Spatial Gossip and Resource Location Protocols

David Kempe[*], Jon Kleinberg[†] and Alan Demers
Dept. of Computer Science
Cornell University, Ithaca NY 14853
email: {kempe,kleinber,ademers}@cs.cornell.edu

September 30, 2004

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, such as 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 mechanisms: 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 mechanisms, called *spatial gossip*, 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 mechanisms and their analysis resolve a conjecture of Demers et al.

We further show an application of our gossip mechanisms to a basic *resource location problem*, in which nodes seek to rapidly learn of the nearest copy of a *resource* in a network. This problem, which is of considerable practical importance, can be solved by a very simple protocol using Spatial Gossip, whereas we can show that no protocol built on top of uniform gossip can inform nodes of their approximately nearest resource within poly-logarithmic time. The analysis relies on an additional useful property of spatial gossip, namely that information travels from its source to sinks along short paths not visiting points of the network far from the two nodes.

[*]Supported by an NSF Graduate Fellowship.

[†]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.

1 Introduction

Gossip mechanisms

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 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]; such information should spread quickly through the network. As another application, 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.

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 computing hosts positioned at the lattice points of a $\sqrt{N} \times \sqrt{N}$ region of the plane. 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. (This assumption is usually warranted in distributed system settings where nodes represent computing hosts connected by a significant underlying network structure. It becomes less realistic in wireless or sensor networks, where the cost tends to increase in the communication distance.) All the mechanisms we develop here are built on top of such a point-to-point communication mechanism.

Here is a basic problem we may wish to solve in such an environment: If a 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 mechanisms 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 mechanisms 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 *mechanism*, by which nodes choose other nodes for (point-to-point) communication; and (ii) a gossip-based *protocol* built on top of a gossip mechanism, 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 mechanism 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 mechanism as simply generating a labeled graph $\mathcal{H}$ on which the protocol then operates.

The propagation time

The two mechanisms discussed above have quite different performance bounds, with uniform gossip "filling in" the set of points exponentially faster than neighbor flooding. The neighbor flooding mechanism, however, exhibits desirable features that uniform gossip is lacking:

1. 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.

2. Messages travel between nodes on a shortest path.

In the examples given above, and in applications exhibiting any kind of spatial locality, these properties 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. In particular, if we think of more complex applications in which data must be locally aggregated, then the "unstructured" communication patterns of uniform gossip may prevent us from designing simple protocols.

Our work was initially motivated by the following question: Is there a gossip mechanism — 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 $f_{\mathcal{M}}(\cdot)$ is a *propagation time* for a given gossip mechanism $\mathcal{M}$ 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 $O(f_{\mathcal{M}}(d))$ steps.

Our question can now be phrased as follows: Is there a gossip mechanism with a propagation time that is polynomial in $\log d$? In addition, we would like the path along which the information travels to be not much longer than the distance d.

A gossip mechanism 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 $\mathbb{R}^D$. We will make this notion precise in Section 2; essentially, it is a point set in $\mathbb{R}^D$ (for a fixed constant D) with the property that every unit ball contains $\Theta(1)$ points.

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

Theorem 1.1 *Let P be a set of points with uniform density in $\mathbb{R}^D$, for a fixed constant D. For every $\varepsilon > 0$, there is a randomized gossip mechanism on the points in P with a propagation time that is $O(\log^{1+\varepsilon} d)$. In addition, the lengths of the paths that the information travels is $d + o(d)$.*

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 mechanism 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 mechanism 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.

If we were only interested in the first property, i.e., dissemination time polynomial in $\log d$, but not in short paths, an alternate mechanism would provide the desired guarantees (we thank an anonymous referee for this suggestion): In each round, each node chooses a *distance bound* b according to some distribution, and then chooses uniformly at random a node at distance at most b. If $b = 2^{2^k}$ with probability k^{-2}, then the resulting dissemination time is $O(\log d \cdot (\log \log d)^2)$; if $b = 2^k$ with probability k^{-2}, then the dissemination time is $O(\log^3 d)$. The idea is that this mechanism, at each scale, resembles uniform gossip, slowed down by a factor of $1/(\log \log d)^2$ resp. $1/(\log d)^2$.

As such, the mechanism does not meet our second requirement, that of short paths. Whether or not such a mechanism can be used to construct efficient protocols will depend on the nature of the problem and protocols. We will see that the length bounds are useful for locating nearby resources, but that an alternate, and slightly weaker, guarantee can be obtained by merely relying on the fast dissemination time.

Building on the analysis of spatial gossip, we show how our mechanisms 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 mechanisms, 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 mechanism 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 mechanism, 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 mechanism 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 cluster of computers may be running the server component of a client-server application, or have free memory or idle machine time.) 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 mechanism 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 SP. 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 mechanism 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'. (If another resource appears between r and u during the interval, then u may learn about r or one of the new resources.)

- In higher dimensions, it has the following approximate guarantee. Again, let r and u be two nodes at distance d. If r acquires a resource at time t, then with high probability node u will know of a resource-holder within distance $d + o(d)$ by time $t + O(\log^{1+\varepsilon} d)$.

Both of these guarantees rely not only on the dissemination time of the gossip mechanism, but also on the existence of short paths that do not visit remote parts of the network. It is therefore natural to ask whether these properties are in fact crucial for the design of resource location protocols, or merely an easy way to prove good bounds. In partial answer to this question, we can prove the resource location problem cannot be solved using uniform gossip. Specifically, we prove that, using bounded-size messages, uniform gossip cannot solve the monotone resource-location problem within time poly-logarithmic in the number n of nodes, even in an approximate sense.

Finally, we define the *non-monotone* version of the problem, in which 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 a protocol 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 Preliminaries

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

Let $B_{x,d} = \{y \mid d_{x,y} \le d\}$ denote the ball around x of radius d. We say that an (infinite) set of points P in $\mathbb{R}^D$ has *uniform density*, with parameters β_1 and β_2, if every ball of radius $d \ge 1$ 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 mechanisms 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 mechanisms 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, 12] 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 $1 \le i < k$. 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 mechanism $\mathcal{M}$, 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 lg the base-2 logarithm.

3 Spatial Gossip in $\mathbb{R}^D$

We consider gossip mechanisms based on inverse-polynomial probability distributions. The mechanisms are parameterized by an exponent ρ satisfying $1 < \rho < 2$. Let $x \ne y$ be two nodes at distance $d = d_{x,y}$, and let $p^{(\rho)}_{x,y}$ denote the probability that x calls y. Then, we let $p^{(\rho)}_{x,y} := c_x (d+1)^{-D\rho}$, i.e., the probability that y is called decreases polynomially in the distance between x and y. c_x is chosen such that $\sum_y p^{(\rho)}_{x,y} \le 1$. 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 mechanism by $\mathcal{M}^\rho$.

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 reciprocal of the normalizing constant at point x will be at most

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

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 $\mathcal{M}^\rho$ when $1 < \rho < 2$. We state the result in the following form.

Theorem 3.1 *Fix a ρ with $1 < \rho < 2$. If x and x' are at distance d, then with probability at least $1 - O(\log^{-\kappa} d)$, there is a time-respecting path P from x to x' within time $O(\log^{\frac{1}{1-\lg(\rho)}} d \cdot \log\log d)$, such that all nodes on P are in the smallest ball containing x and x'.*

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 path taken by 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', and illustrated in Figure 1. In the example, we want to prove the existence of a time-respecting path from x to x', which are at distance k. This is a sub-instance of the inductive proof, which initially set out to prove that there is a time-respecting path between some nodes at distance $d \geq k$.

We consider nodes within distance $k^{\rho/2}$ of x and x' to be near them, and show that during a time interval I of length $\eta = O(\kappa \log\log d)$, some node u near x calls a node u' near x', thus establishing a *long-range contact*, with sufficiently high probability $1 - O(\log^{\frac{-1}{1-\lg\rho}-\kappa} d)$. Note that d is the distance of the two nodes at the *outermost* stage of the inductive proof, not the current stage.

By applying induction on the subpaths from x to u, and from u' to x', we can then conclude the existence of a time-respecting x-x' path with the probability claimed in the theorem. The time satisfies the recurrence relation $g(k) = \eta + 2g(k^{\rho/2})$, which has solution $g(k) = O(\eta \cdot \log^{\frac{1}{1-\lg\rho}} k)$. Finally, we can take a Union Bound over all stages of the inductive proof to show that the probability of all required long-range contacts being made is $1 - O(\log^{-\kappa} d)$.

The central (inductive) idea explained above is captured in the following Lemma. For ease of notation, we omit the constant factor η from the definition of g, and let g be the solution of the recurrence $g(k) = 1 + 2g(k^{\rho/2})$.

Lemma 3.2 *Let x and x' be at distance k, $\hat{B}$ some ball containing x and x', and η and γ constants with $\eta = O(\frac{-1}{\beta_1^2 c}\log\gamma)$ (where the O hides dependencies only on D and ρ). Then, if t and t' are two points in time with $t' - t \geq \eta g(k)$, there is a time-respecting path P from x to x' with probability at least $1 - \gamma g(k)$, such that all time labels are in $[t, t')$ and all nodes in P lie in $\hat{B}$.*

Proof. Because g is the solution of the recurrence $g(k) = 1 + 2g(k^{\rho/2})$, we can divide the time interval $[t, t')$ into three parts $[t, s)$, $[s, s')$ and $[s', t')$, of lengths $\eta g(k^{\rho/2})$, η, and $\eta g(k^{\rho/2})$, respectively.

The base case is when $k \leq 2$ (or some other fixed constant): in that case, only a constant number (depending on β_2 and D) of nodes can be closer to x than x', so after some constant

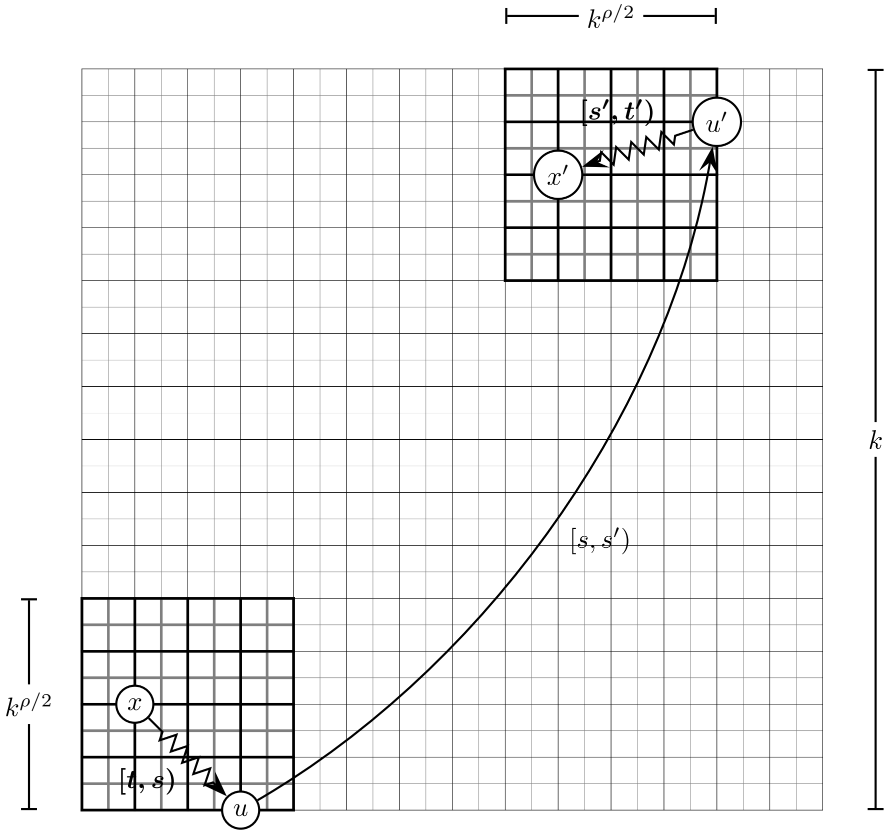

Figure 1: Illustration of the proof of Theorem 3.1: The darker regions are the nodes "near" x resp. x'. We first argue the existence of the long-range arc, then use induction to prove the existence of the zig-zag paths.

number of attempts, any constant probability $1 - \gamma g(2)$ of x calling x' directly is exceeded. From now on, we consider the case that $k > 2$.

Let B and B' be balls of diameter $k^{\rho/2}$, containing x and x', respectively, and each of them contained in $\hat{B}$. For any node $u \in B$, by the Induction Hypothesis, there is a time-respecting x-u path with labels in $[t, s)$ with probability at least $1 - \gamma g(k^{\rho/2})$. Similarly, for any node $u' \in B'$, there is a time-respecting u'-x' path with labels in $[s', t')$ with probability at least $1 - \gamma g(k^{\rho/2})$. Because the three time intervals $[t, s)$, $[s, s')$ and $[s', t')$ are disjoint, the existence of the x-u path and the u'-x' path are independent, and also independent of any outcome of random events in the time interval $[s, s')$.

We now have a look at the probability of making a long-range contact from B to B' during the time interval $[s, s')$. Each point $u' \in B'$ is at distance at most k from any point $u \in B$, so each point $u \in B$ in each round has probability at least $c \cdot |B'| \cdot k^{-D\rho}$ of calling some point $u' \in B'$. Since all these attempts are independent (over nodes $u \in B$ and times in $[s, s')$), and B and B' each contain at least $\beta_1 \cdot (k/2)^{D\rho/2} = \Omega(\beta_1 k^{D\rho/2})$ points by the uniform growth requirement, the probability that no call from B to B' occurs during the time interval $[s, s')$ is at most

$$(1 - c|B'|k^{-D\rho})^{|B|(s'-s)} \quad \leq \quad e^{-\Omega(c\eta\beta_1^2)} \quad = \quad \gamma$$

Assuming that the high-probability event happened, and some node $u \in B$ called $u' \in B'$ at some time in $[s, s')$, there are time-respecting paths P from x to u during the interval $[t, s)$, and P' from u' to x' during $[s', t')$, with probability at least $1 - \gamma g(k^{\rho/2})$ each. By the Union Bound, the probability that the labeled edge (u, u') as well as P and P' exist is at least $1 - \gamma - 2\gamma g(k^{\rho/2}) = 1 - \gamma g(k)$, completing the inductive proof of the Lemma. ■

The theorem now follows quite easily, by substituting concrete values for η and γ.

Proof of Theorem 3.1. First, a simple inductive proof shows that indeed $g(k) = O(\log^{\frac{1}{1-\lg \rho}} k)$. In particular, by choosing $\gamma = O(\log^{\frac{-1}{1-\lg \rho}-\kappa} d)$, we will obtain the claimed high-probability guarantee.

In order to apply Lemma 3.2, it now suffices to choose $\hat{B}$ to be the smallest ball containing both x and x', as well as $\eta = O(\frac{-1}{\beta_1^2 c} \log \gamma) = O(\kappa \log \log d)$, where the O-term now depends on all parameters of the metric space (D, β_1, β_2), as well as on ρ, but not on κ or the distance d. This then gives the claimed guarantee. ■

Of course, we can get rid of the $\log \log d$ term by increasing the exponent for the log-term to $\log^{\frac{1}{1-\lg \rho}+\varepsilon}$ for some small ε.

3.1 Other values of the parameter ρ

It is natural to ask about the behavior of our inverse-polynomial gossip mechanisms for different values of the exponent ρ. We have found that the set of possible values for ρ can be divided into three regimes 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 d alone, and will depend on n. This is contrary to one of our goals in the design of our gossip distributions, and in particular includes the case $\rho = 0$, which is identical to uniform gossip.

Theorem 3.3 *Consider the metric space of n points $\{0, \ldots, n-1\}$ on the line with distance function $d_{x,y} = |y - x|$. Assume that at time 0, some information originates at point 0. If $\rho < 1$, let $T = \frac{1-\rho}{2} \lg n$, otherwise $T = \frac{1}{2} \lg \lg n$. In either case, node 1 will have obtained the information originating with 0 within T rounds of $\mathcal{M}^\rho$ with probability at most $o(1)$.*

Proof. Let us call a point $x \in \{0, \ldots, n-1\}$ *marked* at time t if there is a time-respecting path from 0 to x with all time labels in $[0, t]$. That is, a point is marked if it would have received information originating with 0 if all messages were forwarded.

To define a probability distribution with probabilities $c_x |y - x|^{-\rho}$, the normalizing constant c_x at each node x has to be $O(n^{\rho-1})$ for $\rho < 1$, and $O(\frac{1}{\lg n})$ for $\rho = 1$. (For each node x, there is at least one node y within distance d for each $d \leq n/2$ — summing over all such nodes y gives the bound). Let c_ρ be this upper bound on all normalizing constants.

The number of marked nodes can at most double in each round, because a node can only become marked by being called by another marked node. Therefore, there are at most 2^T pairs (x, t) such that node x was marked at time $t \leq T$. Only nodes x that are marked at time $t \leq T$ could make node 1 marked by time T, and each of them calls node 1 with probability at most c_ρ. By the Union Bound, taken over all such pairs (x, t), we find that at time T, node 1 is marked with probability at most $2^T \cdot c_\rho = O(n^{(\rho-1)/2}) = o(1)$ (in the case $\rho < 1$) resp., $2^T \cdot c_\rho = O((\lg n)^{-1/2}) = o(1)$ (in the case $\rho = 1$), completing the proof. ■

For $\rho > 2$, the time can be bounded as a function of the distance d only, but that function is polynomial, not poly-logarithmic. The reason is that long-range contacts become too rare, so $\mathcal{M}^\rho$ resembles flooding. Formally, we prove the following:

Theorem 3.4 *Consider the metric space $\mathbb{N} = \{0, 1, 2, \ldots\}$ with distance $d_{x,y} = |y - x|$. Let 0 be the source of some information at time 0. Writing $\varepsilon = \frac{\rho-2}{2\rho} > 0$, the probability that a node x at distance $2d$ from 0 has received the information by time d^ε is $o(1)$.*

Proof. Because each node x has a neighbor (at distance 1), the normalizing constant for the distribution at each node is at most $c_x \leq 1$. Let $\mathcal{E}_{x,d,t}$ denote the event that point x at time t calls a point at distance more than d. Then, $\text{Prob}[\mathcal{E}_{x,d,t}] \leq d^{1-\rho}$, simply by bounding the (infinite) sum over distances by an integral.

Let $q = \frac{2}{\rho} < 1$ (thus $\varepsilon = \frac{1-q}{2}$). Also, let $\mathcal{F}_t(s, s')$ denote the event that in round t, some point x with $x \leq s$ calls some point y with $y > s'$. Specifically, we are interested in the events $\mathcal{F}_t := \mathcal{F}_t(d + td^q, d + (t+1)d^q)$, for $t = 0, \ldots, d^\varepsilon - 1$.

To bound the probability of the events $\mathcal{F}_t$, notice that $\mathcal{F}_t \subseteq \bigcup_{x<d+td^q} \mathcal{E}_{x,d^q,t}$. Applying the Union Bound with the bound $\text{Prob}[\mathcal{E}_{x,d^q,t}] \leq d^{q(1-\rho)} = d^{q-2}$ from above, we obtain that $\text{Prob}[\mathcal{F}_t] \leq (d + td^q)d^{q-2} \leq 2d^{q-1}$. We take the Union Bound again, this time over all $t = 0, \ldots, d^\varepsilon - 1$, and find that the probability that any of the events $\mathcal{F}_t$ occurs is at most $2d^\varepsilon \cdot d^{q-1} = 2d^{-\varepsilon}$.

In the absence of the events $\mathcal{F}_t$, a simple inductive proof on the time t shows that no point $x > d + td^q$ can be marked (in the sense used in the proof of Theorem 3.3) by time t. In particular, no point $x > 2d \geq d + d^\varepsilon d^q$ can be marked at time d^ε.

Hence, the probability that a point x at distance $2d$ has received information originating with 0 within d^ε steps of $\mathcal{M}^\rho$ is at most $2d^{-\varepsilon} = o(1)$, completing the proof. ■

This leaves only the "transitional case" $\rho = 2$, which turns out to have very interesting behavior; an analysis similar to the one for $\rho \in (1, 2)$ shows that for every $\varepsilon > 0$, the gossip mechanism $\mathcal{M}^\rho$ with $\rho = 2$ has a propagation time that is $O(d^\varepsilon)$, but we do not know whether it has a propagation time that is polynomial in $\log d$.

4 A more general setting

In the proofs of Theorem 3.1 and Lemma 3.2, we only used relatively few properties of the metric space $\mathbb{R}^D$. 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 (possibly infinite) 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 finite subset $S \subseteq X$, there is a disc $D \in \mathcal{D}$ with $S \subseteq D$.
2. 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 $|D| \leq \beta \cdot b$.
3. 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 $\mathbb{R}^D$, 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 $\mathcal{M}^\rho$. For points $x, y \in X$, let b be the minimum cardinality of any disc $D \in \mathcal{D}$ containing both x and y. Define $p_{x,y}^{(\rho)} := c_x b^{-\rho}$, where c_x is again a normalizing constant at point x. From the second 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 $c_x b^{-\rho}$ or more to the total probability mass at point x.

Hence, the reciprocal of $c_x \sum_{y \neq x} p_{x,y}^{(\rho)}$ 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 $\mathcal{M}^{\rho}$ as follows.

Theorem 4.1 *Let $1 < \rho < 2$. Let $x, x' \in X$ be two points, D a disc of size b containing both x and x', and t, t' two points in time with $t' - t \in \Omega(\log^{\frac{1}{1-\lg \rho}} b)$. Then, with probability at least $1 - O(\log^{-\kappa} b)$, there is a time-respecting path from x to x', all of whose nodes are contained in D, and with time labels in $[t, t')$.*

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 $O()$ in the theorem hides polynomial dependencies on both β and μ. The proof of this theorem closely follows the proof of Theorem 3.1, and we therefore omit the details here. The crucial observation is that the repeated application of Axiom 3 allows us to obtain the smaller discs on which we can apply induction. The larger the constant μ, the smaller those disks will be, so it will take more rounds to establish a long-range link.

4.1 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 three axioms described above hold for this collection of discs. Hence, if leaf nodes communicate according to the gossip mechanism in Theorem 4.1, 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 mechanism actually used in the Astrolabe system is essentially equivalent to $\mathcal{M}^{\rho}$ with $\rho = 2$. The results about exponent $\rho = 2$ mentioned in Section 3.1 also hold for the more general disc model. Therefore, information will spread through a k-leaf subtree, with high probability, in time $O(k^{\varepsilon})$ for every $\varepsilon > 0$.

5 Resource Location Protocols

In the introduction, we discussed the basic *resource location problem*. We have nodes in $\mathbb{R}^D$ 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.")

We consider two versions of this problem: In the *monotone* version, nodes never lose a resource once they acquire it, whereas in the *non-monotone* version, copies of resources may appear and disappear. The more general version requires more elaborate protocols and analysis.

The resource location problem can be solved quite easily if arbitrarily large messages are allowed. Messages whose size is polynomial in the number of nodes are, however, highly impractical — if nothing else, the running time for the protocol would depend polynomially on the system size even if the number of rounds is poly-logarithmic. We therefore restrict our attention to protocols with smaller message sizes, specifically, protocols that forward only a constant number of node names in each message.

Our protocols will rely on an underlying mechanism for gossip on point sets of uniform density in $\mathbb{R}^D$; we do not make any assumptions about this mechanism other than versions of the probabilistic propagation guarantee which we proved $\mathcal{M}^\rho$ to have in Section 3. The most basic required guarantee can be expressed as follows:

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

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 mechanism, and will directly affect the guarantees provided by the resource location protocol. In Section 3, we established that (*) holds for inverse polynomial distributions with exponents $\rho \in (1, 2)$, with $f_{\mathcal{M}}(d) \in O(\kappa \log^{1+\varepsilon} d)$, giving us high probability guarantee $1 - O(\log^{-\kappa} d)$. Alternately, the guarantees are also met by the "interleaved uniform gossip" mechanism. Probably the simplest protocol that only forwards a constant number of node names in each round is the following protocol SP. In it, 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.

At time $t = 0$, set $N_x(0) = \perp$.
In each time step t
1. If x holds a resource, then set $N_x(t) = x$.
2. Choose a node x' according to $\mathcal{M}$, and send $N_x(t)$ to x'.
3. Let M_t be the set of all messages x received in step t.
4. Let $N_x(t+1)$ be the node in $M_t \cup \{N_x(t)\}$ closest to x,
   breaking ties in favor of $N_x(t)$ (otherwise arbitrarily).

Figure 2: The resource location protocol SP

Notice that nodes maintain a very "egocentric" view of the world, discarding all information except what is most useful for them. Yet, as we will see, the guarantees provided by spatial gossip will be enough to show very good guarantees on the protocol's behavior. We analyze SP and some variations of it in the following Section 6.

Subsequently, in Section 7, we prove the perhaps surprising fact that no protocol with constant message sizes can provide good guarantees for resource location in poly-logarithmic time using uniform gossip as an underlying mechanism. This shows that the more "structured" communication patterns of spatial gossip have additional advantages beyond providing dissemination time that depends solely on the distance d.

Finally, we address the challenges resulting from the non-monotone version of the resource location problem. In order to deal with disappearing resources, further modifications to the protocol will be necessary. Basically, information about resources has an expiration date, and needs to be refreshed periodically — we discuss the resulting protocol in Section 8.

6 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.

6.1 Analysis for the line

For dimension $D = 1$, i.e., on the line, the protocol SP solves the resource location problem exactly, in poly-logarithmic time, so long as the underlying mechanism $\mathcal{M}$ satisfies the guarantee (*) from Section 5.

Theorem 6.1 *Let x be any node, and x_R a resource at distance $d = d_{x,x_R}$. Let $t' = t + \kappa f_{\mathcal{M}}(d)$, 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 (*) provided by the underlying mechanism $\mathcal{M}$ 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 $N_x(t') = x_R$. ∎

6.2 Monotone Resource Location in higher dimensions

In higher dimensions, the protocol SP may not inform nodes of their truly closest resource as quickly as $\kappa f_{\mathcal{M}}(d)$. 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 informing all of 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 mechanism $\mathcal{M}$: 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 $f_{\mathcal{M}}(d)$ and a length function $\ell_{\mathcal{M}}(d)$ such that for any ball B of diameter d, any two nodes $x, x' \in B$, and any time t, $\mathcal{H}_{\mathcal{R},B,[t,t+\kappa f_{\mathcal{M}}(d))}$ contains a time-respecting path P from x to x' of path distance $d(P) \leq \ell_{\mathcal{M}}(d)$, with high probability.

Below, we prove that the mechanisms $\mathcal{M}^\rho$ with $\rho \in (1, 2)$ satisfy this property with $\ell_{\mathcal{M}}(d) = d+o(d)$, thus establishing the second part of Theorem 1.1.

We can now state the approximation guarantee for the protocol SP:

Theorem 6.2 *Let x be any node, and x_R a resource at distance $d = d_{x,x_R}$. Let $t' = t + \kappa f_{\mathcal{M}}(d)$, and $x'_R = N_x(t')$. Then, $d_{x,x'_R} \leq \ell_{\mathcal{M}}(d)$ with high probability.*

Hence, for sufficiently large d, the inverse polynomial gossip mechanisms $\mathcal{M}^\rho$ 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 $d(P) \leq \ell_{\mathcal{M}}(d)$. 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

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

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

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

completing the proof. ■

By taking a closer look at the analysis in the proof of Theorem 3.1, we can show that the mechanisms $\mathcal{M}^\rho$ (for $1 < \rho < 2$) have the claimed property of short paths with $\ell_{\mathcal{M}^\rho}(d) = d + o(d)$. Hence, the time-respecting paths are almost shortest paths between the two nodes.

Lemma 6.3 *Let B be a ball with diameter d, t an arbitrary time, and $t' = t + \kappa f_{\mathcal{M}^\rho}(d)$. Then, with probability at least $1 - O(\log^{-\kappa} d)$, 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 Lemma 3.2 consists of one long-range connection of length at most d, and two subpaths P and P' which lie inside balls B and B' of diameter at most $d^{\rho/2}$, and are of the same form. Hence, we obtain the recurrence $\ell_{\mathcal{M}^\rho}(d) \le d + 2\ell_{\mathcal{M}^\rho}(d^{\rho/2})$. Using standard substitution techniques (and writing $n = \frac{\lg\lg d}{1-\lg\rho}$), we find that this recurrence has the solution

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

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 $\ell_{\mathcal{M}^\rho}(d) \le d + (\lg d)^{\frac{1}{1-\lg(\rho)}} \frac{\lg\lg d}{1-\lg\rho} \cdot d^{\rho/2}$, 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 mechanism. This is done by having nodes send larger messages in each time step. Using larger messages thus lets us combine our protocol with more gossip mechanisms, including the "interleaved uniform gossip" one.

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.

A more precise description of the protocol is as follows: 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 $t' \ge t$.

In each round t, each node x selects a communication partner according to the mechanism $\mathcal{M}$, 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 $R_x(t+1) := \{y \in M \mid d_{x,y} \le \xi d\}$, 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 6.4 *Let x be any node, and x_R a resource at distance $d = d_{x,x_R}$. Let $t' = t + \kappa f_{\mathcal{M}}(d)$, and x'_R be a node in $R_x(t')$ with minimal distance to x (among all nodes in $R_x(t')$). Then, $d_{x,x'_R} \le (1 + \frac{2}{\xi-1})d$ 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 $(G, \lambda) = \mathcal{H}_{\mathcal{R},B,[t,t')}$. By (*), this network contains a strictly time-respecting path P from x_R to x with high probability. 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

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

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

7 Impossibility of Resource Location with Uniform Gossip

We saw in Section 6 that by using Spatial Gossip with exponent $\rho \in (1,2)$, information about (approximately) closest resources reaches each node in time poly-logarithmic in its distance from the resource. For Uniform Gossip, such a guarantee is certainly too much to hope for. After all, it will, with high probability, take $\Omega(\log n)$ steps until a node x has a time-respecting path to its immediate neighbor on the line. Nevertheless, one could hope that all nodes could find out about their (approximately) closest resources in time poly-logarithmic in n, the number of nodes. While this is not as desirable as a dependence solely on the distance, it would still scale well in the size of the system. Our protocols in the previous sections relied on useful properties of spatial gossip, but it is not clear a priori that these properties are really required.

In this section, we prove that even the weaker guarantee is unattainable so long as the message size is bounded by a constant. We place the following restrictions on the protocols. Only a node itself can claim that it possesses a copy of the resource, by generating an authenticated *resource message*, which may be forwarded but not altered by other nodes. In order to consider x_R its closest resource, a node x must have received a resource message that originated with x_R. There may be additional messages sent, but we will be concerning ourselves mostly with the delivery of resource messages. Communication partners are chosen uniformly and independently at random, and in particular are independent of the messages that the nodes hold at the time of random choices. The only other assumption we place on the protocol is that nodes do not know the outcome of future random choices.

We let α denote the (constant) bound on the number of messages forwarded during any one communication connection, and assume that the protocol runs for $T = \log^k n$ rounds.

We construct the following resource location instance on the line. Let $\epsilon < 1$ and $\delta = n^\epsilon$, and consider a set of n nodes in $\mathbb{R}$ positioned at the points $\{1, 2, \ldots, n\}$. There is a resource-holder r_j at each point of the form $2j\delta$ for natural numbers $j = 1, 2, \ldots, n^{1-\epsilon}/2$, and x_j denotes the node at the point $2j\delta - 1$. Obviously, the closest resource to x_j is r_j, and any other resource is further away from x_j by at least by a factor of δ. In particular, we are interested in how many of the nodes x_j will receive a resource message from their corresponding resource r_j within poly-logarithmic time.

Theorem 7.1 *After T rounds, the expected number of nodes x_j having received a resource message from their corresponding resource r_j is $o(n^{1-\epsilon})$.*

The intuition behind the proof is as follows: Only αn copies of resource messages are made (and forwarded) during any one round, so within T rounds, the average number of copies per resource message is αT. By the Pigeon Hole Principle, only very few messages can exist in significantly more copies — all other resource messages are *rare*. If x_j is a node in need of a rare message, it must at some point communicate with one of the few — say, $O(\alpha T)$ — owners, and within T rounds, this happens only with probability $O(\alpha T^2/n)$ by the Union Bound. Thus the expected number of resource messages received by their x_j is small, namely mostly the few messages that existed in large numbers.

Proof. Let N be a random variable denoting the number of nodes x_j that have received a resource message from resource r_j by time T. We say that x *knows about* r_j at time t (and use $\mathcal{K}_t(x, j)$ to denote this random event), if, at time t or earlier, x has received a resource message originating with r_j. By writing $K_t(x, j)$ for the 0-1 indicator variable of that event, we can express N as a sum of indicator variables: $N = \sum_j K_T(x_j, j)$.

We let $\mathcal{C}_t(x, y)$ denote the event that nodes x and y communicate in round t, either because x called y, or because y called x. The probability of that event is at most $\frac{2}{n}$ for uniform gossip.

If x_j knows about r_j at time T, it must at some time $t \leq T$ have communicated with a node y that knew about r_j at time t, so

$$\mathcal{K}_T(x_j, j) \quad \subseteq \quad \bigcup_{t,y} \mathcal{C}_t(x_j, y) \cap \mathcal{K}_t(y, j).$$

Applying the Union Bound, and the fact that the choice of communication partners is made independently of the data held, this yields

$$\text{Prob}[\mathcal{K}_T(x_j, j)] \quad \leq \quad \textstyle\sum_{t,y} \text{Prob}[\mathcal{C}_t(x_j, y)] \cdot \text{Prob}[\mathcal{K}_t(y, j)] \quad \leq \quad \frac{2}{n} \cdot \sum_{t,y} \text{E}\left[K_t(y, j)\right].$$

By linearity of expectation, the expected number of nodes x_j knowing about their closest resource r_j by time T is just the sum of the probabilities for the events $\mathcal{K}_T(x_j, j)$, taken over all j.

$$\text{E}\left[N\right] \quad = \quad \textstyle\sum_j \text{Prob}[\mathcal{K}_T(x_j, j)] \quad \leq \quad \frac{2}{n} \cdot \sum_{j,t,y} \text{E}\left[K_t(y, j)\right] \quad = \quad \frac{2}{n} \cdot \text{E}\left[\sum_{t,j,y} K_t(y, j)\right].$$

At time 0, only the resources r_j know about themselves, so $\sum_{j,y} K_0(y, j) = n^{1-\epsilon}/2$. During each round of communication, exactly n calls are made, and each call transmits at most α messages. Hence, there are at most $n \cdot \alpha$ pairs (j, y) such that $K_t(y, j) = 0$ and $K_{t+1}(y, j) = 1$. By induction, it is easy to see that for all times $t \leq T$,

$$\textstyle\sum_{y,j} K_t(y, j) \quad \leq \quad t\alpha \cdot n + n^{1-\epsilon}/2 \quad \leq \quad (\alpha T + 1) \cdot n$$

Therefore, the above expectation can be bounded as

$$\text{E}\left[N\right] \quad \leq \quad \frac{2}{n} \cdot \text{E}\left[\textstyle\sum_t (\alpha T + 1) \cdot n\right] \quad = \quad \Theta(\alpha T^2) \quad = \quad o(n^{1-\epsilon}),$$

completing the proof. ■

Hence, only a vanishingly small fraction of the nodes x_j finds out about their closest resource, and for all other x_j, the closest resource they know about is at least by a polynomial factor n^ϵ further away than the truly closest one, where ϵ can be made arbitrarily large, so long as it is smaller than 1.

It is interesting to note that the sole obstacle seems to lie in the fact that uniform gossip is "too unstructured". Nodes do not know which of their messages to forward, since the location of a node has nothing to do with its future likely communication connections. We conjecture that the resource location problem could be solved using uniform gossip and small messages if nodes knew about future random outcomes.

In fact, this inability to forward messages to particular destinations under uniform gossip lies at the heart of a more general impossibility result, encompassing a wider range of gossip mechanisms as well as additional problems, including approximating a minimum spanning tree by gossip. The details are beyond the scope of this work, and appear in a separate paper [11].

8 Non-Monotone Resource Location

Designing resource-location protocols becomes more complex if resources may disappear over time. In that case, information about the disappearance needs to propagate through the system, 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 x_R. 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 $\delta(v_i) - \delta(v_1) \leq h(d_{v1,v_i})$ for all nodes v_i on the path. Now, the requirement on the underlying protocol can be stated as the following low-probability guarantee:

> There is a constant r and 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 $\Omega(\log^{-r} d)$.

Below, we will show that the inverse polynomial gossip algorithms $\mathcal{M}^\rho$ satisfy this property with $r = \frac{1}{1-\lg(\rho)}$ and time-out function $h(d) = O(f_{\mathcal{M}^\rho}(d))$. First, however, we will see how we can exploit this property to build a protocol for the non-monotone resource location problem.

8.1 A protocol for non-monotone resource location

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, but does not believe in x_R at time $t+1$. Using the time-out function $h(\cdot)$, we define the time-out function $h'(d) = \Theta(h(d) \cdot \kappa \cdot \log^r d \cdot \log\log d)$ for the protocol, where κ is again a measure of the desired probability guarantee, and the constants in Θ are chosen so that the algebra below works out nicely.

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 $\hat{x} = \text{argmin}_{x \neq x'}(d_{x,x'})$. 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 $\mathcal{M}$.

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 8.1 *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 at any time during 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 $1 - O(\log^{-\kappa} d)$.

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 $\Theta(\kappa \log^r \cdot d \log\log d)$. (Here, the constants are chosen so that the calculations below work out nicely.) Now, fix one such j, and the associated interval $I = [t_j, t_{j+1})$ of length $h(d)$.

By assumption on $\mathcal{M}$, 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 $\Omega(\log^{-r} d)$. 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

$$(1 - O(\log^{-r} d))^{\Omega(\kappa \log^r d \cdot \log \log d)} \leq e^{-\Omega(\log^{-r} d \cdot \kappa \log^r d \cdot \log \log d)} = O(\log^{-\kappa} d),$$

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 $d + o(d)$ from 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 8.1 and 6.2, and therefore omitted here. The $o(d)$ term is necessary there, as the time-out free paths, much like the time-respecting ones in Theorem 6.2, may visit nodes at distance $d + o(d)$ from x, and x may believe in one of those nodes, even if it disappears later. Alternately, if the path times out, then x may not believe in any nearby node.

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

In the remainder of this section, we argue that the inverse polynomial gossip algorithms $\mathcal{M}^\rho$ from Section 3 actually satisfy the above property of producing time-out free paths, with time-out function $h(d) = O(\log^{1+\varepsilon} d)$. Hence, the protocol for non-monotone resource location presented above will have time-out function and dissemination time bound $O(\log^{2+\varepsilon} d)$. Notice that this exhibits yet another useful property of $\mathcal{M}^\rho$ as a primitive for protocol design, which is not matched by interleaving uniform gossip stages.

The proof is quite similar to the one of Theorem 3.1. We want to establish that there is a time-out free strictly time-respecting path from a node x to x' at distance d, within an allotted time interval. We again divide the time interval into three disjoint parts. In the middle part, a long-range connection between two nodes u and u' close to x resp. x' will be established with sufficiently high probability. Inductively, we will show the existence of a time-out free x-u path during the first part of the interval.

There are two major differences between the proofs. The major additional difficulty arises because the proof of Lemma 3.2 did not make any guarantees as to how much time elapsed between the time when u was reached by the (inductively guaranteed) path, and the long-range contact. If u was very close to x, and a long time elapsed, then u would time out on x, and the concatenated path would not be time-out free. We solve this problem by requiring that not only is u at most at distance $d^{\rho/2}$ from x, but also that there is a comparable minimum distance of $\frac{1}{\alpha} d^{\rho/2}$ between them (for some constant $\alpha \geq 1$ to be determined later), so u has sufficient time before it will time out on x. We will show that there still are enough candidates u so that the long-range contact will happen with high probability. Along the same lines, it is necessary to make the length of the middle interval depend on d (which it did not in the proof of Lemma 3.2), because otherwise, we could not guarantee that u would not time out on x during the middle interval.

The other change regards the subpath from u' to x'. Since all of the nodes on that subpath are quite far away from x, there is no risk of any of them timing out on x, so long as the time-out function is sufficiently large. Therefore, we need not invoke the induction hypothesis, but instead invoke Theorem 3.1 directly, which gives better guarantees.

We let g be the solution of the recurrence $g(d) = g(d^{\rho/2}) + \Theta(\log\log d) + \Theta(r \log^r d^{\rho/2} \cdot \log\log d^{\rho/2})$, and define the time-out function $h(d) = g(\alpha \cdot d^{2/\rho})$, for the same constant α as above. By a simple inductive proof, it can be verified that $g(d)$ (and hence $h(d)$ also) is in $O(\log^r d \cdot \log\log d)$.

Theorem 8.2 *Let x and x' be at distance d, and t, t' be two times with $t' - t = g(d)$. Then, with probability at least $\Omega(\log^{-r} d)$, there is a time-respecting, time-out free x-x' path (with respect to $h(\cdot)$) such that all nodes on the path are in the smallest ball containing x and x', and all time labels are in $[t, t')$.*

Proof. If d is smaller than some (sufficiently large) constant, the number of points x' within distance d of x is a constant, so in a constant number of rounds, we can obtain arbitrarily high constant probability of a call from x to x', which is a time-respecting path. By making the constant in the time-out function large enough, the path will also be time-out free. This takes care of the base case of the induction, and we assume from now on that d is larger than this constant.

The definition of g ensures that a time interval $[t, t')$ of length $g(d)$ can be divided into three disjoint parts $[t, s)$, $[s, s')$ and $[s', t')$ of respective lengths $g(d^{\rho/2})$, $\Theta(\log\log d)$, and $\Theta(r \log^r d^{\rho/2} \cdot \log\log d^{\rho/2})$.

For any node u within distance $d^{\rho/2}$ of x, the induction hypothesis guarantees a time-out free strictly time-respecting x-u path P during the interval $[t, s)$, with probability at least $\Omega(\log^{-r} d^{\rho/2})$. On the other hand, for any node u' within distance $d^{\rho/2}$ of x', we can apply Theorem 3.1, to show that there is a time-respecting (not necessarily time-out free) u'-x' path P' with probability at least $1 - O(\log^{-\Theta(r)} d^{\rho/2})$. Now, all that we need is a long-range contact from some u to some u' such that the concatenated path does not time out anywhere.

By choosing $d' = \frac{1}{\alpha} \cdot d^{\rho/2}$ with sufficiently large α (this is the α used in the definition of the time-out function $h(d)$), we obtain that there are $\Omega(\beta_1 d^{\rho/2})$ points u at distance between d' and $d^{\rho/2}$ from x. For instance, $\alpha \geq (\frac{2\beta_2}{\beta_1})^{1/D}$ will suffice. Now, analogous to the proof of Lemma 3.2, the probability that none of these nodes u calls any node u' within distance at most $d^{\rho/2}$ from x' is at most $e^{-\Omega(c\beta_1^2(s'-s))} = \log^{-\Theta(r)} d$.

By the Union Bound, the probability that one of the three parts (the long-range connection, the (inductive) time-out free path P or the path P') don't exist is at most $1 - \Omega(\log^{-r} d^{\rho/2}) + O(\log^{-\Theta(r)} d) + O(\log^{-\Theta(r)} d^{\rho/2})$. Noticing that $\log^{-r} d^{\rho/2} = 2\log^{-r} d$, and that $\log^{-\Theta(r)} d + \log^{-\Theta(r)} d^{\rho/2} \leq \log^{-r} d$ for sufficiently large d and constants in the $\Theta(r)$ expression, we can bound the probability of the non-existence of any of the three parts by $1 - \Omega(\log^{-r} d)$.

Finally, it remains to verify that the concatenated path $P \cdot (u, u') \cdot P'$ is time-out free. Because $d - d^{\rho/2} \geq d'$ for sufficiently large d, all nodes $v \in P'$, as well as u and u', are at distance at least d' from node x, and their departure time is at most t'. On the other hand, the departure time from x is at least t, and because $t' - t \leq g(d) = h(d')$, none of these nodes time out on x, so the resulting concatenated path is time-out free, completing the proof. ■

9 Conclusions

In this paper, we investigated a family of gossip distributions called *Spatial Gossip*, in which the probability for communication connections between two nodes depends on their distance in an inverse polynomial way. We saw that for appropriate choices of the degree of the polynomial, the resulting gossip mechanism has very good propagation behavior: the time for a node y to receive information originating with another node x depends poly-logarithmically on their distance $d_{x,y}$. Thus, spatial gossip combines the desirable features of *uniform gossip* and *flooding*: as for uniform gossip, information spreads exponentially fast, but it shares with flooding the fact that the time to inform a node depends only on the distance from the source, not the total number n of nodes.

We investigated the resource location problem as an example of an important practical application in which the power of spatial gossip becomes apparent. In the resource location problem, nodes should be kept informed about their closest copy of some desirable resource, in the presence of the arrival and departure of new copies. The strong guarantees we obtained for spatial gossip, including the dissemination time and the length of propagation paths, allowed us to design extremely simple protocols for solving the resource location problems while sending only small messages. The power of spatial gossip became particularly apparent from the fact that no protocol with small messages can solve the resource location problem by using uniform gossip as a communication mechanism.

This paper leaves open a number of interesting questions for future research. We currently do not have a good understanding of the kind of guarantees provided by spatial gossip with exponent $\rho = 2$. While we can show that the dissemination time is $O(d^{\varepsilon})$ for any $\varepsilon > 0$, we neither have a lower bound nor a poly-logarithmic upper bound. This case is of particular interest in view of the fact that a gossip mechanism basically equivalent to this case is used in the practically deployed system Astrolabe[15].

In the proof of the guarantees for Spatial Gossip, we also required (and used heavily) that the points be spaced "roughly uniformly" in a metric space. While we can prove the same guarantees in some special cases of non-uniform metric spaces, we are far from having a general proof. Using a similar approach to the one taken in Section 4, it is possible to define appropriate gossip distributions for more general metric spaces. It would be very interesting to have a proof for the general case.

Finally, there are many more problems in distributed computing which can be solved by gossip-based protocols. We believe that it would be worth investigating whether spatial gossip proves to be a more appropriate mechanism to build protocols on than uniform gossip.

Acknowledgements

We would like to thank two anonymous referees for many suggestions improving this paper.

References

[1] D. Agrawal, A. El Abbadi, and R. Steinke. Epidemic algorithms in replicated databases. In *Proc. 16th ACM Symp. on Principles of Database Systems*, 1997.

[2] K. Birman, M. Hayden, O. Ozkasap, Z. Xiao, M. Budiu, and 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. Sturgis, D. Swinehart, and D. Terry. Epidemic algorithms for replicated database maintenance. In *Proc. 7th ACM Symp. on Operating Systems Principles*, 1987.

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

[5] F. Göbel, J. Orestes Cerdeira, and H.J. Veldman. Label-connected graphs and the gossip problem. *Discrete Mathematics*, 87:29–40, 1991.

[6] I. Gupta, R. van Renesse, and K. Birman. Scalable fault-tolerant aggregation in large process groups. In *Proc. Conf. on Dependable Systems and Networks*, 2001.

[7] S. Hedetniemi, S. Hedetniemi, and A. Liestman. A survey of gossiping and broadcasting in communication networks. *Networks*, 18:319–349, 1988.

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

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

[10] R. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking. Randomized rumor spreading. In *Proc. 41st IEEE Symp. on Foundations of Computer Science*, 2000.

[11] D. Kempe and J. Kleinberg. Protocols and impossibility results for gossip-based communication mechanisms. In *Proc. 43rd IEEE Symp. on Foundations of Computer Science*, 2002.

[12] D. Kempe, J. Kleinberg, and A. Kumar. Connectivity and inference problems for temporal networks. *J. of Computer and System Sciences*, 64, 2002.

[13] M. Lin, K. Marzullo, and S. Masini. Gossip versus deterministic flooding: Low message overhead and high reliability for broadcasting on small networks. Technical Report CS99-0637, University of California at San Diego, 1999.

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

[15] R. van Renesse. Scalable and secure resource location. In *33rd Hawaii Intl. Conf. on System Sciences*, 2000.

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