Rumors Across Radio, Wireless, and Telephone [*]

Jennifer Iglesias [†]
Carnegie Mellon University
Pittsburgh, USA
jiglesia@andrew.cmu.edu

Rajmohan Rajaraman
Northeastern University
Boston, USA
rraj@ccs.neu.edu

R. Ravi
Carnegie Mellon University
Pittsburgh, USA
ravi@andrew.cmu.edu

Ravi Sundaram
Northeastern University
Boston, USA
koods@ccs.neu.edu

July 13, 2015

Abstract

We study the problem of computing a minimum time schedule to spread rumors in a given graph under several models: In the radio model, all neighbors of a transmitting node listen to the messages and are able to record it only when no other neighbor is transmitting; In the wireless model (also called the edge-star model), each transmitter is at a different frequency to which any neighbor can tune to, but only one neighboring transmission can be accessed in this way; In the telephone model, the set of transmitter-receiver pairs form a matching in the graph. The rumor spreading problems assume a message at one or several nodes of the graph that must reach a target node or set of nodes. The transmission proceeds in synchronous rounds under the rules of the corresponding model. The goal is to compute a schedule that completes in the minimum number of rounds.

We present a comprehensive study of approximation algorithms for these problems, and show several reductions from the harder to the easier models for special demands. We show a new hardness of approximation of $\Omega(n^{\frac{1}{2}-\epsilon})$ for the minimum radio gossip time by a connection to maximum induced matchings. We give the first sublinear approximation algorithms for the most general case of the problem under the wireless model; we also consider various special cases such as instances with symmetric demands and give better approximation algorithms. Our work exposes the relationships across the models and opens up several new avenues for further study.

1

1 Introduction

Problems modeling rumor spread are central to the design of coordination networks that seek to keep demand pairs of vertices in contact over time. The prototypical example is the *broadcast* problem where a message in a root node must be sent to all the other nodes via connections represented by an undirected graph. We assume that communication proceeds in synchronized rounds. When more than one message is being disseminated, we assume that in each round each node can transmit an unlimited number of messages in one communication. A subset generalization of broadcast is called the *Multicast* problem: a subset of nodes is specified as terminals and the goal is to spread the rumor from the root only to this subset, using other non-terminal nodes if needed in the process. An all-to-all generalization of the broadcast problem is termed *gossip*: every node has its own piece of information that must be communicated to all nodes, and the goal is to have all the information spread to all the nodes in the minimum number of rounds. Gossip and broadcast are special cases of a more general demand model that we may call *multicommodity multicast*: in this most general version, we are given a set of source-sink pairs so that each source has a rumor that must be sent to the corresponding sink. Recall that messages from many sources can all be aggregated and exchanged in one round between any pair that can communicate, and the goal is to minimize the number of rounds. In this paper, we will study a specialization of the multicommodity demand model called the *symmetric multicommodity* where for every source-sink pair, we also have the symmetric requirment that the sink wants to send its rumor to the source; thus, the demand pairs are unordered in this case. The more general version will be called the *asymmetric multicommodity* demand model to distinguish it from the symmetric demands case.

1.1 Models: Telephone, Radio, and Edge-Star, a New Model from Wireless

Different communication models result in different constraints on the set of edges on which messages can be transmitted in a single round. The two most widely studied models are the telephone and radio models: In the **telephone** model, in each round, a node can communicate with at most one other node, thus the edges on which communication occurs is a matching; In the **radio** model, a set of transmitters broadcast the message out but only their neighbors who are adjacent to exactly only one transmitter can successfully receive the message (while interference prevents other neighbors from receiving the message): the set of edges through which the messages are sent in any round in this model is a set of stars centered at the transmitters, where each leaf of each star has that star's center as its unique neighbor among all the star centers.

In this paper, we expand the study of rumor spreading problems by introducing a new model based on wireless communications between nodes, which we call the **edge-star** model. We assume that during each round of wireless communication, each transmitter can choose its own channel or frequency distinct

2

from that of all other transmitters. The input undirected graph represents pairs of nodes that are within wireless range of each other. Receiving nodes that are in the vicinity of many different transmitting nodes can choose to tune into the frequency of one of them. In this way, the set of edges in which communication happens in every round is a set of stars which are defined a subset of edges of the input graph. Note that unlike the radio model, there is no requirement that a receiver be adjacent to exactly one transmitter.

1.2   Previous Work

The radio broadcast and gossip problems have been extensively studied (see the work reviewed in the survey [10]). The best-known scheme for radio broadcast is by Kowalski and Pelc [11] which completes in time $O(D + \log^2 n)$, where n is the number of nodes, and D is the diameter of the graph and is a lower bound to get the message across the graph from any root. The $O(\log^2 n)$ term is also unavoidable as demonstrated by Alon et al. [1] in an example with constant diameter that takes $\Omega(\log^2 n)$ rounds for an optimal broadcast scheme to complete. Elkin and Korsartz [5] also show that this additive log-squared term is best possible unless $NP \subseteq DTIME(n^{\log \log n})$.

The best bound for radio gossip known so far, however, is $O(D + \Delta \log n)$ steps in an n-node graph with diameter D and maximum degree Δ [9]. The maximum degree is not a lower bound on the gossip time, and indeed no previous results are known about the approximability for radio gossip, which is mentioned as an open problem in [10].

In the telephone model, the first poly-logarithmic approximation for minimum broadcast time was achieved by Ravi [13] and the current best known approximation ratio is $O(\frac{\log n}{\log \log n})$ due to Elkin and Korsartz [6]. The best known lower bound on the approximation ratio for telephone broadcast is $3 - \epsilon$ [4].

In his study of the telephone broadcast time problem, Ravi [13] introduced the idea of finding low poise spanning trees to accomplish broadcast: the *poise* of a spanning tree of an undirected graph is the sum of its diameter and its maximum degree. In the course of deriving a poly-logarithmic approximation, Ravi also showed how a tree of poise P in an n-node graph can be used to complete broadcast starting from any node in $O(P \cdot \frac{\log n}{\log \log n})$ steps - we will use this observation later.

1.3   Our contributions

We give the first results on the approximability of gossip and multicommodity multicast problems in the radio model. We introduce the edge-star model based on wireless channels and give the first approximation results for minimum time rumor spreading by relating them to their analogs in the telephone model.

1. We show that it is NP-hard to approximate gossip in the radio model within a factor of $O(n^{1/2-\epsilon})$ in an n-node graph. This result is derived by isolating a gathering version of the broadcast problem in the radio

3

model and relating it in a simple bipartite graph to induced matchings (Section 2).

2. We obtain an $O(\frac{\log n}{\log \log n})$ approximation algorithm for gossip in the edge-star model by reducing the problem to the broadcast problem in the telephone model (Section 3.1).

3. We consider the special case where the underlying graph is a tree, and show that the multicommodity multicast in the edge-star model reduces to the broadcast problem in the telephone model, thus proving an $O(\frac{\log n}{\log \log n})$ approximation (Section 3.2).

4. We show that the case of edge-star symmetric multicommodity multicast problem has the same optimal solution (up to poly-log factors) as telephone multicommodity multicast, yielding a $2^{O(\log \log n \sqrt{\log n})}$ approximation (Section 3.3).

5. We give an $O(n^{\frac{2}{3}})$-approximation for the general (asymmetric) multicommodity multicast problem in the edge-star model (Section 3.4).

Table 1 contains a summary of our results in context.

	Broadcast	Gossip	Multicommodity
Radio	$D + O(\log^2 n)$ [11]	$O(D + \Delta \log n)$ [9]	Unknown
		$\Omega(n^{1/2-\epsilon})$ hard*	$\Omega(n^{1/2-\epsilon})$ hard*
Edge-star	OPT$= D$	OPT$\cdot O(\frac{\log n}{\log \log n})$*	OPT$\cdot \tilde{O}(2^{\sqrt{\log n}})$*(symmetric)
			OPT$\cdot O(n^{\frac{2}{3}})$* (asymmetric)
Telephone	OPT$\cdot O(\frac{\log n}{\log \log n})$ [7]	OPT$\cdot O(\frac{\log n}{\log \log n})$ [7]	OPT$\cdot \tilde{O}(2^{\sqrt{\log n}})$ [12]

Table 1: A summary of upper and lower bounds achieved in the different problems. We prove the results marked * in the table.

Remark: As we were preparing this submission, we learned of recent independent results related to the radio gossip problem. Halldorsson et al. (private communication, Halldorsson, July 2015) recently studied the Radio Aggregation Scheduling problem which is a gathering version of the rumor spreading problem in the radio model. The set of edges in which communication occurs in every round is a matching with the additional property that if the edges within receivers and within senders are ignored, the communicating edges form an induced matching. In this model they prove a tight $\Theta(n^{1-\epsilon})$-approximation for their radio aggregation scheduling. Our results were derived independently and we have not seen a description of their methods.

4

2   Lower bound for gossip in the radio model

In this section, we show it is NP-hard to approximate gossip in the radio model within a factor of $O(n^{1/2-\epsilon})$. This also implies the same hardness result for multicommodity multicast under the radio model, because gossip is a special case of multicommodity multicast. In order to show these hardness results, we first consider the smallest set of induced matchings which cover the vertices of a bipartite graph.

Definition 1 *An* induced matching *is a matching of some vertices U in a graph G, such that $G[U]$ is a matching. (We use $G[U]$ to mean the graph G induced on the vertex set U.) In other words, in the graph G only the matching edges are present between the nodes in U.*

A covering set of induced matchings *(CSIM) is a set of induced matchings which cover all the vertices in the graph. The size of covering set of induced matchings is defined to be the number of induced matchings.*

First, we will show the hardness of finding a minimum CSIM by a reduction from coloring. Then we will use the hardness of minimum sized CSIM to prove the hardness results for radio gossip.

Theorem 1 *It is NP-hard to approximate CSIM to within a $n^{1/2-\epsilon}$ factor for any constant $\epsilon > 0$.*

Proof: Given a coloring instance $G = (V, E)$, we first turn this into a bipartite graph, where we want to find a CSIM. For each $v \in V$ we make $n + 1$ copies of v in each side of the partition; $v_1^L, v_2^L, \ldots v_{n+1}^L$ for L and $v_1^R, v_2^R, \ldots v_{n+1}^R$ for R. We use the edges $E_v = \{(v_i^L, v_i^R) | v \in V, i \in [n+1]\}$ also referred to as the straight edges and also $E_e = \{(u_i^L, v_j^R) | uv \in E, i, j \in [n+1]\}$ called the cross edges. Now $G' = (L, R, E_v \cup E_e)$ is the bipartite graph for which we want to find a CSIM. Figure 2 shows an example construction.

Let χ be the number of colors in an optimal coloring in G. Let λ be the number of sets in a minimal CSIM in G'.

We now show that $\lambda \leq \chi \leq n$. Let C_i be a set of vertices of color i in the coloring. If we take the edges $\{(v_j^L, v_j^R) | v \in C_i, j \in [n+1]\}$, they are an induced matching. Each vertex has one straight edge in G', and if a vertex is used in the matching, then its straight edge is used. So, we only need to show that no cross edges go between vertices in this matching. If a cross edge (u_j^L, v_k^R) did exist, then $(u, v) \in E$ but then u, v couldn't be the same color. So, for each color we have defined an induced matching. These induced matchings cover all the nodes since every node receives some color in the coloring on G'.

Now we will show that $\chi \leq \lambda$ or $n + 1 \leq \lambda$. Let $S_1, S_2, \ldots S_\lambda$ be the induced matchings covering G'. Assume that there is some $v \in V$ that has all of its corresponding vertices in G' matched via cross edges. Then we can only have at most one cross edge per matching. If an induced matching has (v_i^L, u_ℓ^R) and $(v_j^L, w_{\ell'}^R)$ then this is not an induced matching since (v_j^L, u_ℓ^R) is an edge. Therefore in this case to match all the v_i^L in some induced matching, we will
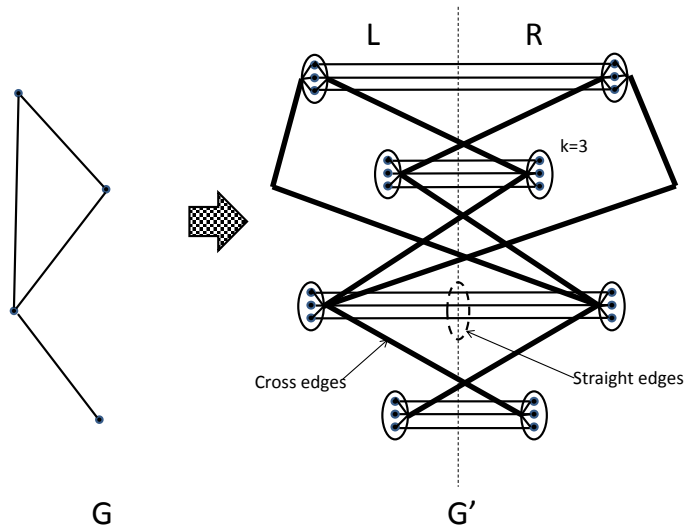
Figure 1: Here is an example of the construction of G' from G. The thick edges represent complete bipartite subgraphs.

need at least $n + 1$ induced matchings. Now consider each $v \in G$ has one of its straight edges used in some induced matching. Let S_j be the first induced matching containing a straight edge adjacent to some v_L^i. In S_j, because some v_i^L is matched via its straight edge, then no v_ℓ is matched via a cross edge. So, in G color v with the jth color. This is a valid coloring. If some (v_i^L, v_i^R) and (u_ℓ^L, u_ℓ^R) were both in the same induced matching, then there can't be the edge (u, v) in the original graph G.

Combining the above two parts we get that $\chi = \lambda$.

We begin with a graph G such that it is NP-hard to distinguish if there is a coloring of size $|V(G)|^\epsilon$ from if the coloring requires at least $|V(G)|^{1-\epsilon}$ colors [8]. Therefore, in the graph G' we created, it is NP-hard to distinguish if there is a set of induced matchings that cover the vertices of size n^ϵ or n. We have n^2 vertices in G' though. So, in a graph with n vertices it is NP-hard to approximate the number of induced matchings needed to cover the vertices within a factor of $O(n^{1/2-\epsilon})$. $\qquad \square$

Now that we have developed the hardness result for CSIM, we will use the graph we created for CSIM, to create instances of radio gossip.

Corollary 1 *It is NP-hard to approximate radio gossip to within a $n^{1/2-\epsilon}$ factor for any constant $\epsilon > 0$.*

Proof: We convert the induced matching instance to a gossip problem in a similar fashion to above. We can consider that we have the bipartite graph

G' and we build a complete binary tree with it's leaves being the v_i^L. Now the terminal nodes in the gossip problem are all the nodes. Now to communicate the message to all other nodes, then each node v_i^R must at some point be the only node trying to talk to some node on the other side of the bipartition. In other words, we need to have induced matchings at each point in order for the v_i^R to propogate their messages to some other node without interference. Therefore, we need at least as many induced matchings as it takes to cover the graph to complete the gossip. Call this number C; we can now achieve gossip in time $2C + 3 \log n$ as follows. We do this by doing the induced matchings so that each vertex v_i^R communicates its message to someone on the other side of the partition. Next we propagate the message up the binary tree to the root node. This takes time at most $2 \log n$ since at each node of the path in the binary tree, a message can be delayed only for two step, and the path length is logarithmic. Then we broadcast the message down the tree. This takes time $\log n$ since we can use the edge-star model to just broadcast all the gathered messages from the root along the down-stars in one time step per leve;. Lastly, we need to communicate the message back to the v_i^R, which takes time C. We know that radio gossip takes time at least C and can be done in time $2C + 3 \log n$ on this graph.

Therefore, it is NP-hard to approximate radio gossip better than a factor of $O(n^{1/2-\epsilon})$ otherwise, we could approximate the CSIM within the same factor.

$\square$

3   The Edge-Star Model

In this section, we consider the edge-star model which generalizes the telephone model. We focus on three specific classes of problems; gossip, symmetric multicommodity multicast, and asymmetric multicommodity. In the *symmetric multicommodity* problem, we are given a set of demand pairs, and if (s_i, t_i) is a demand, then (t_i, s_i) is also a demand. In the asymmetric multicommodity case, there are no restrictions on which demand pairs are present.

In Section 3.1, we first obtain an $O(\frac{\log n}{\log \log n})$ approximation algorithm for gossip in the edge-star model by reducing the problem to the broadcast problem in the telephone model. Next, in Section 3.2, we consider the special case where the underlying graph is a tree. In this special case, then we show that the multicommodity multicast in the edge-star model reduces to the broadcast problem in the telephone model, yielding an $O(\frac{\log n}{\log \log n})$ approximation. In Section 3.3, we show that the case of edge-star symmetric multicommodity multicast problem has the same optimal solution (up to poly-log factors) as telephone multicommodity multicast, yielding an $\tilde{O}(2^{\sqrt{\log n}})$ approximation. Lastly, in Section 3.4, we give an $O(n^{\frac{2}{3}})$-approximation for the general (asymmetric) multicommodity multicast problem in the edge-star model.

3.1 Gossip

Here we show an $O(\frac{\log n}{\log \log n})$ approximation for edge-star gossip. First, we show that a solution to the gossip problem in the edge-star model gives a solution to the broadcast problem in the telephone model of the same length. Next, we show that using a solution for the broadcast problem in telephone we can get a solution of twice the length to the gossip problem in the edge-star model. This show that their optimal solutions differ in cost by a factor of at most two.

Lemma 1 *The optimal broadcast time in the telephone model is no more than the optimal gossip time in the edge-star model.*

Proof: Let $\mathcal{S}$ denote an optimal schedule for gossip in the edge-star model that completes in T rounds. Let r denote the root node for the broadcast problem in the telephone model. Fix a node v. Let P_v denote a path taken by the message from v to arrive at r in the schedule S. Let E_t denote the set of all directed edges in $\cup_v P_v$ that are activated in round t in S. By definition of the edge-star model, if (u_1, v_1) and (u_2, v_2) are in E_t, then $v_1 \neq v_2$. Furthermore, by our choice of the paths, we obtain that (i) for any distinct (u_1, v_1) and (u_2, v_2) in E_t, $u_1 \neq u_2$; and (ii) the edges of P_v appear in order of increasing time in the collection of E_ts.

   We now argue that a reverse schedule in which the activated sets are given by $E'_t = \text{REV}(E_{T-t})$ forms a broadcast schedule from the root, where $\text{REV}(X)$ equals $\{(v, u) : (u, v) \in X\}$ for any set X of directed edges. In any round t, for any distinct (u_1, v_1) and (u_2, v_2) in E_t, we have $u_1 \neq u_2$ and $v_1 \neq v_2$; therefore, $\text{REV}(E_t)$ is a matching. Since the edges of P_v appear in order of increasing time in the collection of E_ts, the edges of the $\text{REV}(P_t)$ appear in order of increasing time in the collection of E'_ts. Consequently, the message from the root is delivered to each node in T rounds. □

Lemma 2 *The optimal gossip time in the edge-star model is no more than twice the optimal broadcast time in the telephone model.*

Proof: The proof mirrors the proof of Lemma 1. Let $\mathcal{S}$ denote an optimal schedule for broadcast from root r in the telephone model that completes in T rounds. Fix a node v. Let P_v denote a path taken by the message from r to arrive at v in the schedule S. Let E_t denote the set of all directed edges in $\cup_v P_v$ that are activated in round t in S. By definition of the telephone model, for distinct (u_1, v_1) and (u_2, v_2) in E_t, $u_1 \neq u_2$ and $v_1 \neq v_2$. Furthermore, by our choice of the paths, we obtain that the edges of P_v appear in order of increasing time in the collection of E_ts.

   We now argue that a reverse schedule in which the activated sets are given by $E'_t = \text{REV}(E_{T-t})$ forms a schedule for gathering in the edge-star model. In any round t, for any distinct (u_1, v_1) and (u_2, v_2) in E_t, we have $u_1 \neq u_2$ and $v_1 \neq v_2$; therefore, $\text{REV}(E_t)$ is a matching, and is a valid set of edges to activate in the edge-star model in round $T - t$. Since the edges of P_v appear in order of increasing time in the collection of E_ts, the edges of the $\text{REV}(P_t)$ appear in order

8

of increasing time in the collection of E'_is. Consequently, the message from any node v is delivered to the root in T rounds.

Once the root has all the messages, we can complete the gossip by running the broadcast schedule. Since any schedule in the telephone model is valid in the edge-star model, it follows that this broadcast completes in T rounds. We thus have a gossip schedule that completes in the edge-star model in $2T$ rounds.
$\square$

There exists an $O(\frac{\log n}{\log \log n})$ approximation for telephone broadcast [7]. Therefore this same approximation holds for the edge-star gossip problem.

3.2 Multicommodity multicast on a tree

In this part, we consider the multicommodity multicast problem in the edge-star model in the special case where our host graph is a tree. Here we give a reduction to telephone broadcast. When the host graph is a tree, the path taken by any message is known, so we simply need to coordinate the communications.

Lemma 3 *There is an $O(\frac{\log n}{\log \log n})$ approximation for the edge-star multicommodity multicast problem in a tree.*

Proof: We will start by choosing some vertex r to be the root of the tree. Let the optimal solution take time D (we can try all $2n$ possible values for D only losing a polynomial factor in runtime). Now for each demand pair, (s_i, t_i) the message will have to go from s_i to $lca(s_i, t_i)$, and then from the $lca(s_i, t_i)$ down to t_i. Bringing all the messages down the tree from $lca(s_i, t_i)$ to t_i can be done in time $D + 1$; we spend $D + 1$ time steps alternating between the odd layers broadcasting their messages down and the even layers broadcasting their message down. SInce each layer is a collection of edge-disjoint stars, it can be implemented in one round in the edge-star model.

The hard part is bringing the messages up from s_i to $t'_i = lca(s_i, t_i)$. So, we will consider that we simply have the constraints of the form (s_i, t'_i). First we will break the tree up into sets of $2D$ consecutive layers starting every D layers. This guarantees that every constraint (s_i, t_i) is in some set of $2D$ layers.

Now consider some $2D$ layers in the tree. Look at the union of all the (s_i, t'_i) paths in these layers. These form a forest, where each tree has depth at most $2D$ and each node has a max degree of D; so each tree has poise $4D$ (recall that the poise is the sum of the maximum degrree and the diameter). Therefore these trees can gather to their uppermost node in time $O(D\frac{\log n}{\log \log n})$.

We can run all the gathers to satisfy (s_i, t'_i) in two groups; we run every other set of $2D$ layers in the tree simultaneously as they are disjoint layers. Hence, in time $O(D\frac{\log n}{\log \log n})$, we can satisfy the demands (s_i, t'_i). In time $D + 1$, then we can satisfy the demands (t'_i, t_i). Therefore in time $O(D\frac{\log n}{\log \log n})$ we satisfy all the (s_i, t_i) demands. $\square$

3.3   Symmetric Multicommmodity Multicast

Note that the symmetric multicommodity multicast problem in the telephone model is equivalent (within constant factors) to the general multicommodity multicast problem [13, 3] for which an $O(2^{\sqrt{\log k}})$ approximation algorithm is known, where k is the number of terminals [12]. We show a reduction from the symmetric multicommodity multicast problem in the edge-star model to the symmetric multicommodity multicast problem in the telephone model, losing an additional $O(\frac{\log^3 n}{\log \log n})$ factor in the approximation ratio in an n-node graph.

Theorem 2 *Given a ρ-approximation for the symmetric multicommodity multicast problem on k terminal pairs in an n-node undirected graph under the telephone model, we can design an $O(\rho \cdot \log^2 k \cdot \frac{\log n}{\log \log n})$ approximation for the same problem in the edge-star model.*

Proof: Given an optimal solution to symmetric multicommodity multicast in the edge-star model, we demonstrate a solution to the symmetric multicommodity multicast problem in the telephone model with a poly-log multiplicative loss in performance. Consider an input instance with demand pairs $\{s_i, t_i\}$ for $i = 1 \cdots k$ on an undirected graph G. Consider an optimal schedule for the edge-star symmetric multicommodity multicast problem on this instance. This defines for each pair s, t, a pair of paths from one node to the other where the edges of the paths are labeled in increasing time order denoting the periods in which these edges participated in an information transmission. Suppose the optimal time for multicasting is L; then these paths are of length at most L. Also, given the in-degree one bound for the edge-star model (each receiver can listen to at most one transmitter in this wireless model), the indegree of the sugraph representing the union of these optimal transmissions is also at most L. Our goal is to use these paths to aggregate the messages from a set of these pairs into a subset of carefully selected terminals using a reverse broadcast scheme, and then transmit the aggregated messages back to the corresponding mates of these sources. Both these steps of gathering and sending will be accomplished using multicommodity multicast instances in the telephone model.

To define the aggregation pattern, define an auxiliary graph H with one node per *demand pair s_i, t_i*. This graph is only for the sake of argument so we will use optimal paths in the edge-subgraph multicommodity multicast scheme in defining it. Note that the optimal transmission paths for a pair represent two paths: one from s_i to t_i and the second from t_i to s_i, where these two paths may share edges. Concatenated together they define what we will call an "optimal cycle" for this pair. Define an edge between two pairs if their optimal cycles intersect at a node. In Figure 2, we can see an example of when optimal cycles intersect. Thus H defines the conflict or interference between the demand pairs in the optimal multicommodity multicast schedule in the edge-subgraph model.

We now use a network decomposition procedure [2] on H to decompose the k demand pairs into $\log_2 k$ disjoint layers with the following property: the set of nodes in each layer can be decomposed into node-disjoint shallow trees, i.e., each tree in one of the layers has diameter at most $2 \log_2 k$. This decomposition
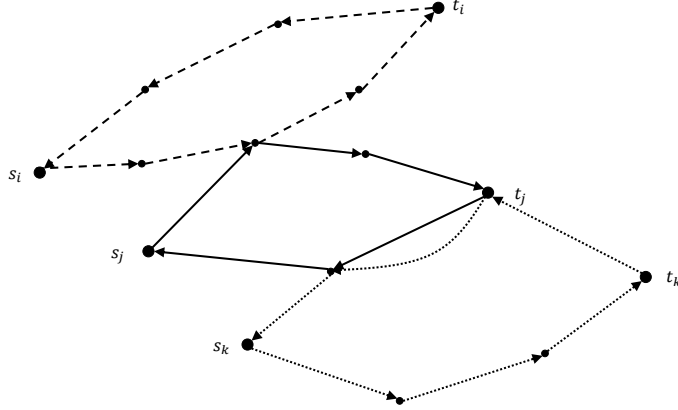
10

Figure 2: Here is an example of the optimal paths between some (s_i, t_i) pairs. Here we see that the (s_j, t_j) pair intersects (s_i, t_i) and (s_k, t_k), but (s_i, t_i) and (s_k, t_k) do not intersect.

is done as follows: pick any vertex v in H and build a BFS tree from v. Now let i be the smallest depth such that the number of nodes at depth i or less is more than the number of nodes at depth $i + 1$. Put v and everything within distance i of v into the current layer. Now remove v and it's BFS tree up to depth $i + 1$ from H. Repeat this process to form each layer. Once H is empty, let U be the vertices not yet assigned to a layer. Then start forming a new layer from the graph $H[U]$.

This process assigns at least half of the remaining nodes to the current layer, hence we build at most $\log_2 k$ layers. The diameter of each component in a layer is at most $2 \log_2 k$, because as we move down the BFS tree the number of nodes contained in it double at each step.

Now we can use these layers to define our gathering problems. Consider one layer i and one tree $T_{i,j}$ in this layer in the decomposition. This represents a shallow subgraph in H, so let us root this at a demand pair denoted P_{ij}. By following the paths in this subgraph from every other node to P_{ij}, we can replace their intersections with paths in the optimal multicast originating at each terminal s in any of the pairs to one of the two terminals, say t_{ij} in the pair P_{ij}. This defines one of the gathering trees gathering to the terminal t_{ij}. By construction, the in-degree of any node in the gathering tree is at most L and the distance from any node to the root t_{ij} is at most $O(L \log k)$. Note that by the disjointness of the subgraphs in one layer i, all the gather trees are node disjoint. For each gather tree T_{ij}, we now set up a gathering multicast problem with all the terminals in the tree going to the root t_{ij}. Note that since each

tree has total degree + diameter at most $O(L \log k)$, the poise of each tree is bounded by $O(L \log k)$ and thus each of these trees has a gathering schedule in the telephone model taking at most $O(Poise \cdot \frac{\log n}{|log \log n})$ steps in an n-node graph [13]. This gives a feasible solution to the set of all gathering problems in one layer i running in time $O(L \cdot \log k \cdot \frac{\log n}{\log \log n})$. Repeating this over the layers finally gives a set of gathering problems in the telephone model that complete in total time $O(L \cdot \log^2 k \cdot \frac{\log n}{\log \log n})$.

Note that the same schedules can be reversed to send all the gathered information in each tree to all the terminals in a tree finishing the requirements. Employing a ρ-approximation for this multicommodity multicast problem in the telephone model proves the theorem. □

3.4   Asymmetric Multicommodity Multicast

For the edge-star asymmetric multicommodity multicast problem, we will use the network decomposition used in the previous proof, along with telephone broadcast in trees with small poise.

Theorem 3 *There is an $\tilde{O}(n^{\frac{2}{3}})$-approximation for the asymmetric multicommodity multicast problem in the edge-star model.*

Proof: We develop the algorithm in two phases. First, we design an $O(\sqrt{p})$-approximation algorithm for the case with p demand pairs (note that p can be up to $O(n^2)$ in an n-node graph). Then we combine this with an algorithm that satisfies all the demands in the in-neighborhood of a node in the demand graph with high indegree to get the final result.

A Greedy Algorithm. To design the $\tilde{O}(\sqrt{p})$-approximation algorithm, we use a greedy method: assume that the value of the optimal multicast time is L (we can try all the $2n$ possible guesses in parallel to dispense this assumption with a polynomial running-time overhead). For every unsatisfied demand pair (s_i, t_i) (note that demand pairs are ordered in the asymmetric case), we look for a path of length at most L from s_i to t_i. If we find one, we add it to the greedy collection and delete all the nodes in this path. Suppose we are able to collect g paths for the pairs denoted G in the greedy phase until we can find no more paths of small length for the remaining demands.

Now it must be the case that all optimal paths for the remaining demands in $P \setminus G$ must intersect the greedy paths. This implies that for every demand pair (s, t) in $P \setminus G$, we can follow its optimal path to its intersection with one of the greedy paths, say for the pair (s_i, t_i), and then continue in the greedy path to t_i. In this way, every demand source in $P \setminus G$ can be routed and assigned to one of the sinks in the greedy pairs G in a collection of paths: each such path has length at most $2L$ (coming from at most L steps to the intersection with the greedy collection and another L from the intersection to the sink at the end of this greedy path); also the indegree of the collection of these paths is at most $L+1$ since they arise from the optimal collection plus the greedy subgraph which adds at most one to each node's indegree. We now set up a dummy broadcast

problem (following Nikzad and Ravi [12]) by hooking up the set of sinks at the end of the greedy paths, say $T(G)$, as leaves in a complete binary tree with new dummy nodes and a dummy root t. We solve for the broadcast problem in this graph from the dummy root t to all the sources s_i in all the pairs. By the above construction, there exists a tree of poise $O(L + \log n)$ that connects all the sources to this root. From the result of Ravi [13], this implies a broadcast scheme completing in $O(L \frac{\log n}{\log \log n})$. Using an α-approximation algorithm for broadcast in the telephone model, we get a tree that assign the sources in P to the sinks in $T(G)$ in $O(\alpha \cdot L \frac{\log n}{\log \log n})$ steps. Let us denote the set of sinks in $T(G)$ by $t'_1, \cdot, t'_g$ and the set of sources assigned to a sink t'_i by S_i.

The remaining task is to send back the messages gathered from S_i at t'_i to the sinks corresponding to the sources in S_i - let us denote this sink set by T_i. Note that by construction, all the sinks in T_i are at a distance at most $O(\alpha \cdot L \frac{\log n}{\log \log n})$ from t'_i by following the paths to the corresponding source s and then concatenating the undirected path to its mate t. However, these local broadcasts must obey the edge-subgraph condition of having indegree at most one which is tricky to enforce.

If the number of greedy pairs $g = |G|$ is at least $\sqrt{p}$, we simply satisfy these pairs and move to the next iteration: the number of such iterations is at most $\sqrt{p}$ and each iteration can be implemented in $O(L)$ steps (running the disjoint greedy path schedules in parallel). If the number of pair is less than $\sqrt{p}$, we can carry out the broadcast from each greedy sink t'_i to its sink set T_i in time $O(\alpha \cdot L \frac{\log n}{\log \log n})$ by reversing the gathering in the earlier broadcast tree and extending it to the corresponding sinks. Processing these trees one after another, we use a total of $O(\sqrt{p} \cdot \alpha \cdot L \frac{\log n}{\log \log n})$. Since α is sublogarithmic [6], we finally get an $\tilde{O}(\sqrt{p})$-approximation as claimed.

A Local Algorithm. For the second ingredient we observe that if the in-degree of any node v in the demand graph is δ, then we can satisfy all the demand requirements of the predecessors of v in the demand graph $In(v)$ in time $\tilde{O}(L)$. Note that since all the terminals in $In(v)$ send their message to v, the union of the directed paths that transmit these messages in the optimal solution have distance at most L from the terminals to v and induce an in-degree of at most L. This defines a tree of poise $O(L)$ and hence enables us to find a broadcast scheme that gathers all the messages from $In(v)$ at v in time $\tilde{O}(L)$. By reversing this broadcast tree and then following the optimal paths from each terminal in $In(v)$ to its other sinks, we can find a tree of depth (not poise) at most $\tilde{O}(L)$ rooted at v where these messages are gathered. Since v is the only node sending out the gathered messages, we can send all these messages to their intended sinks in a breadth-first tree in time $\tilde{O}(L)$. Note that we have taken care of all the demands originating in $|In(v)|$ nodes.

Combining the two algorithms. We can now combine the two algorithms as follows: As long as p, the number of demand pairs in the n-node graph, is at least $\Omega(n^{\frac{4}{3}})$, we use the local algorithm. By averaging over the indegrees that partition the demand pairs, there exists a node of indegree at least $\Omega(n^{\frac{1}{3}})$ in the demand graph. The local algorithm thus satisfies the demands originating

in at least this many nodes in one iteration. The number of iterations is thus at most $n^{\frac{2}{3}}$ each taking $\tilde{O}(L)$ multicast steps. On the other hand, when p drops below $O(n^{\frac{4}{3}})$, we use the greedy algorithm to get an approximation ratio of $\tilde{O}(\sqrt{p}) = \tilde{O}(n^{\frac{2}{3}})$ giving the result. $\qquad\Box$

4 Conclusion

We have obtained new results in the approximability of rumor spreading problems in the well-studied radio model as well as a new model motivated by wireless communications, which we call the edge-star model. For the radio model, we present an $\Omega(n^{1/2-\epsilon})$ hardness of approximation bound for radio gossip, making progress on an open problem mentioned in [10]. For the edge-star model, we present an $O(\log n/\log\log n)$ approximation algorithm for gossip, an $O(2^{\sqrt{\log n}\log\log n})$ approximation algorithm for symmetric multicommodity multicast, and an $O(n^{2/3})$ approximation algorithm for asymmetric multicommodity multicast. Our approximation algorithms expose relationships between the edge-star model and the well-studied telephone model.

Our work leaves several interesting open problems. Among the nine cells listed in the matrix of Table 1 of Section 1, only radio broadcast and edge-star broadcast are resolved. Significant gaps between the best known upper and lower bounds on approximability remain for telephone broadcast, the gossip problem under all three models, and the multicommodity multicast problem under all three models. In the edge-star model, the symmetric and asymmetric versions of the multicommodity multicast problem are distinct, and both are open, in terms of the best approximation factor achievable in polynomial-time.

References

[1] N. Alon, A. Bar-Noy, N. Linial, and D. Peleg. A lower bound for radio broadcast. *Journal of Computer and System Sciences*, 43:290298, 1991.

[2] B. Awerbuch and D. Peleg. Sparse partitions. In *Foundations of Computer Science, 1990. Proceedings., 31st Annual Symposium on*, pages 503–513. IEEE, 1990.

[3] A. Bar-Noy, S. Guha, J. Naor, and B. Schieber. Multicasting in heterogeneous networks. In *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing, Dallas, Texas, USA, May 23-26, 1998*, pages 448–453, 1998.

[4] M. Elkin and G. Kortsarz. A combinatorial logarithmic approximation algorithm for the directed telephone broadcast problem. *SIAM J. Comput.*, 35(3):672–689, 2005.

[5] M. Elkin and G. Kortsarz. Polylogarithmic additive inapproximability of the radio broadcast problem. *SIAM J. Discrete Math*, 19(4):881–899, 2005.

[6] M. Elkin and G. Kortsarz. Sublogarithmic approximation for telephone multicast. *J. Comput. Syst. Sci.*, 72(4):648–659, 2006.

[7] M. Elkin and G. Kortsarz. An improved algorithm for radio broadcast. *ACM Transactions on Algorithms (TALG)*, 3(1):8, 2007.

[8] U. Feige and J. Kilian. Zero knowledge and the chromatic number. *J. Comput. Syst. Sci.*, 57:187–199, 1998.

[9] L. Gasieniec, D. Peleg, and Q. Xin. Faster communication in known topology radio networks. *Distributed Computing*, 19(4):289–300, 2007.

[10] L. Gsieniec. On efficient gossiping in radio networks. In S. Kutten and J. erovnik, editors, *Structural Information and Communication Complexity*, volume 5869 of *Lecture Notes in Computer Science*, pages 2–14. Springer Berlin Heidelberg, 2010.

[11] D. R. Kowalski and A. Pelc. Optimal deterministic broadcasting in known topology radio networks. *Distributed Computing*, 19(3):185–195, 2007.

[12] A. Nikzad and R. Ravi. Sending secrets swiftly: Approximation algorithms for generalized multicast problems. In *Automata, Languages, and Programming*, pages 568–607. Springer, 2014.

[13] R. Ravi. Rapid rumor ramification: Approximating the minimum broadcast time. In *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on*, pages 202–213. IEEE, 1994.