the network.

For nodes arranged with uniform density in Euclidean space, we present natural gossip algorithms that satisfy such a guarantee: new information is spread to nodes at distance d, with high probability, in $O(\log^{1+\varepsilon} d)$ time steps. Such a bound combines the desirable qualitative features of *uniform gossip*, in which information is spread with a delay that is logarithmic in the full network size, and *deterministic flooding*, in which information is spread with a delay that is linear in the distance and independent of the network size. Our algorithms and their analysis resolve a conjecture of Demers et al. We show an application of our gossip algorithms to a basic *resource location problem*, in which nodes seek to rapidly learn of the nearest copy of a *resource* in a network.

*Department of Computer Science, Cornell University, Ithaca NY 14853. Email: kempe@cs.cornell.edu.
Supported by an NSF Graduate Fellowship.

†Department of Computer Science, Cornell University, Ithaca NY 14853. Email: kleinber@cs.cornell.edu.
Supported in part by a David and Lucile Packard Foundation Fellowship, an ONR Young Investigator Award, NSF ITR/IM Grant IIS-0081334, and NSF Faculty Early Career Development Award CCR-9701399.

‡Department of Computer Science, Cornell University, Ithaca NY 14853. Email: ademers@cs.cornell.edu.

1 Introduction

Gossip algorithms

The dynamic behavior of a network in which information is changing continuously over time requires robust and efficient mechanisms for keeping nodes updated about new information. For example, we may have a network of sensors measuring properties of the physical world, and performing computations on the collective set of measurements in a distributed fashion (see e.g. [4, 6, 8, 9]). As measured values change, we would like for them to be propagated through the network rapidly. Or we may have a distributed network of computing hosts that need to be informed about significant changes in the load on machines, or the appearance of new *resources* in the network [15]; again, we would like such information to be spread quickly through the network.

For reasons of reliability and scalability, we do not want central control for updates to reside with a small set of nodes. Rather, we seek mechanisms by which nodes communicate in a relatively homogeneous fashion with one another, so as to spread information updates. *Gossip protocols* [3, 7] are mechanisms of this type in which nodes communicate with one another according to some underlying deterministic or randomized algorithm, exchanging information in each communication step. In a variety of contexts, the use of randomization for such tasks has been found to provide better reliability and scalability than more regimented deterministic approaches (see e.g. [1, 2, 3, 6, 13, 15, 16]).

It is useful to consider some of the issues that arise in a very simple version of our first example. Suppose we have a network of N sensors positioned at the lattice points of a $\sqrt{N} \times \sqrt{N}$ region of the plane, monitoring conditions about the underlying environment. We assume that there is an underlying mechanism that supports an abstraction of point-to-point communication: in a single virtual "step," any node can send a message to any other node, regardless of the distance separating them in the plane. All the algorithms we develop here are built on top of such a point-to-point communication mechanism; for our purposes, the fact that this point-to-point communication may actually be implemented by multi-hop transmission of packets is abstracted away.

Here is a basic problem we may wish to solve in such an environment: If a sensor node x detects abnormal conditions, it will generate an alarm message m that needs to be *propagated* to all other nodes in the network. Consider the following two different approaches to this problem.

Uniform gossip. In each step, each node u chooses a node v uniformly at random, and forwards to v all the alarm messages it knows about. A well-known result states that with high probability, all nodes will receive a copy of a given message m within $O(\log N)$ steps of its initial appearance [7, 10, 14].

Neighbor flooding. In each step, each node u chooses one of its closest neighbors v in the plane, according to a round-robin ordering, and forwards to v all the alarm messages it knows about. Clearly, any node v that is at distance d from the origin of a message m will receive a forwarded copy of m within $O(d)$ steps of its initial appearance. However, the time it takes for all nodes to obtain a given message under this scheme is $\Theta(\sqrt{N})$.

Both of these algorithms follow the gossip paradigm: in each time step u picks some other node v (either deterministically or at random) and communicates with v. We will refer to this as u *calling* v. Moreover, both algorithms are very simple, since each node is essentially following the same local rule in each time step, independently of message contents. In the protocols we consider, the analysis will use only information that u passes to the node v it calls, ignoring any information that u may obtain from v. In other words, all our protocols work in the *push* model of communication.

In this discussion, it is crucial to distinguish between two conceptual "layers" of protocol design: (i) a basic gossip *algorithm*, by which nodes choose other nodes for (point-to-point) communication; and (ii) a gossip-based *protocol* built on top of a gossip algorithm, which determines the contents of the messages that are sent, and the way in which these messages cause nodes to update their internal states. We can view a gossip algorithm as generating a labeled graph $\mathcal{H}$ on the N nodes of the network; if u communicates with v at time t, we insert an edge (u, v) with label t into $\mathcal{H}$. When we consider more complex gossip-based protocols below, it is useful to think of a run of the underlying gossip algorithm as simply generating a labeled graph $\mathcal{H}$ on which the protocol then operates.

The propagation time

The two algorithms discussed above have quite different performance bounds, with uniform gossip "filling in" the set of points exponentially faster than neighbor flooding. The neighbor flooding algorithm, however, exhibits a desirable feature that uniform gossip is lacking: *messages are propagated to nodes with a delay that depends only on their distance from the origin of the message, not on the total number of nodes in the system.* In the examples above, and in applications exhibiting any kind of spatial locality, this can be very important: when an alarm is triggered, we may well want to alert nearby nodes earlier than nodes further away; similarly, as resources appear in a network, we may want nodes to learn more quickly of the resources that are closer to them. With uniform gossip, it is likely that a node adjacent to the source of an alarm will only be alerted after news of the alarm has traveled extensively through the network.

Our work was initially motivated by the following question: Is there a gossip algorithm — preferably a simple one — that exhibits the best qualitative features of both uniform gossip and neighbor flooding, by guaranteeing that a message can be propagated to any node at distance d from its originator, with high probability, in time bounded by a polynomial in $\log d$? The crucial point is that such a bound would be poly-logarithmic, yet independent of N. To make this question precise, we introduce the following definition. We will say that a function $f_{\mathcal{A}}(\cdot)$ is a *propagation time* for a given gossip algorithm $\mathcal{A}$ if it has the following property: Whenever a new piece of information is introduced at u, there is a high probability[1] that some sequence of communications will be able to forward it to v within $O(f_{\mathcal{A}}(d))$ steps.

Our question can now be phrased as follows: Is there a gossip algorithm with a propagation time that is polynomial in $\log d$?

[1] When we write "with high probability" here, we mean with probability at least $1-(\log(d+1))^{-\kappa}$, where κ may appear in the constant of the expression $O(f_{\mathcal{A}}(d))$.

A gossip algorithm with good propagation time

Our first result is an affirmative answer to this question. Rather than the lattice points of the plane, we consider a more general setting — that of a point set with *uniform density* in $\mathbf{R}^D$. We will make this notion precise in the next section; essentially, it is a point set in $\mathbf{R}^D$ with the property that every unit ball contains $\Theta(1)$ points.

Theorem 1.1 *Let P be a set of points with uniform density in $\mathbf{R}^D$. For every $\varepsilon > 0$, there is a randomized gossip algorithm on the points in P with a propagation time that is $O(\log^{1+\varepsilon} d)$.*

In this Theorem, the $O(\cdot)$ includes terms depending on ε and the dimension D. However, the propagation bound is independent of the number of points in P, and in fact holds even for infinite P.

The algorithm achieving this bound is equivalent to one proposed by Demers et al. [3], who considered it in the context of concerns different from ours. We fix an exponent ρ strictly between 1 and 2. In each round, each node u chooses to communicate with a node v with probability proportional to $d_{u,v}^{-D\rho}$, where $d_{u,v}$ denotes the distance between u and v. Demers et al. had conjectured that this algorithm would propagate a single message to all nodes in a D-dimensional grid of side length N with high probability in time polynomial in $\log N$; a special case of Theorem 1.1 (obtained by choosing a finite grid of side length N for our metric space) yields a proof of this conjecture.

We also show how our algorithms and their analysis can be used to provide a partial resolution to open questions about the behavior of *Astrolabe*, a network resource location service developed by van Renesse [15]. Astrolabe relies on an underlying gossip mechanism, and by generalizing the setting in which we cast our algorithms, we are able to give bounds on the rate at which messages spread through this system.

Alarm-spreading and resource location

Following our discussion above, the inverse-polynomial gossip algorithm provides a "transport mechanism" on which to run a variety of protocols. Perhaps the most basic application is a simple version of the alarm-spreading example we discussed at the outset. Suppose that at each point in time, each node u can be in one of two possible states: `safe` or `alarm`. In each time step, u calls another node v according to an underlying gossip algorithm, and transmits its current state. If u is in the `alarm` state, then v will also enter the `alarm` state when it is called by u. All nodes remain in the `alarm` state once they enter it. By using the inverse-polynomial gossip algorithm from Theorem 1.1, we obtain the following guarantee for this protocol: if x is a node in P, and some node at distance d from x undergoes a transition to the `alarm` state at time t, then with high probability x will enter the `alarm` state by time $t + O(\log^{1+\varepsilon} d)$.

A more interesting problem than the simple spreading of an alarm is that of *resource location*. As before, we have a set of points P with uniform density. As time passes, nodes may acquire a copy of a *resource*. (For example, certain nodes in a sensor network may be receiving data from an external source; or certain nodes in a cluster of computers may be running the server component of a client-server application.) At any given time, each node

u should know the identity of a resource-holder (approximately) closest to it; we wish to establish performance guarantees asserting that u will rapidly learn of such a resource.

If at every time step, all nodes forward the names of all resource holders they know about, then the propagation bounds from Theorem 1.1 immediately imply that nodes will learn of their closest resource within time $O(\log^{1+\varepsilon} d)$, where d is the distance to the closest resource. The disadvantage of this protocol is that the message sizes grow arbitrarily large as more resources appear in the network.

Resource location: Bounding the message size

Naturally, protocols that require the exchange of very large messages are not of significant interest for practical purposes. For problems in which there are natural ways of aggregating information generated at individual nodes, it is reasonable to hope that strong guarantees can be obtained by protocols that use messages of bounded size (or messages that contain a bounded number of node names).

We focus on the above *resource location* problem as a fundamental problem in which to explore the power of gossip-based protocols that use bounded messages, in conjunction with the algorithm from Theorem 1.1.

The problem has both a *monotone* and a *non-monotone* variant. In the *monotone* version, a node never loses its copy of the resource once it becomes a resource-holder. For this version of the problem, we consider the following simple gossip protocol. Each node u maintains the identity of the closest resource-holder it knows about. In a given time step, u calls a node v according to the gossip algorithm from Theorem 1.1 and transmits the identity of this resource-holder. Finally, the nodes update the closest resource-holders they know about based on this new information. Note that the protocol only involves transmitting the name of a single node in each communication step; moreover, u transmits the same value regardless of the identity of the node it calls.

Despite its simplicity, we can show that this protocol satisfies strong performance guarantees in the monotone case.

- In one dimension, it has the following property. Let r and u be two nodes at distance d. If r holds the closest copy of the resource to u during the time interval $[t, t']$, and if $t' - t \geq \Omega(\log^{1+\varepsilon} d)$, then with high probability u will learn about r by time t'.
- In higher dimensions, it has the following approximate guarantee. Again, let r and u be two nodes at distance d. If r acquires a resources at time t, then with high probability node u will know of a resource-holder within distance $d + o(d)$ by time $t + O(\log^{1+\varepsilon} d)$.

In the the *non-monotone* version of the problem, a node may lose a resource it previously held. Designing gossip protocols in this case is more difficult, since they must satisfy both a *positive requirement* — that nodes rapidly learn of nearby resources — and a *negative requirement* — that nodes rapidly discard names of nodes that no longer hold a resource. We provide an algorithm for the one-dimensional case, establishing that precise formulations of the positive and negative requirements can be maintained with a delay that is polynomial in $\log d$. Weaker versions of the positive and negative requirements, incorporating an approximation guarantee, can be obtained in higher dimensions.

2 Spatial Gossip

Definitions

We will present our gossip algorithms and analysis for nodes positioned at points in $\mathbf{R}^D$; for two such nodes x and y, we define their distance $d_{x,y}$ using any L_k metric. Below, we will discuss how the results can be generalized to other settings.

Let $B_{x,d} = \{y \mid d_{x,y} \leq d\}$ denote the ball around x of radius d. We say that an (infinite) set of points P in $\mathbf{R}^D$ has *uniform density*, with parameters β_1 and β_2, if every ball of radius $d \geq 1$ contains between $\beta_1 d^D$ and $\beta_2 d^D$ points of P. (This includes balls not centered at points of P.) As stated, our definition only makes sense for infinite point sets. However, we can easily extend it to finite point sets, and our algorithms and analysis apply to this case as well with essentially no modifications. For simplicity, however, we will focus on infinite point sets here.

As we discussed in the introduction, the gossip algorithms we study are designed to produce "communication histories" with good propagation behavior. A useful model for stating and proving properties of such communication histories is that of temporal networks and strictly time-respecting paths in them, proposed in [5, 11] as a means to describe how information spreads through a network over time. A *directed temporal network* is a pair (G, λ), where $G = (V, E)$ is a directed graph (possibly with parallel edges), and $\lambda : E \to \mathbb{N}$ a time-labeling of edges. A *strictly time-respecting path* $P = e_1, \ldots, e_k$ from u to v is a path in G such that $\lambda(e_i) < \lambda(e_{i+1})$ for all $1 \leq i < k$. Hence, strictly time-respecting u-v paths are exactly those paths along which information from u can reach v.

For a given run $\mathcal{R}$ of a gossip algorithm $\mathcal{A}$, the associated temporal network $\mathcal{H}_{\mathcal{R}}$ is the pair $((V, E), \lambda)$, where V is the set of all nodes in the system, and E contains an edge $e = (u, v)$ labeled $\lambda(e) = t$ if and only if u called v at time t in $\mathcal{R}$. Note that there may be an infinite number of parallel copies of the edge (u, v), with different labels — however, no two parallel copies can have the same label. For a subset $V' \subseteq V$ of nodes, and a time interval I, we use $\mathcal{H}_{\mathcal{R},V',I}$ to denote the temporal network $((V', E'), \lambda|_{E'})$, with $E' = \{e \in E \cap (V' \times V') \mid \lambda(e) \in I\}$. That is, $\mathcal{H}_{\mathcal{R},V',I}$ is the communication history of the run $\mathcal{R}$, restricted to a specific time interval and a specific group of participating nodes.

Throughout, ln denotes the natural logarithm, and ld the base-2 logarithm.

The Gossip Algorithms

We consider gossip algorithms based on inverse-polynomial probability distributions. The algorithms are parameterized by an exponent ρ satisfying $1 < \rho < 2$. Let $x \neq y$ be two nodes at distance $d = d_{x,y}$, and let $p_{x,y}^{(\rho)}$ denote the probability that x calls y. Then, we let $p_{x,y}^{(\rho)} := c_x(d+1)^{-D\rho}$, i.e. the probability that y is called decreases polynomially in the distance between x and y. c_x is chosen such that $\sum_y p_{x,y}^{(\rho)} \leq 1$. c_x might be the normalizing constant for the distribution; however, we want to retain the freedom to choose c_x smaller, to model the fact that messages might get lost with constant probability. We make the restriction that $c := \inf_x c_x$ is strictly greater than 0. We denote the resulting gossip algorithm by $\mathcal{A}^\rho$.

Let us quickly verify that the probability distribution is indeed well-defined at each point x, i.e. that c_x can be chosen strictly greater than 0. Notice that the total probability mass at point x is at most

$$\int_{z=0}^{\infty} D\beta_2 z^{D-1} \cdot (z+1)^{-D\rho} dz$$
$$\leq \quad D\beta_2 \int_{z=1}^{\infty} z^{D(1-\rho)-1} dz < \infty,$$

because $\rho > 1$ and $D > 0$.

The main result of this section is a proof of Theorem 1.1, giving poly-logarithmic bounds on the propagation time of $\mathcal{A}^\rho$ when $1 < \rho < 2$. We state the result in the following form.

Theorem 2.1 *Fix a ρ with $1 < \rho < 2$, and define the function $f_{\mathcal{A}^\rho}(d) = \alpha \cdot (\,\mathrm{ld}\,(d+1))^{\frac{1}{1-\mathrm{ld}\,(\rho)}}\,\mathrm{ld}\,\mathrm{ld}\,(d+1)$. Then, α can be chosen such that for every ball B of diameter $d > 1$, every time t, and nodes $x, x' \in B$, the temporal network $\mathcal{H}_{\mathcal{R},B,[t,t+\kappa f_{\mathcal{A}^\rho}(d))}$ contains a strictly time-respecting x-x' path with probability at least $1 - (\,\mathrm{ld}\,(d+1))^{-\kappa}$.*

Intuitively, this theorem states that information from any node x can reach any node x' with high probability within time poly-logarithmic in their distance. Also, it guarantees that the information stays within a small region containing both x and x' on its way from x to x'.

The proof is by induction on the distance between the nodes x and x'. We will show that during a certain time interval I, a node u close to x will call a node u' close to x', and then we will use induction to guarantee paths from x to u before I, and from u' to x' after I. The following definitions formally capture the random events that we are interested in.

For given x and x', let $\mathcal{E}_{\mathcal{H},x,x'}$ denote the event that the temporal network $\mathcal{H}$ contains a strictly time-respecting x-x' path. Let $\hat{E}(U, U', I)$ be the set of all potentially available labeled edges from U to U' with labels from the interval I; that is, $\hat{E}(U, U', I)$ contains one element for each (u, u', t), where $u \in U, u' \in U'$, and $t \in I$ is a time. Let $\prec$ be any total order on $\hat{E}(U, U', I)$. Let $\mathcal{F}_{\mathcal{H},\hat{E},e}$ denote the event that e is the smallest edge from $\hat{E}$ (with respect to $\prec$) contained in $\mathcal{H}$, and $\mathcal{F}_{\mathcal{H},\hat{E},\perp}$ the event that $\mathcal{H}$ contains no edge from $\hat{E}$. Notice that the events $\mathcal{F}_{\mathcal{H},\hat{E},e}$ are disjoint for different e, and $\mathcal{F}_{\mathcal{H},\hat{E},\perp} = \overline{\bigcup_{e \in \hat{E}} \mathcal{F}_{\mathcal{H},\hat{E},e}}$.

Let $\sigma = \frac{1}{2}\min\{\beta_1, 1\} \cdot 4^{-D}$. First, we prove a lemma bounding from above the probability of the event $\mathcal{F}_{\mathcal{H},\hat{E},\perp}$.

Lemma 2.2 *Let S and S' be sets of size at least $\sigma(k+1)^{\frac{D\rho}{2}}$ both contained in a ball B_k of radius at most k, and let I be a time interval of length at least $\frac{2^{D\rho}}{c\sigma^2}\ln\tau$. Let $\hat{E} = \hat{E}(S, S', I)$, and $\mathcal{H} = \mathcal{H}_{\mathcal{R},B_k,I}$. Then, $Pr[\mathcal{F}_{\mathcal{H},\hat{E},\perp}] \leq 1/\tau$.*

Proof. Any two points $u \in S$ and $u' \in S'$ are at distance at most $2k$, so at any time $t \in I$, u calls u' with probability at least $c_u(2k+1)^{-D\rho} \geq c(2k+2)^{-D\rho}$. Hence, for any fixed $u \in S$ and $t \in I$, $\mathcal{H}$ contains a u-S' edge labeled t with probability at least $c|S'|(2k+2)^{-D\rho}$.

As the random choices are independent for all $u \in S$ and $t \in I$, the probability $\Pr[\mathcal{F}_{\mathcal{H},\hat{E},\perp}]$ that $\mathcal{H}$ contains no S-S' edge e with label $\lambda(e) \in I$ is at most

$$(1 - c|S'|(2k+2)^{-D\rho})^{|S||I|} \leq e^{-c|S'||S||I|(2k+2)^{-D\rho}}$$
$$\leq e^{-c\sigma^2 2^{-D\rho}\frac{2^{D\rho}}{c\sigma^2}\ln\tau} = 1/\tau.$$

■

The central (inductive) idea explained above is captured in the following Lemma. We let $r = \frac{1}{1-\operatorname{ld}(\rho)}$ be the (claimed) exponent in the propagation bound, and define $g(k) := \max\{0, 2(\operatorname{ld}(k+1))^r - 1\}$. $\gamma < 1$ is a constant (to be determined only for the proof of Theorem 2.1), and $\eta = \frac{2^{D\rho}}{c\sigma^2} \operatorname{ld} \frac{1}{\gamma}$.

Lemma 2.3 *Let x and x' be at distance k, B_k a ball of radius k containing x and x', and $[t, t')$ a time interval of length at least $\eta g(k)$. Let $\mathcal{H}_k = \mathcal{H}_{\mathcal{R},B_k,[t,t')}$ be the corresponding temporal network. Then, the probability of the event $\mathcal{E}_{\mathcal{H}_k,x,x'}$ that there is a strictly time-respecting x-x' path in $\mathcal{H}_k$ is at least $1 - \gamma g(k)$.*

Proof. For $k = 0$, the ball B_k consists only of one point x, so $x = x'$, and as the empty path is a time-respecting x-x' path, such a path exists with probability $1 = 1 - \gamma g(0)$.

For $k \geq 1$, we let $k' = \lfloor k^{\rho/2} - 1 \rfloor$ (notice that $k' < k$), and B and B' balls of radius k' with $x \in B \subseteq B_k$, and $x' \in B' \subseteq B_k$. Let $s = t + \eta g(k')$, $s' = s + \eta$, $I = [s, s')$, $\hat{E} = \hat{E}(B, B', I)$, and $\mathcal{H} = \mathcal{H}_{\mathcal{R},B,[t,s)}, \mathcal{H}' = \mathcal{H}_{\mathcal{R},B',[s',t')}$.

Because $s - t = \eta g(k')$, we can apply the induction hypothesis to $\mathcal{H}$, and conclude that $\Pr[\mathcal{E}_{\mathcal{H},x,u}] \geq 1 - \gamma g(k')$, for any $u \in B$. Using the assumption that $t' - t \geq \eta g(k)$, we obtain that $t' - s' = t' - t - \eta g(k') - \eta \geq \eta(g(k) - g(k') - 1)$. If $k' = 0$, then $2g(k') + 1 = 1 \leq g(k)$, and otherwise,

$$2g(k') + 1 \leq 4(\operatorname{ld} k^{\rho/2})^r - 2 + 1 = 2(\operatorname{ld} k)^r - 1 \leq g(k).$$

In either case, $2g(k') + 1 \leq g(k)$, and therefore $t' - s' \geq \eta g(k')$, so we can apply the induction hypothesis to $\mathcal{H}'$ as well, to show that $\Pr[\mathcal{E}_{\mathcal{H}',u',x'}] \geq 1 - \gamma g(k')$, for any $u' \in B'$.

Whenever there is an edge $e = (u, u') \in \hat{E}$, and strictly time-respecting paths P and P' from x to u in $\mathcal{H}$ and from u' to x' in $\mathcal{H}'$, the concatenated path PeP' is a strictly time-respecting x-x' path in $\mathcal{H}_k$. Because the time-intervals $[t, s)$, $[s, s')$ and $[s', t')$ are disjoint, and all random choices made by $\mathcal{A}^\rho$ are independent, the three events $\mathcal{E}_{\mathcal{H},x,u}$, $\mathcal{F}_{\mathcal{H}_k,\hat{E},e}$ and $\mathcal{E}_{\mathcal{H}',u',x'}$ are independent for fixed x, x' and $e = (u, u')$. Using that the events $\mathcal{F}_{\mathcal{H}_k,\hat{E},e}$ are disjoint for distinct e, we obtain that

$$\begin{aligned}
&\Pr[\mathcal{E}_{\mathcal{H}_k,x,x'}] \\
&\quad\geq \sum_{e=(u,u')\in\hat{E}} \Pr[\mathcal{E}_{\mathcal{H},x,u} \cap \mathcal{F}_{\mathcal{H}_k,\hat{E},e} \cap \mathcal{E}_{\mathcal{H}',u',x'}] \\
&\quad= \sum_{e=(u,u')\in\hat{E}} \Pr[\mathcal{E}_{\mathcal{H},x,u}] \cdot \Pr[\mathcal{E}_{\mathcal{H}',u',x'}] \cdot \Pr[\mathcal{F}_{\mathcal{H}_k,\hat{E},e}] \\
&\quad\geq \sum_{e=(u,u')\in\hat{E}} (1 - \gamma g(k')) \cdot (1 - \gamma g(k')) \cdot \Pr[\mathcal{F}_{\mathcal{H}_k,\hat{E},e}] \\
&\quad= (1 - \gamma g(k'))^2 \cdot (1 - \Pr[\mathcal{F}_{\mathcal{H}_k,\hat{E},\perp}]) \\
&\quad\geq 1 - 2\gamma g(k') - \Pr[\mathcal{F}_{\mathcal{H}_k,\hat{E},\perp}]
\end{aligned}$$

If $k' \geq 1$, then $k' = \lfloor k^{\rho/2} - 1 \rfloor \geq \frac{1}{4}(k^{\rho/2}+1) \geq \frac{1}{4}(k+1)^{\rho/2}$, so both B and B' contain at least $\beta_1(k')^D \geq \beta_1 4^{-D}(k+1)^{\frac{D\rho}{2}}$ points. Otherwise, $k' = 0$, and B and B' contain exactly one point (x or x', resp.). As $\lfloor k^{\rho/2} - 1 \rfloor = 0$, we know that $k^{\rho/2} < 2$, so $(k+1)^{\rho/2} \leq k^{\rho/2} + 1 < 3$, and $4^{-D}(k+1)^{\frac{D\rho}{2}} \leq 1$. Therefore, B and B' contain at least $\frac{1}{2}\min\{\beta_1, 1\} \cdot 4^{-D}(k+1)^{\frac{D\rho}{2}}$ points in either case. Because I has size $|I| \geq \frac{\sigma}{c} \operatorname{ld}\frac{1}{\gamma} \geq \frac{\sigma}{c}\ln\frac{1}{\gamma}$, we can apply Lemma 2.2, with B, B', B_k, and I, to obtain that $\Pr[\mathcal{F}_{\mathcal{H}_k,\hat{E},\perp}] \leq \gamma$. Using $2g(k')+1 \leq g(k)$ (which we showed above), we now conclude that $\Pr[\mathcal{E}_{\mathcal{H}_k,x,x'}] \geq 1 - \gamma g(k)$, completing the proof. ■

The Theorem now follows quite easily.

Proof of Theorem 2.1. Choose $\gamma = (\operatorname{ld}(d+1))^{-r-\kappa}$, $\alpha = 4r \cdot \frac{2^{D\rho}}{c\sigma^2}$, and let $t' = t + \kappa f_{\mathcal{A}^\rho}(d)$.

By substituting the definitions of the function g and the constants, we can verify that $t' - t = \kappa f_{\mathcal{A}^\rho}(d) \geq \eta g(d)$. (In this calculation, we use that $2\kappa r \geq \kappa + r$, which holds because $r \geq 1$ and $\kappa \geq 1$.) We can therefore apply the above Lemma 2.3 to x and x' (and thus with $k = d$), to obtain that with probability at least $1 - \gamma(\operatorname{ld}(d+1))^r = 1 - (\operatorname{ld}(d+1))^{-\kappa}$, the temporal network $\mathcal{H}_{\mathcal{R},\hat{B},[t,t')}$ contains a strictly time-respecting x-x' path, completing the proof. ■

We can get rid of the $\operatorname{ld}\operatorname{ld}(d+1)$ term by increasing the exponent r slightly.

A more general setting

In the proofs of Theorem 2.1 and Lemma 2.3, we only used relatively few properties of the metric space $\mathbf{R}^D$. In fact, we can generalize the results to hold for point sets without an underlying metric, provided only that we have an appropriate notion of what a "ball" is. Specifically, let X be a set of points, β and $\mu > 1$ two designated constants, and $\mathcal{D}$ a collection of finite subsets of X, called *discs*, that satisfy the following axioms.

1. For every $x \in X$, there is a disc $D \in \mathcal{D}$ with $x \in D$.
2. For any two discs D_1, D_2, there is a disc $D \supseteq D_1 \cup D_2$.
3. If $D_1, D_2, \ldots$ are discs with $|D_i| \leq b$ and $x \in D_i$ for all i, then there exists a disc D with $D_i \subseteq D$ for all i, and $|D| \leq \beta \cdot b$.
4. If D is a disc with $x \in D$ and $|D| > 1$, then there is a disc $D' \subseteq D$ with $x \in D'$ and $|D| > |D'| \geq \frac{1}{\mu}|D|$.

We can think of the collection of discs as serving the role of balls in a metric space with point set X. With this interpretation, the above axioms are satisfied by point sets of uniform density in $\mathbf{R}^D$, with the constant β chosen to be exponential in D. However, they are also satisfied by certain other natural metric spaces and set systems, including a version of van Renesse's *Astrolabe* system [15] that we discuss below.

For a given disc collection, we can define probabilities $p^{(\rho)}$ for the inverse polynomial gossip algorithms $\mathcal{A}^\rho$. For points $x, y \in X$, let b be the minimum cardinality of any disc $D \in \mathcal{D}$ containing both x and y. By the first two axioms, such a disc exists, and therefore, b is well-defined. Now, define $p^{(\rho)}_{x,y} := c_x b^{-\rho}$, where c_x is again a normalizing constant at point

x. From the third axiom, we obtain that for any point x, at most βb points y can lie in a disc D that also contains x and has size $|D| \leq b$. Therefore, at most βb points can contribute $b^{-\rho}$ or more to the total probability mass at point x.

Hence, the total probability mass $\sum_{y \neq x} p^{(\rho)}_{x,y}$ at any point x is at most $\sum_{b=1}^{\infty}(\beta b - \beta(b-1)) \cdot b^{-\rho} = \beta \sum_{b=1}^{\infty} b^{-\rho} \leq \frac{\beta\rho}{\rho-1}$, and in particular finite, so the distribution is well-defined. For this distribution, we can state the propagation guarantee of $\mathcal{A}^{\rho}$ as follows.

Theorem 2.4 *Let $1 < \rho < 2$. Let $x, x' \in X$ be two points, and D a disc of size b containing both x and x'. Then, the temporal network $\mathcal{H}_{\mathcal{R},D,[t,t+\kappa f_{\mathcal{A}^{\rho}}(b))}$ contains a strictly time-respecting x-x' path with probability at least $1 - (\operatorname{ld}(b+1))^{-\kappa}$, for every time t.*

In other words, x' will learn of information originating at x with a propagation delay that is poly-logarithmic in the size of the smallest disc containing both of them. The proof of this Theorem closely follows the proof of Theorem 2.1, and we omit it here.

Other values of the parameter ρ

It is natural to ask about the behavior of our inverse-polynomial gossip algorithms for different values of the exponent ρ. We have found that the set of possible values for ρ can be divided into three parts with qualitatively distinct behavior. For $\rho \leq 1$, we do not actually obtain a well-defined probability distribution for infinite point sets; for a finite point set, the propagation time cannot be bounded as a function of the distance alone (i.e. independent of the size of the system). For $1 < \rho < 2$, we have the behavior analyzed above; and for $\rho > 2$ we prove that the propagation time is at least $\Omega(d^{\varepsilon})$ for an exponent $\varepsilon > 0$ depending on ρ.

This leaves only the "transitional case" $\rho = 2$, which turns out to have very interesting behavior; we are able to prove that for every $\varepsilon > 0$, the gossip algorithm $\mathcal{A}^{\rho}$ with $\rho = 2$ has a propagation time that is $O(d^{\varepsilon})$, but do not know whether it has a propagation time that is polynomial in $\log d$. We note that all these results for varying values of ρ also hold in the more general setting of disc collections, with bounds in terms of distances replaced by bounds in terms of disc sizes.

Astrolabe

Astrolabe is a network resource location service that uses a gossip mechanism for spreading information [15]; we refer the reader to this paper for more detail than we are able to provide here. One reasonable model of the structure of an Astrolabe system is as follows: computing nodes are positioned at the leaves of a uniform-depth rooted tree T of constant internal node degree; there is an underlying mechanism allowing for point-to-point communication among these leaf nodes. It is desirable that information originating at a leaf node x should be propagated more rapidly to leaf nodes that share lower common ancestors with x than to those that share higher common ancestors.

The gossip mechanism in Astrolabe can be modeled using disc collections as follows: The underlying point set X is equal to the leaves of the tree T, and there is a disc D corresponding to the leaves of each rooted subtree of T. It is easy to verify that the four axioms described above hold for this collection of discs. Hence, if leaf nodes communicate according to the gossip algorithm in Theorem 2.4, then a piece of information originating at a leaf node x

will spread to all the leaves of a k-node subtree containing x, with high probability, in time polynomial in $\log k$. For reasons of scalability, the gossip algorithm actually used in the Astrolabe system is essentially equivalent to $\mathcal{A}^\rho$ with $\rho = 2$; and so, by the result mentioned above, information will spread through a k-leaf subtree, with high probability, in time $O(k^\varepsilon)$ for every $\varepsilon > 0$.

3 Resource Location Protocols

In the introduction, we discussed the basic *resource location problem*. We have nodes in $\mathbf{R}^D$ as before; as time passes, nodes may acquire copies of a *resource*, and we wish for each node to rapidly learn the identity of a resource-holder (approximately) closest to it. (By abuse of terminology, we will sometimes interchange the terms "resource-holder" and "resource.")

Our protocols for resource location will rely on an underlying algorithm for gossip on point sets of uniform density in $\mathbf{R}^D$; we do not make any assumptions about this algorithm other than versions of the probabilistic propagation guarantee which we proved $\mathcal{A}^\rho$ to have in the previous section. The most basic required guarantee can be expressed as follows:

> There is a monotonically non-decreasing time-bound $f_\mathcal{A}(d)$ such that for any ball B of diameter d, any two nodes $x, x' \in B$, and any time t, the temporal network $\mathcal{H}_{\mathcal{R},B,[t,t+f_\mathcal{A}(d))}$ contains a time-respecting path from x to x' with high probability.

This asserts that information from a source x with high probability reaches a destination x' "sufficiently fast" via a path not involving any node too far away from either x or x' (as can be seen by choosing B to be the smallest ball containing both x and x'). The nature of the "high probability" guarantee may depend on the algorithm, and will directly affect the guarantees provided by the resource location protocol. In the preceding section, we established that this guarantee holds for inverse polynomial distributions with exponents $\rho \in (1, 2)$, with $f_\mathcal{A}(d) \in O(\kappa \log^{1+\varepsilon} d)$, giving us high probability guarantee $1 - (\,\mathrm{ld}\,(d+1))^{-\kappa}$.

Monotone Resource Location

We begin by considering the *monotone* case, in which any node that is a resource-holder at time t is a resource-holder for all $t' \geq t$. Our protocol should guarantee that within "short time", a node learns of its closest resource (once it becomes available), and subsequently will never believe any resource further away to be its closest resource. An approximation guarantee would be to require that a node learns of a resource that is "not too much further" away from it than its closest resource.

A simple protocol for the line

If resources never disappear, and all points lie on a line (i.e. the dimension is $D = 1$), a very simple protocol will ensure that each node learns of its closest resource "quickly" (as quantified by $f_\mathcal{A}(d)$), with high probability.

The protocol is as follows: Each node x locally maintains the node $N_x(t)$, the closest node which x knows to hold a copy of the resource at time t. Initially, at time $t = 0$, $N_x(0)$

is set to a null state $\perp$. When a resource appears at a node x at time t, $N_x(t') = x$ for all $t' \geq t$. In each round t, each node x selects a communication partner according to the algorithm $\mathcal{A}$, and sends the value $N_x(t)$. Let M_t be the set of all messages that x received in a round t. Then, it updates $N_x(t+1)$ to be the closest node in $M_t \cup \{N_x(t)\}$ (ties broken in favor of $N_x(t)$, and otherwise arbitrarily).

For the line, we can prove the following strong guarantee about the performance of this protocol:

Theorem 3.1 *Let x be any node, and x_R a resource at distance $d = d_{x,x_R}$. Let $t' = t + \kappa f_{\mathcal{A}}(d)$, and assume that x_R was the (unique) closest resource to x throughout the interval $[t, t']$. Then, $N_x(t') = x_R$ with high probability,*

Proof. For the proof, let B be the smallest interval containing both x_R and x, and consider the temporal network $(G, \lambda) = \mathcal{H}_{\mathcal{R},B,[t,t')}$. The guarantee of the underlying algorithm $\mathcal{A}$ states that with high probability, (G, λ) contains a time-respecting path from x_R to x. Let $x_R = v_1, \ldots, v_k = x$ be the vertices on any such time-respecting path, and $e_1 = (v_1, v_2), \ldots, e_{k-1} = (v_{k-1}, v_k)$ its edges with time labels $t_i = \lambda(e_i)$. Let x_i be the message that was sent from v_i to v_{i+1} at time t_i.

By induction, we will establish that $x_i = x_R$. For $i = 1$, this is clearly true. For the inductive step from i to $i + 1$, consider the message x_i received by v_{i+1} at time t_i, and the message x_{i+1} sent by v_{i+1} at time t_{i+1}. Because v_{i+1} lies in the smallest interval containing x_R and x, v_{i+1} lies on a shortest path from x_R to x. Therefore, x_R being the closest resource to x throughout $[t, t']$ implies that it is also the closest resource to v_{i+1}. By the choices of the protocol and the induction hypothesis, $N_{v_{i+1}}(t_i) = x_R$, and as x_R is still the closest resource to v_{i+1} at time t_{i+1}, we obtain $x_{i+1} = N_{v_{i+1}}(t_{i+1}) = x_R$.

At time t_{k-1}, node x receives the message $x_{k-1} = x_R$, and as $t \leq t_{k-1} < t'$, x_R is the closest resource to x from time t_{k-1} until t', so $N_x(t') = x_R$. ■

Analysis of the protocol in higher dimensions

In higher dimensions, this simple protocol may not inform nodes of their truly closest resource as quickly as $\kappa f_{\mathcal{A}}(d)$. Intuitively, we want the message about a node x_R with a resource to quickly reach every node x such that x_R is the closest resource to x. That is, we are interested in "filling in" the Voronoi region of the node x_R. However, if the Voronoi region is very long and narrow, most calls made by nodes inside the region will be to nodes outside the region. Hence, the time depends on the angles of the corners of the Voronoi region.

If we wish to make guarantees avoiding such specific properties of the actual distribution of resources, we can obtain an approximation guarantee by slightly strengthening the requirement on the algorithm $\mathcal{A}$. Our stronger requirement will be that not only is there a strictly time-respecting path, but its total "length" is bounded by some function of the distance d under consideration. Intuitively, this means that messages with high probability do not take very long "detours" on their way from a source to a destination.

More formally, for a path P with vertices $v_1, \ldots, v_k$, let its *path distance* be $d(P) = \sum_{i=1}^{k-1} d_{v_i,v_{i+1}}$. Then, we can state the requirement as:

There is a time-bound $f_{\mathcal{A}}(d)$ and a length function $\ell_{\mathcal{A}}(d)$ such that for any ball B of diameter d, any two nodes $x, x' \in B$, and any time t, $\mathcal{H}_{\mathcal{R},B,[t,t+\kappa f_{\mathcal{A}}(d))}$ contains a time-respecting path P from x to x' of path distance $d(P) \leq \ell_{\mathcal{A}}(d)$, with high probability.

Below, we prove that the algorithms $\mathcal{A}^\rho$ with $\rho \in (1,2)$ satisfy this property with $\ell_{\mathcal{A}}(d) = d + o(d)$.

We can now state the approximation guarantee of the resource location protocol that we previously analyzed for line:

Theorem 3.2 *Let x be any node, and x_R a resource at distance $d = d_{x,x_R}$. Let $t' = t + \kappa f_{\mathcal{A}}(d)$, and $x'_R = N_x(t')$. Then, $d_{x,x'_R} \leq \ell_{\mathcal{A}}(d)$ with probability at least $1 - (\operatorname{ld}(d+1))^{-\kappa}$.*

Hence, for sufficiently large d, the inverse polynomial gossip algorithms $\mathcal{A}^\rho$ will guarantee a $(1 + o(1))$-approximation to the closest resource within poly-logarithmic time, with high probability.

Proof. Let B be any smallest ball containing both x_R and x, and consider the temporal network $(G, \lambda) = \mathcal{H}_{\mathcal{R},B,[t,t')}$. With high probability, this network contains a strictly time-respecting x_R-x path P of path distance $d(P) \leq \ell_{\mathcal{A}}(d)$. Let $x_R = v_1, \ldots, v_k = x$ be the vertices of this path, $e_i = (v_i, v_{i+1})$ its edges with labels $t_i = \lambda(e_i)$, and x_i the message sent from v_i to v_{i+1} at time t_i.

We prove by induction that for all i, $d_{v_i,x_i} \leq \sum_{j=1}^{i-1} d_{v_j,v_{j+1}}$. Clearly, this holds for $i = 1$, since $x_1 = x_R = v_1$. For the step from i to $i+1$, notice that $t_i < t_{i+1}$. The protocol then ensures $d_{v_{i+1},x_{i+1}} = d_{v_{i+1},N_{v_{i+1}}(t_{i+1})} \leq d_{v_{i+1},N_{v_{i+1}}(t_i+1)} \leq d_{v_{i+1},x_i}$. By the triangle inequality, $d_{v_{i+1},x_i} \leq d_{v_{i+1},v_i} + d_{v_i,x_i}$, and applying the induction hypothesis to i yields that

$$d_{v_{i+1},x_{i+1}} \leq d_{v_{i+1},v_i} + \sum_{j=1}^{i-1} d_{v_j,v_{j+1}} = \sum_{j=1}^{i} d_{v_j,v_{j+1}}.$$

Using the behavior of the protocol at node $x = v_k$ and time t_k, we know that

$$d_{x,x'_R} \leq d_{x,v_{k-1}} + \sum_{j=1}^{k-2} d_{v_j,v_{j+1}} = d(P) \leq \ell_{\mathcal{A}}(d),$$

completing the proof. ∎

By taking a closer look at the analysis in the proof of Theorem 2.1, we can show that the algorithms $\mathcal{A}^\rho$ (for $1 < \rho < 2$) have the claimed property of short paths with $\ell_{\mathcal{A}^\rho}(d) = d + o(d)$.

Lemma 3.3 *Let B be a ball with diameter d, t an arbitrary time, and $t' = t + \kappa f_{\mathcal{A}^\rho}(d)$. Then, with probability at least $1 - (\operatorname{ld}(d+1))^{-\kappa}$, the temporal network $(G, \lambda) = \mathcal{H}_{\mathcal{R},B,[t,t')}$ contains a strictly time-respecting path $\hat{P}$ from x to x' of path distance at most $d + o(d)$ for any two nodes $x, x' \in B$.*

Proof. The path $\hat{P}$ constructed in the proof of Theorem 2.1 consists of one "jump" edge of length at most d, and two subpaths P and P' which lie inside balls B and B' of diameter at most $\lfloor d^{\rho/2} - 1 \rfloor \leq d^{\rho/2}$, and are of the same form. Hence, we obtain the recurrence $\ell_{\mathcal{A}^\rho}(d) \leq d + 2\ell_{\mathcal{A}^\rho}(d^{\rho/2})$. Using standard substitution techniques (and writing $n = \frac{\mathrm{ld}\,\mathrm{ld}\, d}{1 - \mathrm{ld}\,\rho}$), we find that this recurrence has the solution

$$\ell_{\mathcal{A}^\rho}(d) = \sum_{k=0}^{n} 2^{n-k+(\frac{2}{\rho})^k}.$$

We can bound this from above by splitting off the last term ($k = n$) of the sum, and bounding all other terms from above by $2^{n+(\frac{2}{\rho})^{n-1}}$, obtaining that $\ell_{\mathcal{A}^\rho}(d) \leq d + (\,\mathrm{ld}\, d)^{\frac{1}{1-\mathrm{ld}\,(\rho)}} \frac{\mathrm{ld}\,\mathrm{ld}\, d}{1-\mathrm{ld}\,\rho} \cdot d^{\rho/2}$, which is bounded by $d + o(d)$, as claimed. ■

There is an alternate way to obtain approximation guarantees for resource location in higher dimensions, without strengthening the assumptions on the underlying gossip algorithm. This is done by having nodes send larger messages in each time step.

For a scaling parameter $\xi > 1$, a node x at time t stores the identity of the closest resource-holder x_R, and a set $R_x(t)$ consisting of all resource-holders x'_R that x has heard about whose distance to x is at most $\xi \cdot d_{x,x_R}$. In each time step, nodes communicate their sets $R_x(t)$, and then update them based on any new information they receive.

Here is a more precise description of the protocol: The local state of any node x consists of a set $R_x(t)$ of nodes which x knows to hold a copy of the resource at time t. Initially, $R_x(0) = \emptyset$, and whenever a resource appears at a node x at time t, $x \in R_x(t')$ for all $t' \geq t$.

In each round t, each node x selects a communication partner according to the algorithm $\mathcal{A}$, and sends the entire set $R_x(t)$. Let M_t be the set of all gossip messages that x received in a round t, and $M = (\bigcup_{m \in M_t} m) \cup R_x(t)$ the set of all resources about which x knows at time t. Let $d = \min_{y \in M} d_{x,y}$. Then, x updates $R_x(t+1) := \{y \in M \mid d_{x,y} \leq \xi d\}$, i.e. to be the set of all nodes no more than ξ times as far away from x as the nearest resource.

The sets $R_x(t)$ sent and locally stored could potentially be large, but if we believe the resources to be spaced relatively evenly, the local storage and messages should only contain a constant number of nodes (for constant ξ). This protocol yields the following approximation guarantee:

Theorem 3.4 *Let x be any node, and x_R a resource at distance $d = d_{x,x_R}$. Let $t' = t + \kappa f_{\mathcal{A}}(d)$, and x'_R be a node in $R_x(t')$ with minimal distance to x (among all nodes in $R_x(t')$). Then, $d_{x,x'_R} \leq (1 + \frac{2}{\xi - 1})d$ with high probability.*

Hence, this Theorem shows how we can smoothly trade off message size against better approximation guarantees. Notice that the runtime of the gossip protocol is not directly affected by the desired better guarantees (only via the larger message size).

Proof. Let B be a smallest ball containing both x and x_R. Consider the temporal network $(G, \lambda) = \mathcal{H}_{\mathcal{R},B,[t,t')}$. With high probability, this network contains a strictly time-respecting path P from x_R to x. Let $v_1, \ldots, v_k$ be the vertices of this path, $e_i = (v_i, v_{i+1})$ its edges with labels $t_i = \lambda(e_i)$, and m_i the message sent from v_i to v_{i+1} at time t_i.

We prove by induction that for all i, $R_{v_i}(t_{i-1}+1)$ contains a resource at distance at most $(1+\frac{2}{\xi-1})d$ from x. With $i=k$, this will clearly imply the theorem.

The claim holds for $i=1$, since $x_R \in R_{v_1}(t)$ and $d_{x,x_R}=d$. For the step from i to $i+1$, let x_i be the node in $m_i = R_{v_i}(t_i)$ closest to v_i, and consider two cases:

1. If $x_i \in m_{i+1} = R_{v_{i+1}}(t_{i+1})$, then the claim holds for $i+1$ because it held for i by hypothesis.

2. If $x_i \notin m_{i+1}$, then m_{i+1} must contain a node x_{i+1} such that $d_{v_{i+1},x_i} > \xi d_{v_{i+1},x_{i+1}}$ (otherwise, v_{i+1} would have retained x_i). By the triangle inequality, $d_{v_{i+1},x_i} \leq d + d_{x,x_i}$, and using the induction hypothesis to bound the distance between x and x_i, we get

$$\begin{aligned} d_{v_{i+1},x_{i+1}} &\leq \frac{d_{v_{i+1},x_i}}{\xi} \leq \frac{d+d_{x,x_i}}{\xi} \\ &\leq \frac{d+(1+\frac{2}{\xi-1})d}{\xi} = \frac{2}{\xi-1}d, \end{aligned}$$

so $d_{x,x_{i+1}} \leq d + \frac{2}{\xi-1}d = (1+\frac{2}{\xi-1})d$. ■

Non-Monotone Resource Location

The situation becomes more complex if resources may disappear over time. In that case, information about the disappearance needs to propagate through the system as well, to ensure that nodes do not store outdated information. However, we do not want to send messages for every disappearing resource, since this would again increase the size of messages too much.

Rather, we use a time-out scheme to ensure that nodes find out about the disappearance of resources implicitly. That is to say, when a node x has not heard about its closest resource x_R for a sufficiently long time, x concludes that x_R is no longer a resource-holder; x stops sending information about x_R, and becomes receptive to learning about new resources even if they are further away than the one previously considered closest. To implement the timing mechanism that we referred to above, we will assume that each node has access to the global time t.

Of course, it is crucial to state what the time-out function $h(\cdot)$ should be. We still want nodes to find out about their (approximately) closest resources in time depending solely on their distance from the resource. However, we now also want to require that nodes find out about the disappearance of their closest resource within similar time bounds, in particular in time depending only poly-logarithmically on their distance from the resource. That is, $h(\cdot)$ should be a function of the distance d, but not the size of the underlying node set.

In view of this requirement, the existence of time-respecting paths might not be sufficient to ensure that information about a resource actually reaches the desired destination. We have not imposed any bounds on the amount of time that may lie between the labels of two adjacent edges of the path (i.e. the time information spends at one node), and if this time is too long, the node may "time out" on the resource, i.e. decide that it does not hold a resource any more. We therefore want to require the existence of "time-out free" paths.

For a time-respecting path $P = v_1, \ldots, v_k$, with edges $e_i = (v_i, v_{i+1})$, the *departure time* from node $i < k$ is $\delta(v_i) = \lambda(e_i)$. A time-respecting path P is called *time-out free* (with respect to a time-out function $h(d)$) if $\delta(v_i) - \delta(v_1) \leq h(d_{v_1,v_i})$ for all nodes v_i on the path. We let $\mathcal{T}_{\mathcal{H}_k,x,x'}$ denote the event that the temporal network $\mathcal{H}_k$ contains a time-out free x-x' path. Now, the requirement on the underlying protocol can be stated as follows (where we write $r = \frac{1}{1-\operatorname{ld}(\rho)}$, as before):

> There is a non-decreasing time-out function $h(d)$ such that for any ball B of diameter d, any two nodes $x, x' \in B$ and time t, the temporal network $\mathcal{H}_{\mathcal{R},B,[t,t+\kappa h(d))}$ contains a time-out free, strictly time-respecting path from x to x' with probability at least $\frac{1}{2}(\operatorname{ld}(d+1))^{-r}$, i.e. $\Pr[\mathcal{T}_{\mathcal{H}_{\mathcal{R},B,[t,t+\kappa h(d))}}] \geq \frac{1}{2}(\operatorname{ld}(d+1))^{-r}$.

Below, we will show that the inverse polynomial gossip algorithms $\mathcal{A}^\rho$ satisfy this property with time-out function $h(d) = O(f_{\mathcal{A}^\rho}(d))$. First, however, we will see how we can exploit this property to build a protocol for the non-monotone resource location problem.

Notice that if $h'(d) \geq h(d)$ for all d, then a path that is time-out free with respect to $h(\cdot)$ is also time-out free with respect to $h'(\cdot)$. We can therefore always choose the time-out function larger, without decreasing the probability. We will use this fact to design a protocol with high-probability guarantees, even though the above guarantee is only low-probability in itself.

The local state $S_x(t)$ of a node x at time t is either the set $\{\langle x_R, \tau\rangle\}$ consisting of a single time-stamped message containing the name of some node x_R and the time-stamp τ, or the empty set $\emptyset$. If $S_x(t) = \{\langle x_R, \tau\rangle\}$, then x_R is x's current estimate of the closest resource-holder; we will say that x *believes in* x_R *at time* t. We say that a node x *times out* on x_R at time t if x believes in x_R at time t, and x does not believe in x_R at time $t+1$. Using the time-out function $h(\cdot)$, we define the time-out function $h'(d) = h(d) \cdot \kappa \cdot (\frac{1}{2}\operatorname{ld}(d+1))^r \operatorname{ld}\operatorname{ld}(d+1)$ for the protocol, where κ is again a measure of the desired probability guarantee.

Each node executes the following protocol:

- If x holds a copy of the resource at time t, then its local state is set to $S_x(t) := \{\langle x, t\rangle\}$.
- Otherwise, let M_t be the set of all messages received at time $t-1$, plus the previous state $S_x(t-1)$. If M_t contains a message $\langle x', \tau\rangle$ with $t - \tau \leq h'(d_{x,x'})$, let $\hat{x}$ be such that $\hat{x} \neq x$, and $\hat{x}$ minimizes $d_{x,x'}$ among all such messages. Then, let $\hat{\tau}$ be maximal such that $\langle \hat{x}, \hat{\tau}\rangle \in M_t$, and set $S_x(t) := \{\langle \hat{x}, \hat{\tau}\rangle\}$. If M_t contains no such message $\langle x', \tau\rangle$, set $S_x(t) := \emptyset$.
- Send $S_x(t)$ to a node y chosen according to $\mathcal{A}$.

We will show that on the line, this protocol ensures that nodes learn quickly about both appearance and disappearance of their closest resource, in the following sense.

Theorem 3.5 *Let t be an arbitrary time, x and x_R nodes at distance $d = d_{x,x_R}$, $t' = t+h'(d)$, and $t'' = t + 2h'(d)$.*

(1) If x_R did not hold a copy of the resource during any of the interval $[t, t']$, then x does not believe in x_R at time t'.

(2) If x_R has held the copy of the resource uniquely closest to x throughout the interval $[t, t'']$, then x believes in x_R at time t'' with probability at least $1 - (\operatorname{ld}(d+1))^{-\kappa}$.

In terms of the analysis of systems, the negative (first) condition corresponds to *safety*, i.e. ensuring that wrong or outdated information will not be held for too long, while the positive (second) condition corresponds to *liveness*, i.e. ensuring that information will eventually reach any node. Notice that the safety guarantee is in fact deterministic, while the liveness guarantee is probabilistic.

Proof. Property (1) follows directly. If x_R was not a resource-holder during any of the interval $[t, t']$, then no messages $\langle x_R, \tau\rangle$ for $\tau \in [t, t']$ were generated, so no such message can have reached x.

The remainder of the proof is concerned with Property (2). Let B be the smallest interval containing x and x_R, and $t_j = t' + j \cdot h(d)$, where $h(\cdot)$ is the original time-out function, and j ranges from 0 to $\kappa \cdot (\frac{1}{2}\operatorname{ld}(d+1))^r \operatorname{ld}\operatorname{ld}(d+1)$. Now, fix one such j, and the associated interval $I = [t_j, t_{j+1})$ of length $h(d)$.

By assumption on $\mathcal{A}$, there is a time-out free x_R-x path $P = v_1, \ldots, v_k$ (with $v_1 = x_R$ and $v_k = x$) in the temporal network $\mathcal{H}_{\mathcal{R},B,I}$ with probability at least $\frac{1}{2}(\operatorname{ld}(d+1))^{\frac{-1}{1-\operatorname{ld}(\rho)}}$. Let us suppose that such a time-out free path does exist. Let $e_i = (v_i, v_{i+1})$, and m_i the message sent along e_i at time $\lambda(e_i)$. We will show by induction that $m_i = \langle x_R, \tau\rangle$ for some $\tau \geq \lambda(e_1) \geq t_j$, for all i.

In the base case $i = 1$, this is obvious, since x_R was assumed to hold a resource at time $\delta(v_1) \in I$. For the inductive step from i to $i+1$, we know by induction hypothesis that $m_i = \langle x_R, \tau\rangle$, with $\tau \geq \lambda(e_1) = \delta(v_1)$. There could be two "obstacles" to v_{i+1} sending a message m_{i+1} of the same form: (1) messages about other resources closer to v_{i+1} than x_R, and (2) v_{i+1} timing out on x_R at some time $s \in [\delta(v_i), \delta(v_{i+1})]$.

For (1), notice that we assumed x_R to be the unique closest resource to x throughout I. As v_{i+1} lies on a shortest x_R-x path (here, it is crucial that the points are on the line), x_R is also closest to v_{i+1} throughout I. Hence, we can apply the safety property (1) proved above to obtain that at no time $s \in I$, v_{i+1} believes in any x' closer to v_{i+1} than x_R.

For (2), recall that P is a time-out free path, and therefore satisfies $s - \tau \leq s - \delta(x_R) \leq h'(d_{v_{i+1},x_R})$ for all $s \in [\delta(v_i), \delta(v_{i+1})]$. In the protocol, message m_i is therefore always available as a candidate for the next state of v_{i+1}, and since we argued above that no messages for x' closer than x_R are available, the next state is of the form $\langle x_R, \tau'\rangle$ (where $\tau' \geq \tau$). Hence, message m_{i+1} is actually of the form $\langle x_R, \tau'\rangle$ with $\tau' \geq \tau \geq \delta(x_R)$.

Applying this to $v_k = x$, we obtain that at time t_{j+1}, node x believes in x_R. The time-out function $h'(\cdot)$ is so large that for any $s \in [t', t'']$, node x cannot time out on x_R if it ever received a message from x_R with a time-stamp $\tau \geq t'$. That is, if there is a time-out free, strictly time-respecting x_R-x path in the temporal network $\mathcal{H}_{\mathcal{R},B,[t_j,t_{j+1})}$ for any j, then x believes in x_R at time t''. Because the intervals $[t_j, t_{j+1})$ are disjoint for different j, and all random choices made during the protocol are independent, the probability that none of the intervals contain a time-out free path is at most

$$\begin{aligned}
& (1 - (\frac{1}{2}\operatorname{ld}(d+1))^{-r})^{\kappa(\frac{1}{2}\operatorname{ld}(d+1))^r \operatorname{ld}\operatorname{ld}(d+1)} \\
\leq\ & e^{-(\frac{1}{2}\operatorname{ld}(d+1))^{-r}\kappa(\frac{1}{2}\operatorname{ld}(d+1))^r \operatorname{ld}\operatorname{ld}(d+1)} \\
=\ & (\operatorname{ld}(d+1))^{-\kappa},
\end{aligned}$$

completing the proof. ■

In higher dimensions, we can obtain similar approximation bounds to the ones in the monotone case, by requiring that no resources within distance $d + o(d)$ of x disappear in the time interval under consideration, where d is again the distance of node x to its closest resource. The proof is a direct combination of the proofs of Theorems 3.5 and 3.2, and therefore omitted here.

Time-out free paths for $\mathcal{A}^\rho$

In the remainder of this section, we argue that the inverse polynomial gossip algorithms $\mathcal{A}^\rho$ from the Section 2 actually satisfy the above property of producing time-out free paths, with time-out function $h(d) = O(\operatorname{ld}^{1+\varepsilon}(d+1))$. Hence, the protocol for non-monotone resource location presented above will have time-out function and dissemination time bound $O(\operatorname{ld}^{2+\varepsilon}(d+1))$.

As in Section 2, we define the constants $\sigma = \frac{1}{2}\min\{\beta_1, 1\} \cdot 4^{-D}$ and $\alpha = 2r \cdot \frac{2^{D\rho}}{c\sigma^2}$. If we restrict our attention to balls B of diameter at most 64, and make the allowed time interval sufficiently large, we can ensure that there will be a time-respecting x-x' path within B with probability at least $\frac{1}{2}$. If nothing else, x calls x' with constant non-zero probability in every round, and the one edge induced by this call is a time-respecting and time-out free path. By making enough independent trials, the probability that one will succeed (and x will call x') will become at least $\frac{1}{2}$. Let Δ such that for any x and x' at distance at most 64, x will call x' within Δ rounds with probability at least $\frac{1}{2}$. Let $g(k) := (4\alpha + \Delta)(\operatorname{ld}(k+1))^r \operatorname{ld}\operatorname{ld}(k+3)$, and define $h(d) = g\left(\left((\frac{2\beta_2}{\beta_1})^{\frac{1}{D}} \cdot (d+1)\right)^{\frac{2}{\rho}}\right)$ as our time-out function. Notice that $h(d) \in O(\operatorname{ld}^r(d+1)\operatorname{ld}\operatorname{ld}(d+1))$.

Lemma 3.6 *Let x and x' be at distance k, B_k a ball of diameter k containing x and x', and $t' = t + g(k)$. Then, with probability at least $\frac{1}{2}\min\{1, (\operatorname{ld}(k+1))^{-r}\}$, the temporal network $\mathcal{H}_k = \mathcal{H}_{\mathcal{R},B_k,[t,t')}$ contains a time-out free x-x' path (with respect to $h(\cdot)$), i.e. $Pr[\mathcal{T}_{\mathcal{H}_k,x,x'}] \geq \frac{1}{2}\min\{1, (\operatorname{ld}(k+1))^{-r}\}$.*

Proof. The proof is by induction, and similar to the proof of Lemma 2.3. For the base case, we consider $k \leq 64$. We claim that the temporal network $\mathcal{H}_k$ contains a time-out free x-x' path with probability at least $\frac{1}{2} \geq \frac{1}{2}\min\{1, (\operatorname{ld}(k+1))^{-r}\}$. In the case $k = 0$, this follows because x and x' must be identical, and the empty path is a time-out free x-x path that exists with probability 1. If $k \geq 1$, then $g(k) \geq \Delta$, so there is an edge (x, x') with probability at least $\frac{1}{2}$.

For the case $k > 64$, we define $k' = k^{\rho/2} - 1$, similar to before — notice that $k - 1 \geq k' \geq 64^{1/2} - 1 = 7$. We divide the time interval $[t, t']$ into three parts $[t, s)$, $[s, s')$ and $[s', t')$, by setting $s = t + g(k')$, $s' = s + \alpha \operatorname{ld}\operatorname{ld}(k+3)$, and write $I = [s, s')$. As before, $B' \subseteq B_k$ is a ball of radius k' containing x'.

In the proof of Lemma 2.3, we argued that there was a node u close to x calling a node u' close to x' during the interval I. Now, we have to ensure in addition that the time-respecting path does not time out at u. As we do not know when the call from u to u' happens, we might have to deal with paths that "wait" at node u for all of I. If u can be arbitrarily close to x, such a wait would certainly result in a time-out, so we want to make sure that u is sufficiently far away from x.

We choose $k'' = \left(\frac{\beta_1}{2\beta_2}\right)^{\frac{1}{D}} \cdot k'$. as the lower bound on the distance from x to u. With $S \subseteq B_k$ denoting a ball of radius k' containing x, we set $B = S \setminus B_{x,k''}$. Then, we obtain that

$$\begin{aligned}
|B| &\geq |S| - |B_{x,k''}| \\
&\geq \beta_1(k')^D - \beta_2(k'')^D \\
&= \frac{1}{2}\beta_1(k')^D \\
&\geq \sigma \cdot (k+1)^{\frac{D\rho}{2}},
\end{aligned}$$

as was shown in the proof of Lemma 2.3. Finally, we define, as before, $\hat{E} = \hat{E}(B, B', I)$, and $\mathcal{H} = \mathcal{H}_{\mathcal{R},B,[t,s)}$, $\mathcal{H}' = \mathcal{H}_{\mathcal{R},B',[s',t')}$.

Because $s - t = g(k')$, we can apply the induction hypothesis to $\mathcal{H}$, and find that

$$\begin{aligned}
\Pr[\mathcal{T}_{\mathcal{H},x,u}] &\geq \frac{1}{2}\min\{1, (\operatorname{ld}(k'+1))^{-r}\} \\
&= \frac{1}{2}(\operatorname{ld}(k'+1))^{-r} \\
&= (\operatorname{ld} k)^{-r} \\
&\geq (\operatorname{ld}(k+1))^{-r},
\end{aligned}$$

for any $u \in B$.

We want to concatenate P with a time-respecting path P' from $u' \in B'$ to x'. This path P' need not be time-out free in itself, because all of its nodes will be sufficiently far away from x, and hence not time out during the interval $[s', t')$. That is, we can simply invoke Theorem 2.1 for the existence of the path once we have ensured that the interval $[s', t')$ is long enough. To bound from below the length of the time interval $[s', t')$, we use the definitions of t', s', k' and g, to obtain that

$$\begin{aligned}
t' - s' &= t + g(k) - (t + g(k') + \alpha \operatorname{ld}\operatorname{ld}(k+3)) \\
&\geq g(k) - \frac{1}{2}g(k) - \alpha \operatorname{ld}\operatorname{ld}(k+3) \\
&\geq \alpha \operatorname{ld}\operatorname{ld}(k+3)(2(\operatorname{ld}(k+1))^r - 1) \\
&\geq \frac{3}{2}\alpha \cdot (\operatorname{ld}(k+1))^r \operatorname{ld}\operatorname{ld}(k+3) \\
&\geq 3\alpha(\operatorname{ld}(k'+1))^r \operatorname{ld}\operatorname{ld}(k'+1).
\end{aligned}$$

In the second and fifth step, we used $\operatorname{ld}\operatorname{ld}(k'+3) \leq \operatorname{ld}\operatorname{ld}(k+3)$ resp. $\operatorname{ld}\operatorname{ld}(k'+1) \leq \operatorname{ld}\operatorname{ld}(k+3)$, and $(\rho/2)^r = \frac{1}{2}$. In the third step, we simply dropped the $\frac{\Delta}{2}$ term, and in the fourth step, we used that $(\operatorname{ld}(k+1))^r \geq \operatorname{ld}(k+1) \geq 6$. We just showed that $t' - s' \geq 3\alpha(\operatorname{ld}(k'+1))^r \operatorname{ld}\operatorname{ld}(k'+1)$, so by Theorem 2.1, we obtain that with probability at least $1 - (\operatorname{ld}(k'+1))^{-3}$, there is a strictly time-respecting (but not necessarily time-out free) u'-x' path P' in $\mathcal{H}'$ for any u'. By substituting $y = 3$, and verifying that the function $\frac{\frac{1}{2}y^{-1}-y^{-2}}{y^{-3}}$ is monotonically increasing for $y \geq 3$, we see that $y^{-3} \leq \frac{1}{2}y^{-1} - y^{-2}$ for all $y \geq 3$. Then, we can substitute $y = (\operatorname{ld}(k'+1))^r \geq \operatorname{ld}(k'+1) \geq 3$, and obtain that $(\operatorname{ld}(k'+1))^{-3} \leq \frac{1}{2}(\operatorname{ld}(k'+1))^{-r} - (\operatorname{ld}(k'+1))^{-2r}$. Therefore,

$$\begin{aligned}\Pr[\mathcal{E}_{\mathcal{H}',u',x'}] &\geq 1 - \frac{1}{2}(\operatorname{ld}(k'+1))^{-r} + (\operatorname{ld}(k'+1))^{-2r} \\ &\geq 1 - \frac{1}{2}(\operatorname{ld}(k+1))^{-r} + (\operatorname{ld}(k+1))^{-2r}.\end{aligned}$$

Assume that there is an edge $e = (u,u') \in \hat{E}$, a time-out free x-u path P in $\mathcal{H}$, and a strictly time-respecting u'-x' path P' in $\mathcal{H}'$. We want to verify that the concatenated path PeP' is time-out free as well. For the subpath P, this follows from the induction hypothesis. Node u and all nodes on the path P' are at distance at least k'' from x. For all these nodes $v \in P' \cup \{u\}$, the departure time is at most t', whereas the departure time from node x is at least t. Therefore, we obtain that

$$\begin{aligned}\delta(v) - \delta(x) &\leq t' - t = g(k) \\ &= h\left(\left(\frac{\beta_1}{2\beta_2}\right)^{\frac{1}{D}} \cdot (k^{\frac{\varrho}{2}} - 1)\right) \\ &= h(k'') \leq h(d_{v,x}),\end{aligned}$$

so PeP' is time-out free as well.

We thus know that for any edge $e = (u,u') \in \hat{E}$, the event $\mathcal{F}_{\mathcal{H}_k,\hat{E},e} \cap \mathcal{T}_{\mathcal{H},x,u} \cap \mathcal{E}_{\mathcal{H}',u',x'}$ implies the existence of a time-out free x-x' path, i.e. the event $\mathcal{T}_{\mathcal{H}_k,x,x'}$. By the same argument as before, the three events $\mathcal{F}_{\mathcal{H}_k,\hat{E},e}$, $\mathcal{T}_{\mathcal{H},x,u}$ and $\mathcal{E}_{\mathcal{H}',u',x'}$ are independent for any fixed x, x' and edge $e = (u,u')$. Using again that the events $\mathcal{F}_{\mathcal{H}_k,\hat{E},e}$ are disjoint for distinct e, we obtain that

$$\begin{aligned}&\Pr[\mathcal{T}_{\mathcal{H}_k,x,x'}] \\ &\geq \sum_{e=(u,u')\in\hat{E}} \Pr[\mathcal{T}_{\mathcal{H},x,u} \cap \mathcal{F}_{\mathcal{H}_k,\hat{E},e} \cap \mathcal{E}_{\mathcal{H}',u',x'}] \\ &= \sum_{e=(u,u')\in\hat{E}} \Pr[\mathcal{T}_{\mathcal{H},x,u}] \cdot \Pr[\mathcal{E}_{\mathcal{H}',u',x'}] \cdot \Pr[\mathcal{F}_{\mathcal{H}_k,\hat{E},e}] \\ &\geq \sum_{e=(u,u')\in\hat{E}} \Big((\operatorname{ld}(k+1))^{-r} \cdot \Pr[\mathcal{F}_{\mathcal{H}_k,\hat{E},e}] \cdot \\ &\qquad (1 - \frac{1}{2}(\operatorname{ld}(k+1))^{-r} + (\operatorname{ld}(k+1))^{-2r})\Big) \\ &= (\operatorname{ld}(k+1))^{-r} \cdot (1 - \Pr[\mathcal{F}_{\mathcal{H}_k,\hat{E},\perp}]) \cdot\end{aligned}$$

$$
\begin{aligned}
& (1 - \frac{1}{2}(\,\mathrm{ld}\,(k+1))^{-r} + (\,\mathrm{ld}\,(k+1))^{-2r}) \\
\geq\ & (\frac{1}{2}\,\mathrm{ld}\,(k+1))^{-r} + (\,\mathrm{ld}\,(k+1))^{-2r} - \Pr[\mathcal{F}_{\mathcal{H}_k,\hat{E},\perp}].
\end{aligned}
$$

Because the length of the interval I is $|I| = \alpha\,\mathrm{ld}\,\mathrm{ld}\,(k+3) \geq \frac{2^{D\rho}}{c\sigma^2}\ln\,\mathrm{ld}\,((k+1)^{2r})$, we obtain from Lemma 2.2, applied to B, B', B_k, and I, that $\Pr[\mathcal{F}_{\mathcal{H}_k,\hat{E},\perp}] \leq (\,\mathrm{ld}\,(k+1))^{-2r}$, and therefore $\Pr[\mathcal{T}_{\mathcal{H}_k,x,x'}] \geq \frac{1}{2}(\,\mathrm{ld}\,(k+1))^{-r}$, completing the proof. ■

References

[1] D. Agrawal, A. El Abbadi, R. Steinke, "Epidemic algorithms in replicated databases," *Proc. ACM Symp. on Principles of Database Systems* 1997.

[2] K. Birman, M. Hayden, O. Ozkasap, Z. Xiao, M. Budiu, Y. Minsky, "Bimodal multicast," *ACM Transactions on Computer Systems* 17(1999).

[3] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Stuygis, D. Swinehart, D. Terry, "Epidemic algorithms for replicated database maintenance," *Proc. ACM Symp. on Principles of Distributed Computing*, 1987.

[4] D. Estrin, R. Govindan, J. Heidemann, S. Kumar, "Next century challenges: Scalable coordination in sensor networks," *Proc. 5th Intl. Conf. on Mobile Computing and Networking*, 1999.

[5] F. Göbel, J. Orestes Cerdeira, H.J. Veldman, "Label-connected graphs and the gossip problem," *Discrete Mathematics* 87(1991).

[6] I. Gupta, R. van Renesse, K. Birman, "Scalable Fault-tolerant Aggregation in Large Process Groups," *Proc. Conf. on Dependable Systems and Networks*, 2001.

[7] S. Hedetniemi, S. Hedetniemi, A. Liestman, "A survey of gossiping and broadcasting in communication networks," *Networks* 18(1988).

[8] W. Heinzelman, J. Kulik, H. Balakrishnan, "Adaptive protocols for information dissemination in wireless sensor networks," *Proc. 5th Intl. Conf. on Mobile Computing and Networking*, 1999.

[9] J. Kahn, R. Katz, K. Pister, "Next century challenges: Mobile networking for 'smart dust,' " *Proc. 5th Intl. Conf. on Mobile Computing and Networking*, 1999.

[10] R. Karp, C. Schindelhauer, S. Shenker, B. Vöcking, "Randomized rumor spreading," *Proc. IEEE Symp. Foundations of Computer Science*, 2000.

[11] D. Kempe, J. Kleinberg, A. Kumar, "Connectivity and Inference Problems for Temporal Networks," *Proc. 32nd ACM Symp. Theory of Computing*, 2000.

[12] J. Li, J. Jannotti, D. De Couto, D. Karger, R. Morris, "A scalable location service for geographic ad hoc routing," *Proc. 5th Intl. Conf. on Mobile Computing and Networking*, 1999.

[13] M. Lin, K. Marzullo, S. Masini, "Gossip versus deterministic flooding: Low message overhead and high reliability for broadcasting on small networks," UCSD Technical Report TR CS99-0637.

[14] B. Pittel, "On spreading a rumor," *SIAM J. Applied Math.* 47(1987).

[15] R. van Renesse, "Scalable and secure resource location," *Proc. Hawaii Intl. Conf. on System Sciences*, 2000.

[16] R. van Renesse, Y. Minsky, M. Hayden, "A gossip-style failure-detection service," *Proc. IFIP Intl. Conference on Distributed Systems Platforms and Open Distributed Processing*, 1998.