

Average-time complexity of gossiping in radio networks

Bogdan S. Chlebus^{1*}, Dariusz R. Kowalski², and Mariusz A. Rokicki^{1*}

¹ Department of Computer Science and Eng., UCDHSC, Denver, CO 80217, USA.

² Department of Computer Science, University of Liverpool, Liverpool L69 7ZF, UK.

Abstract. Radio networks model wireless synchronous communication with only one wave frequency used for transmissions. In the problem of many-to-all (M2A) communication, some nodes hold input rumors, and the goal is to have all nodes learn all the rumors. We study the average time complexity of distributed many-to-all communication by deterministic protocols in directed networks under two scenarios: of combined messages, in which all input rumors can be sent in one packet, and of separate messages, in which every rumor requires a separate packet to be transmitted. Let n denote the size of a network and k be the number of nodes activated with rumors; the case when $k = n$ is called gossiping. We give a gossiping protocol for combined messages that works in the average time $\mathcal{O}(n/\log n)$, which is shown to be optimal. For the general M2A communication problem, we show that it can be performed in the average time $\mathcal{O}(\min\{k \log(n/k), n/\log n\})$ with combined messages, and that $\Omega(k/\log n + \log n)$ is a lower bound. We give a gossiping protocol for separate messages that works in the average time $\mathcal{O}(n \log n)$, which is shown to be optimal. For the general M2A communication problem, we develop a protocol for separate messages with the average time $\mathcal{O}(k \log(n/k) \log n)$, and show that $\Omega(k \log n)$ is a lower bound.

1 Introduction

Packet radio networks are a class of wireless networks in which only one wave frequency is used for communication. The restricted bandwidth results in a conflict when different messages arrive simultaneously at a node. The main challenge, in developing communication protocols for such networks, is in resolving local conflicts for access to the limited bandwidth.

The networks we consider are directed, which captures a scenario in which a possibility of a direct transmission from node x to node y does not necessarily make it possible for node y to transmit directly to node x . Networks are ad-hoc, in that protocols do not rely on the knowledge of the topology; the only information about the network that may be a part of code of a protocol is the size n , which is the number of nodes. We consider deterministic distributed communication protocols. Initially, some k among the nodes are simultaneously activated with

* The work of this author is supported by the NSF Grant 0310503.

input data; these data are called rumors. The communication task is to make all the nodes in the network learn all the input rumors. This communication task can be called many-to-all communication (M2A). The special case in which $k = n$ is called gossiping. The underlying network is assumed to be strongly connected, so that gossiping is always possible to achieve.

Nodes exchange packets carrying rumors. A node sends at most one packet per round. In the model of *combined messages*, a packet can carry all the rumors. In such a setting, it is natural to have a node send all the rumors learned so far in any transmitted packet. In the model of *separate messages*, a packet can carry only one rumor. With such a restriction, a protocol needs to rely on a mechanism to prioritize rumors so that a node sends a rumor of the highest current priority at a round.

The time of an execution of a protocol is defined to be the first round when the communication goal has been achieved. Such a completion of the communication task is not required to be known by the nodes. The complexity measure we investigate is the average time as a function of the size n of the network. To find the average time for size n , first compute conceptually the durations of executions of the protocol on all strongly connected networks of size n , and next take the average of the times accrued for these networks.

Our results. We give upper and lower bounds on the average-case complexity of gossiping and many-to-all communication. Protocols are distributed and designed for both the models of combined and separate messages. Let n denote the size of a network and k the number of nodes activated with rumors. The summary of the contributions is as follows.

- I. Gossiping with combined messages can be performed in the average time $\mathcal{O}(n/\log n)$, which is shown to be optimal.
- II. We show that M2A communication can be performed in the average time $\mathcal{O}(\min\{k \log(n/k), n/\log n\})$ with combined messages and that $\Omega(k/\log n + \log n)$ is a lower bound.
- III. Gossiping with separate messages can be performed in the average time $\mathcal{O}(n \log n)$, which is shown to be optimal.
- IV. M2A communication can be achieved in the average time $\mathcal{O}(k \log(n/k) \log n)$ with separate messages. We show a lower bound $\Omega(k \log n)$.

Previous work. Broadcasting in radio networks with topology modeled by random graphs was considered by Elsässer and Gašieniec [12], who showed that $\mathcal{O}(\log n)$ expected time was optimum for distributed protocols. This result can be interpreted as giving the optimum average-time complexity of broadcasting. The authors of this paper do not know of any other results related to the average-time complexity of communication in radio networks.

A many-to-many communication problem in radio networks, similar to what we consider, was studied by Gašieniec, Kranakis, Pelc and Xin [13]. The problem is defined as follows: There is a set S of k nodes initialized with rumors and every one among these nodes needs to get to know all the rumors. Networks are

undirected, each node knows the topology of the network but does not know the set S , the maximum distance d among any pair of nodes in S is an additional parameter. A protocol solving this problem of time complexity $\mathcal{O}(d \log^2 n + k \log^3 n)$ was given in [13].

Related work. The model of multi-hop radio networks was introduced by Chlamtac and Kutten [4] who considered sequential algorithms to find an efficient broadcast protocol for a given input network. The first distributed randomized broadcast protocols of sub-quadratic expected-time performance were given by Bar-Yehuda, Goldreich and Itai [2]. The first distributed deterministic explicit broadcast protocol with sub-quadratic time performance was given by Chlebus, Gašieniec, Gibbons, Pelc and Rytter [5]. Alon, Bar-Noy, Linial and Peleg [1] showed that there exists a bipartite graph of n nodes for which any broadcasting protocol requires time $\Omega(\log^2 n)$. The fastest known deterministic distributed broadcasting protocol was given by Czumaj and Rytter [10], who showed that it works in time $\mathcal{O}(n \log^2 D)$, where D is the diameter of the network.

Gossiping was initially studied for the model of combined messages. The first distributed protocol of sub-quadratic time complexity was given by Chrobak, Gašieniec and Rytter [8]. The fastest known distributed deterministic protocol works in time $\mathcal{O}(n^{4/3} \log^4 n)$; it was given by Gašieniec, Radzik and Xin [14]. The best randomized protocol operates in the expected time $\mathcal{O}(n \log^2 n)$, it was given by Czumaj and Rytter [10].

Oblivious gossiping was first studied by Chlebus, Gašieniec, Lingas, and Pagourtzis [6]. The paper gave a deterministic gossiping protocol that works in time $\mathcal{O}(n^{3/2})$ on undirected networks; this was shown to be optimal by Kowalski and Pelc [15]. Randomized oblivious gossiping protocols working in the expected time $\mathcal{O}(n \log^2 n)$ on undirected networks and $\mathcal{O}(\min\{m, D\Delta\} \log^2 n)$ on directed networks, where Δ is the maximum node in-degree, were presented in [6].

The model of separate messages was first considered by Bar-Yehuda, Israeli and Itai [3] and Clementi, Monti and Silvestri [9]. Christersson, Gašieniec, and Lingas [7] considered gossiping in undirected networks; they gave an adaptive deterministic gossiping protocol with time complexity $\mathcal{O}(n^{3/2} \log n)$ and a randomized protocol of the expected time complexity $\mathcal{O}(n \log^2 n)$.

2 Technical preliminaries

A radio network is modeled as a graph $G = (V, E)$, in which the set of vertices V represents the physical nodes of the network, and the set of edges E represents the possibilities of direct transmissions among the nodes. If node x of a radio network can send a message directly to y , then node y is *reachable* from x . For any ordered pair $\langle u, v \rangle$ of nodes in the network, edge $u \rightarrow v$ is in the graph G if and only if node v is reachable from node u . The *size* of the network is defined to be the number of nodes $|V|$, which we usually denote by n .

We assume full synchrony in that all nodes are equipped with local clocks that are clicking at the same rate and indicate the same round numbers. Protocols we consider are for a scenario when all the nodes activated with inputs start

simultaneously at round zero. When some two nodes v and v' transmit simultaneously at a given round, and are both in-neighbors of node x in the reachability graph of the network, then a *conflict* occurs at x . A conflict at x results in all the messages arriving at x interfering with one another so that each is received as garbled. A message is said to be *heard* when it is received as fully readable in its correct form. Radio networks have the following properties:

- (a) If a node performs a transmission, then it transmits a single message.
- (b) The message transmitted by a node is delivered in that round to all the reachable nodes.
- (c) A node can hear a message delivered at a round, if exactly one among its in-neighbors transmitted in this round.

Communication problems. Initially some nodes hold their input data called *rumors*. When node i is initialized with a rumor, then this rumor is denoted by r_i . The goal of communication protocols is to disseminate such input rumors.

In the problem of *gossiping*, each node v is a source for its private input rumor r_v , and the goal is to have all the nodes learn all the rumors. Gossiping may be called all-to-all communication problem.

A generalization of gossiping called *many-to-all* problem, or simply *M2A*, is about a scenario in which only some of the nodes have input rumors. Such nodes are called *activated*. The goal is to have all the nodes in the network get to know all the rumors of the activated nodes. All nodes in the network participate in forwarding messages in the course of an execution of an M2A protocol.

To have a communication problem in radio networks meaningful, we need to assume that the topology of the underlying graph makes the communication task at hand possible to perform. In the case of gossiping and M2A, the graph is assumed to be strongly connected.

Communication protocols. Correctness of an M2A or gossiping protocol means that the communication goal is eventually achieved on any strongly connected network. Nodes running a protocol are not required to reach eventually a state representing the completion of a task. This is assumed in order to decouple termination from complexity considerations. The *time complexity* of a protocol at hand, for a given strongly connected network, is defined to be the first round when the communication goal has been achieved.

Nodes of a network of size n are identified by their unique names. We assume that *names* give a one-to-one correspondence between the nodes and integers in the range $[0, n - 1]$. While designing communication protocols, we assume that the size of the network is known.

A simple protocol called ROUND-ROBIN operates as follows. In round i , the unique node with name k such that $i \equiv k \pmod{n}$ is scheduled to perform a transmission. There are variants of this protocol depending on the size of packets. In the model of combined messages, a node scheduled to transmit at the current round transmits a message with all the rumors it has learned so far. In the model of combined messages, the protocol is augmented by a selection rule to choose a rumor to transmit from among those that have been learned by the given round. Usually the selection is made by resorting to a queuing mechanism.

```

for  $k := 0$  to  $\ell_n$  do
  if  $v$  is in  $\mathcal{G}_k$  then transmit
call ROUND-ROBIN

```

Fig. 1. Protocol GOSSIP-COMBINED-MESSAGES; the code for node v .

Average complexity. We consider the average time complexity of gossiping and M2A communication on strongly-connected networks. This is the same as the expected time complexity when the probabilistic space has all strongly connected networks on n nodes as elementary events, each occurring with the same probability. A random directed network is strongly connected with the probability exponentially close to 1. This fact allows to obtain expected time estimates while working with arbitrary random directed networks. These estimates are the same when conditioned on the networks being strongly connected, provided the time estimates are polynomial. An explicit termination in polynomial time could be obtained for all protocols we develop, since there are polynomial-time worst-case time estimates for these protocols, valid for strongly connected networks.

We do not want M2A protocols to have their performance biased towards specific sets of activated nodes. Therefore we work with the average complexity of M2A protocols defined in an adversarial manner as follows. Suppose there is an adversary who is given a protocol \mathcal{P} for n nodes together with a number $k \leq n$. The adversary chooses a set K of k specific names of nodes to be activated; the goal of the adversary is to show a scenario maximizing the complexity of the protocol. The average complexity of the protocol \mathcal{P} , for n -node networks, is defined to be the average complexity of protocol \mathcal{P} measured when exactly the nodes in K are activated with rumors.

3 Gossiping with combined messages

We show that gossiping can be performed with the average time $cn/\lg n$, for any fixed $c > 1$, and that the average time always has to be at least $cn/\lg n$, for any fixed $c < 1/2$. (The logarithm of x to the base 2 is denoted by $\lg x$.)

Gossiping protocol for combined messages. Let $\ell_n = \lceil n/b \lg n \rceil$, where $b = \frac{1}{2}(1 + \frac{1}{c})$. Observe that the inequalities $1 > b > 1/c$ hold. Define group \mathcal{G}_k , for $0 \leq k \leq \ell_n$, to consist of nodes i , for $0 \leq i < n$, with the property that the congruence $i \equiv k \pmod{\ell_n}$ holds. The size of a group is about $b \lg n$. The sizes of two groups differ by at most 1. We consider an oblivious protocol GOSSIP-COMBINED-MESSAGES, which is given in Figure 1. A transmission by a node contains all the rumors that the node has already learnt in the execution.

Theorem 1. *For any $c > 1$, the average number of rounds to complete gossiping by protocol GOSSIP-COMBINED-MESSAGES on a network of n nodes is smaller than $cn/\lg n$, for a sufficiently large n .*

Proof. Take a node y and group \mathcal{G}_k . Let x be in \mathcal{G}_k . The node y can hear x at the k th round of the first phase when the following two events hold:

- (i) x is an in-neighbor of y ; and
- (ii) no other node in \mathcal{G}_k is an in-neighbor of y .

It follows that the probability of the event that y hears $x \in \mathcal{G}_k$ during the k th round of the first phase of the protocol is $2^{-|\mathcal{G}_k|} = 2^{-b \lg n} = n^{-b}$. Let x and v be two nodes. Node y is called a *relay* for the pair $\langle x, v \rangle$ when the following holds:

- (i) y is an in-neighbor of v ; and
- (ii) y heard x in the first phase.

Observe that

$$\Pr[y \text{ is a relay for } \langle x, v \rangle] = \frac{1}{2n^b},$$

because the events “ y heard x in the first phase” and “ y is an in-neighbor of v ” are independent.

Consider the first t nodes, that is, the nodes i with $0 \leq i < t$. These nodes are scheduled to perform a transmission among the first t rounds of the protocol ROUND-ROBIN in the second phase. Node i may make node v learn the rumor r_x of x if i is a relay for the pair $\langle x, v \rangle$. For all such nodes i making the first t transmission during ROUND-ROBIN and different from x and v , the events “ i is a relay for the pair $\langle x, v \rangle$ ” are independent.

If v has not learnt r_x in the first t rounds of the second phase, then no i such that $0 \leq i < t$ is a relay for the pair $\langle x, v \rangle$. The latter event holds with the probability $(1 - \frac{1}{2n^b})^t$ by independence of the events of being a relay node. It follows that v does not learn r_x in the first t rounds of ROUND-ROBIN with the probability of at most $(1 - \frac{1}{2n^b})^t$.

We use the inequality

$$\left(1 - \frac{1}{s}\right)^s < \exp\left(-1 + \frac{1}{2s}\right), \quad (1)$$

which holds for real $s > 1$. It yields the following estimate:

$$\left(1 - \frac{1}{2n^b}\right)^t < \exp\left(\left(-1 + \frac{1}{4n^b}\right)\frac{t}{2n^b}\right) = \exp\left(-\frac{t}{2n^b}\right) \exp\left(\frac{t}{8n^{2b}}\right). \quad (2)$$

Let $d = \min\{2b, 1\}$. Take $t = n^a$ where $b < a < d$. Now the right-hand side of (2) becomes

$$\exp(-n^{a-b}/2) \exp(n^{a-d}/8) = \exp(-n^{a-b}/2)(1 + o(1)). \quad (3)$$

Consider the event that for any pair of nodes $\langle x, v \rangle$ there is a relay node during the first t rounds of the second phase. The event does not hold with the probability of at most $n^2 \exp(-n^{a-b}/2)(1 + o(1))$ by the estimate (3). If this event holds, then gossiping is completed by round $\frac{n}{b \lg n} + n^a$, which is smaller than $\frac{cn}{\lg n}$ for a sufficiently large n . Otherwise the time of gossiping can be estimated by $\frac{n}{b \lg n} + n^2$. These two estimates contribute to the expected value of the time of protocol \mathcal{A} to complete gossiping, which together is smaller than $\frac{cn}{\lg n}$, for all sufficiently large n .

Lower bound for gossiping with combined messages. We show that any gossiping protocol for the model of combined messages has the average time complexity $\Omega(n/\log n)$. This implies that protocol GOSSIP-COMBINED-MESSAGES is asymptotically optimal.

Theorem 2. *For any $c < 1/2$ and gossiping protocol \mathcal{A} for the model of combined messages, the average number of rounds to complete gossiping by \mathcal{A} on a network of n nodes is larger than $cn/\lg n$, for a sufficiently large n .*

Proof. Let X be the random variable defined on the domain of all directed graphs of n nodes. For such a graph G , run \mathcal{A} on G and let s be the first round when the gossiping has been completed. Define $X(G) = s$. Let an execution of \mathcal{A} be given as a sequence $\langle T_0, T_1, T_2, \dots \rangle$ of transmissions.

We estimate the probability of the event $X > s$, for integer $s > 0$. Take event $H(v, s)$, for node v and round s , which holds when no node has heard from node v by round s . Observe that

$$\Pr[X > s] \geq \Pr[H(v, s)] . \quad (4)$$

We want to estimate the probability that $H(v, s)$ holds.

We start with choosing v . If some node v does not belong to any of the first s transmissions, then such v yields the best possible estimate $\Pr[H(v, s)] = 1$, which also implies that $\mathbb{E}X > s$.

Assume that every node belongs to at least one among the first s transmissions of protocol \mathcal{A} . Next we restrict our attention only to these transmissions. We claim that there is a node, say, v with the property that every transmission T_i that v belongs to, for $i \leq s$, is of a size at least $|T_i| \geq n/s$. This is because otherwise, even if every node belonged to only one transmission, the total number of nodes in the initial segment of s transmissions of \mathcal{A} were smaller than n , which would contradict the assumption that these transmissions include all the nodes.

A node x hears from v at round $i \leq s$, provided $v \in T_i$, when the following two events hold:

- (i) v is an in-neighbor of x in T_i , and
- (ii) no other node $y \neq x$ in T_i is an in-neighbor of x .

This implies that the estimate

$$\Pr[x \text{ hears from } v \text{ at round } i \mid v \in T_i] \leq 2^{-n/s}$$

holds. Node v could belong to a number of transmissions T_i for $i \leq s$, so we use the estimate

$$\Pr[x \text{ hears from } v \text{ in the first } s \text{ rounds}] \leq s2^{-n/s} .$$

Node x was arbitrary, and we need to be concerned with all the nodes. We use the estimate

$$\Pr[\text{some node hears from } v \text{ in the first } s \text{ rounds}] \leq ns2^{-n/s} .$$

The event $H(v, s)$, that no node hears v during the first s transmissions, holds with a probability of at most

$$\Pr [H(v, s)] \geq 1 - ns2^{-n/s} . \quad (5)$$

To estimate the expected value $\mathbb{E}X$ of X , we use the formula

$$\mathbb{E}X = \sum_{k=0}^{\infty} \Pr [X > k] ,$$

which holds true for any random variable X with non-negative integer values. Combining this with the estimates (4) and (5), we obtain the inequality

$$\mathbb{E}X \geq \sum_{k=1}^t (1 - nk \cdot 2^{-n/k}) , \quad (6)$$

for any integer $t > 0$.

Let k_0 be the largest value of k for which the expression $1 - nk \cdot 2^{-n/k}$ is positive. We take the upper bound t on the range of summation in (6) to be close to k_0 .

Next we estimate the magnitude of k_0 as a function of n . Observe that k_0 is the largest k for which the inequality

$$nk \leq 2^{n/k} \quad (7)$$

holds. Take the binary logarithm \lg of both sides of (7) to obtain the equivalent inequality $\lg n + \lg k \leq \frac{n}{k}$, which implies $k_0 = \frac{n}{2 \lg n} (1 + o(1))$. We use the value $t = n/(2 \lg n)$ in the estimate (6) to obtain

$$\mathbb{E}X \geq \sum_{k=1}^{n/(2 \lg n)} (1 - nk \cdot 2^{-n/k}) = \frac{n}{2 \lg n} - n \sum_{k=1}^{n/(2 \lg n)} k \cdot 2^{-n/k} . \quad (8)$$

The function $f(k) = k2^{-n/k}$ is increasing as $k \rightarrow \infty$. The value $f(t) = f(n/(2 \lg n))$ is the largest term in the sum on the right-hand side of (8). Observe that

$$f(t) = \frac{n}{2 \lg n} \cdot 2^{-2 \lg n} = \frac{n}{2 \lg n} \cdot n^{-2} = \frac{1}{2n \lg n}$$

and hence the estimate

$$\sum_{k=1}^{n/(2 \lg n)} k \cdot 2^{-n/k} \leq \frac{n}{2 \lg n} \cdot \frac{1}{2n \lg n} = \frac{1}{4 \lg^2 n}$$

holds. Therefore (8) can be bounded from below as follows:

$$\mathbb{E}X \geq \frac{n}{2 \lg n} - \frac{n}{4 \lg^2 n} = \frac{n}{2 \lg n} \left(1 - \frac{1}{2 \lg n}\right) ,$$

which completes the proof of Theorem 2.

```

for  $i := \lg k$  downto 0 do
  call SELECTOR-SUBROUTINE( $2^i$ )
  continue ROUND-ROBIN for  $10 \lg n$  rounds

```

Fig. 2. Procedure M2A-COMBINED(k).

4 M2A communication with combined messages

Suppose k nodes among n in the network are activated with rumors. We give a protocol with average time complexity $\mathcal{O}(\min\{k \log(n/k), n/\log n\})$. We assume that k is a power of 2.

Two schedules of transmissions \mathcal{P}_1 and \mathcal{P}_2 are said to be *interleaved*, when the consecutive actions as specified by \mathcal{P}_1 are performed in even-numbered rounds, while \mathcal{P}_2 determines the actions for the odd-numbered rounds. Infinite schedules of transmissions are called *protocols*, while finite schedules are called *procedures* in this paper. When a procedure \mathcal{P}_1 is interleaved with a protocol \mathcal{P}_2 , then eventually \mathcal{P}_1 ends. At this point we make the protocol \mathcal{P}_2 take over completely, such that its actions are performed in all the following rounds; this is explicitly marked in the pseudocode of our protocols by the instruction `continue \mathcal{P}_2` . Another mode of using a protocol \mathcal{P} specifies that the schedule of \mathcal{P} is repeatedly executed for an interval of x rounds, then it is frozen. This is indicated in the pseudocode by the instruction `continue \mathcal{P} for x rounds`.

We use families of sets called (n, j) -selectors in [8]. They are defined as follows. A set Y *selects* element v from a set X when $X \cap Y = \{v\}$. A family \mathcal{F} of subsets of $[n] = [0, n - 1]$ is an (n, j) -*selector* when, for any set $X \subseteq [n]$ of size ℓ , at least $|X|/2$ elements in X can be selected by sets in \mathcal{F} . The size of \mathcal{F} is called its *length*. We refer to any used selector \mathcal{F} as a sequence $\mathcal{F} = \langle F_1, F_2, \dots \rangle$ in an arbitrary fixed order.

Selectors are used to determine schedules of transmissions. Given positive integer number ℓ and a (n, ℓ) -selector \mathcal{F} , we define SELECTOR-SUBROUTINE(ℓ) as follows. Node v transmits in round i if $v \in F_i$; rounds are counted from the call of this subroutine. We use $(n, 2^i)$ -selectors of length $\Theta(2^i \log(n/2^i))$, which were proved to exist in [11].

A M2A procedure, representing the case when k may be a part of code, is given in Figure 2. Protocol M2A-COMBINED-MESSAGES is given in Figure 3. Next we analyze the average complexity of the protocol.

A node v is said to be a *unique transmitter* at a round, when v is the only node transmitting at that round. We say that *broadcast of r_v was successful/completed*, or that *node v broadcast successfully*, when every node has received r_v .

Lemma 1. *Suppose that node v transmits its rumor r_v as the unique transmitter, and after this ROUND-ROBIN is executed. Then the broadcast of r_v is completed in at most $10 \lg n$ following rounds of ROUND-ROBIN with the probability of at least $1 - 1/n^3$.*

```

for  $j := 0$  to  $\lg n$  do
  call M2A-COMBINED( $2^j$ ) interleaved with GOSSIP-COMBINED-MESSAGES
continue GOSSIP-COMBINED-MESSAGES

```

Fig. 3. Protocol M2A-COMBINED-MESSAGES.

Theorem 3. *Protocol M2A-COMBINED-MESSAGES, on networks of n nodes with any k activated nodes, works in average time $\mathcal{O}(\min\{k \log(n/k), n/\log n\})$.*

Proof. First we show that the protocol completes M2A by the end of the loop for $j = \lceil \lg k \rceil$ with the probability of at least $1 - 4k/n^3$. The protocol runs M2A-COMBINED(2^j) which involves SELECTOR-SUBROUTINE(2^j). Since there are $k \leq 2^j$ activated nodes, during SELECTOR-SUBROUTINE(2^j) at most 2^{j-1} activated nodes did not transmit as unique transmitters. Being a unique transmitter results in a successful broadcast during the next ROUND-ROBIN part, with probability at least $1 - 1/n^3$ by Lemma 1.

During SELECTOR-SUBROUTINE(2^{j-1}), at most 2^{j-2} activated nodes did not transmit as unique transmitters, since there are at most 2^{j-1} participating nodes with probability at least $1 - 2^j/n^3$. Those which transmitted as unique transmitters have a successful broadcast during the next ROUND-ROBIN rounds with probability at least $1 - 2^j/n^3 - 2^{j-1}/n^3$.

In general, in an execution of SELECTOR-SUBROUTINE(2^i) within M2A-COMBINED(2^j), there are at most 2^i activated nodes for which broadcast was not successful during previous iterations with probability at least $1 - \sum_{a=i+1}^j 2^a/n^3$. Conditioned on this event, during SELECTOR-SUBROUTINE(2^i) at most 2^{i-1} of activated nodes did not transmit as unique transmitters. It follows that after the i th iteration of the loop, at most 2^{i-1} rumors have not been broadcast successfully with probability at least $1 - \sum_{a=i}^j 2^a/n^3$.

Considering only M2A-COMBINED(2^j), it completes M2A for k activated nodes in time $\sum_{i=0}^j \mathcal{O}(2^i \log(n/2^i) + \log n) \leq \mathcal{O}(k \log(n/k))$ with probability at least $1 - \sum_{a=0}^j 2^a/n^3 \geq 1 - 4k/n^3$. Including also previous executions of M2A-COMBINED($2^{j'}$) for $j' < j$ produces time estimate $\sum_{j' \leq j} \mathcal{O}(2^{j'} \log(n/2^{j'})) = \mathcal{O}(k \log(n/k))$.

Since $\mathcal{O}(n^2)$ is the worst-case time bound, the average time of M2A-COMBINED-MESSAGES is $\mathcal{O}(k \log(n/k)) + \mathcal{O}(n^2) \cdot 4k/n^3 = \mathcal{O}(k \log(n/k))$.

Theorem 4. *The average cost of any M2A protocol, for the model of combined messages, executed on network of n nodes with some k of them activated is $\Omega(k/\log n + \log n)$.*

Proof. Let \mathcal{A} be a M2A protocol. Fix a set K of activated nodes, where $|K| = k$. Let $\langle T_0, T_1, \dots \rangle$ be the sequence in which T_i denotes the set of nodes transmitting at round i in the execution of \mathcal{A} . There are two kinds of rounds i :

Case 1: Rounds i in which T_i includes at most $4 \lg n$ nodes in K that transmit for the first time.

Even $k/(4 \lg n)$ such rounds are not sufficient to exhaust all the elements in K .

Case 2: Rounds i in which there are more than $4 \lg n$ nodes from K transmitting for the first time.

We show that with a large probability in any round, up to round $s = k/(4 \lg n)$, there is no successful transmission between any pair of nodes. Take node v . Let $a = |T_i| > 4 \lg n$. The probability that v receives a rumor for a node in T_i at round i is $(a/2)(1/2)^{a-1} > 1/n^3$, for sufficiently large n . It follows that the probability of existence of a node that receives a rumor at round i is smaller than $1/n^2$. The probability that some node receives a rumor by round s is smaller than s/n^2 . The expected value of the number of rounds by completion of the communication task is at least $s \cdot (1 - s/n^2) > s/2$ for $n > 2$.

The complexity of our protocol is close to the lower bound by a factor of $\mathcal{O}(\log n \log(n/k))$.

5 Gossiping with separate messages

We consider now gossiping in the case when input rumors are so large that it takes a separate packet to carry one rumor. We show that gossiping can be performed with the average time $\mathcal{O}(n \log n)$, and that the average time has to be $\Omega(n \log n)$.

Gossiping protocol for separate messages. Every node v maintains a priority queue Queue_v in the private memory. The queue is used to store rumors that v still needs to transmit. There is a set Received_v to store all the rumors learned so far. A newly received message with a rumor that is not stored in Received_v is added to both Received_v and Queue_v . The protocol working according to these rules is called GOSSIP-SEPARATE-MESSAGES; it is given in Figure 4.

Let the nodes be ordered cyclically by their names in $[n] = [0, n-1]$, so that i is followed by number $(i+1) \bmod n$. This ordering governs which nodes transmit in any ROUND-ROBIN type of protocol, like GOSSIP-SEPARATE-MESSAGES in particular.

The priority queue Queue_v has its own queuing discipline. Rumors are ordered cyclically, starting from the own input rumor r_v . This rumor is followed by rumors with larger indices according to their order, that is, r_{v+1} , r_{v+2} , until r_{n-1} , which is then followed by r_0 , r_1 , through the final r_{v-1} .

Theorem 5. *The average number of rounds to complete gossiping by protocol GOSSIP-SEPARATE-MESSAGES on a network of n nodes is $\mathcal{O}(n \log n)$.*

```

initialize Receivedv := Queuev := {rv};
for round i := 0 to ∞ do
  if v ≡ i (mod n) then
    if Queuev nonempty then
      transmit the first rumor r in Queuev
      and remove r from Queuev
    else
      attempt to receive a message;
      if rumor r received then
        if r is not in Receivedv then
          insert r into Queuev and add to Receivedv

```

Fig. 4. Protocol GOSSIP-SEPARATE-MESSAGES; the code for node v .

Proof. Every node transmits every rumor exactly once. The worst-case time complexity of this gossiping protocol is n^2 . The full cycle of n rounds makes an *epoch*. During the first epoch, every node v transmits its input rumor r_v .

Take some rumor r and consider an event $\mathcal{E}_a(r)$ which holds when r has been transmitted by $a \lg n$ different nodes. The probability of the event that some node y has not heard r , conditioned on $\mathcal{E}_a(r)$, is n^{-a} . The probability that some node has not heard r , conditioned on $\mathcal{E}_a(r)$, is at most $n \cdot n^{-a} = n^{1-a}$. The probability that some node has not heard some rumor, conditioned on the events $\mathcal{E}_a(r)$ for all rumors r , is at most n^{2-a} . In the following application we will use $a = 4$ to obtain the probability $n^{2-4} = n^{-2}$.

Consider the following event \mathcal{B} : every rumor was transmitted at least $3 \lg n$ times during the first $b \lg n$ epochs, for some fixed integer b to be determined later. Take a node v and the $b \lg n$ nodes preceding v in the cyclic ordering. If any of these nodes receives rumor r_v in the first epoch from v , then it transmits r_v in the first $b \lg n$ epochs. When v transmits in the first epoch, then every other node receives r_v with probability $1/2$ independently over all the nodes. The expected value of the number of these nodes that receive r_v in the first epoch is $\mu = \frac{b}{2} \lg n$. Take δ determined by the equality $(1 - \delta) \frac{b}{2} \lg n = 3 \lg n$, that is, $\delta = 1 - \frac{6}{b}$. Then by the Chernoff bound, the probability that less than $3 \lg n$ nodes receives rumor r_v in the first epoch is at most

$$\exp\left\{-\left(1 - \frac{6}{b}\right)^2 \frac{b}{4} \lg n\right\} \leq n^{-(1 - \frac{6}{b})^2 \frac{b}{4} \lg e}.$$

Take integer $b > 6$ for which the inequality $(1 - \frac{6}{b})^2 \frac{b}{4} \lg e \geq 3$ holds. This b is sufficient to guarantee that event \mathcal{B} does not hold with the probability of at most n^{-2} .

Conditional on \mathcal{B} , the expected time of gossiping is at most $bn \lg n + n^{-2} \cdot n^3 = n(1 + b \lg n)$, because the worst-case time complexity is n^3 . Since event \mathcal{B} does not hold with probability at most n^{-2} , the unconditional expected time complexity is at most $n(1 + b \lg n) + n^{-2} \cdot n^3 = n(2 + b \lg n)$, for a similar reason.

```

for  $i := \lg k$  downto 0 do
  for  $j := 1$  to  $m(n, 2^i)$  do
    (a) if  $v \in F_j(n, 2^i)$  then transmit rumor  $r_v$ 
        else attempt to hear a message
          ( this is  $j$ th round of SELECTOR-SUBROUTINE( $2^i$ ) )
    (r) continue ROUND-ROBIN-STACK in next  $10 \lg n$  rounds

```

Fig. 5. Procedure M2A-SEPARATE(k); the code for node v .

The average number of rounds to complete gossiping on a network of n nodes is $\Omega(n \log n)$; this is a corollary of a more general lower bound for M2A communication shown in Section 6.

6 M2A communication with separate messages

We give a protocol with average time $\mathcal{O}(k \log(n/k) \log n)$. Let k be a power of 2.

SELECTOR-SUBROUTINE(2^i) is similar to the one described for the protocol with combined messages, in that it uses $(n, 2^i)$ -selector. There are two main differences in how they are used. The first difference is that after *each* round of SELECTOR-SUBROUTINE(2^i) we continue with ROUND-ROBIN-STACK for $10 \lg n$ rounds, while in the case of combined messages we put $10 \lg n$ of ROUND-ROBIN rounds after every used $(n, 2^i)$ -selector. The second difference is that specific rumor needs to be selected for each transmission by a node.

An M2A procedure, representing the case when k may be a part of code, is given in Figure 5. An auxiliary protocol ROUND-ROBIN-STACK used in procedure M2A-SEPARATE(k) is defined as follows. A node maintains a stack of rumors different from its original one. A rumor heard by the node is pushed on its stack. A rumor to transmit is obtained by popping the stack; when the stack is empty, then the node pauses. The stack is initialized to be empty, and is made empty just before ROUND-ROBIN-STACK is to be continued for $10 \lg n$ rounds, see Figure 5.

Protocol M2A-SEPARATE-MESSAGES is given in Figure 6. Next we analyze the average complexity and optimality of the protocol.

Theorem 6. *Protocol M2A-SEPARATE-MESSAGES, on a network of n nodes with k nodes initially activated, has the average time $\mathcal{O}(k \log(n/k) \log n)$.*

Proof. First, M2A task is completed by the end of M2A-SEPARATE(2^j), where $j = \lceil \lg k \rceil$, with probability at least $1 - 4k/n^3$. Consider M2A-SEPARATE(2^j). It follows that during SELECTOR-SUBROUTINE(2^j) of M2A-SEPARATE(2^j) at most 2^{j-1} activated nodes do not transmit as unique transmitters in rounds (a). Those who transmit as unique transmitters in some rounds (a) have also successful broadcasts in the following $10 \lg n$ rounds of ROUND-ROBIN-STACK in code line (r), with probability at least $1 - 1/n^3$ each, by Lemma 1.

```

for  $j = 0$  to  $\lg n$  do
  call M2A-SEPARATE( $2^j$ )
call GOSSIP-SEPARATE-MESSAGES

```

Fig. 6. Protocol M2A-SEPARATE-MESSAGES.

Consider SELECTOR-SUBROUTINE(2^{j-1}), which is the second subroutine of M2A-SEPARATE(2^j). During this part at most 2^{j-2} activated nodes did not transmit as unique transmitters in rounds (a). Conditioned on this event, certain rumors are completed during $10 \lg n$ following rounds of ROUND-ROBIN-STACK in part (r) of the loop, with probability at least $1 - 1/n^3$ each, again by Lemma 1. We continue analyzing subroutines of M2A-SEPARATE(2^j) which are based on $(n, 2^i)$ -selectors for $i = \lg(k/4), \lg(k/8), \dots, 1, 0$. Quantitatively, by the beginning of SELECTOR-SUBROUTINE(2^i) at most 2^i selected nodes have not broadcasted successfully, with the probability of at least $1 - \sum_{a=i+1}^j 2^a/n^3$. Conditioned on this event, during SELECTOR-SUBROUTINE(2^i) at most 2^{i-1} activated nodes did not transmit as unique transmitters in rounds (a), while those which have transmitted as unique transmitters in rounds (a) complete broadcast during next $10 \lg n$ rounds in line (r) of the code, with the probability of at least $1 - 2^i/n^3$. Consequently, by the beginning of SELECTOR-SUBROUTINE(2^{i-1}) at most 2^{i-1} of activated nodes have not complete broadcast, with probability at least $1 - \sum_{a=i}^j 2^a/n^3$.

M2A-SEPARATE(2^j) takes $\sum_{i=0}^{\lg k} \mathcal{O}(2^i \log(n/2^i) \lg n) = \mathcal{O}(k \log(n/k) \log n)$ rounds, and during this procedure M2A task is completed with probability at least $1 - \sum_{a=0}^j 2^a/n^3 \geq 1 - 4k/n^3$.

The number of rounds in M2A-SEPARATE-MESSAGES by the end of execution of M2A-SEPARATE(2^j) is $\sum_{j'=0}^j \mathcal{O}(2^{j'} \log(n/2^{j'}) \log n) = \mathcal{O}(k \log(n/k) \log n)$. The worst-case $\mathcal{O}(n^3)$ can occur with probability at most $4k/n^3$. This justifies the estimate $\mathcal{O}(k \log(n/k) \log n) + \mathcal{O}(n^3) \cdot 4k/n^3 = \mathcal{O}(k \log(n/k) \log n)$ to be an upper bound on the average time.

We also show a lower bound for M2A communication with separate messages.

Theorem 7. *For any M2A protocol for the model of separate messages, the average number of rounds to complete gossiping on a network of n nodes with k nodes initially activated is $\Omega(k \log n)$.*

Corollary 1. *For any gossiping protocol for the model of separate messages, the average number of rounds to complete gossiping on a network of n nodes is $\Omega(n \log n)$.*

Our M2A protocol is within a factor of at most $\mathcal{O}(\log(n/k))$ close to optimality. In the case of $k = \Omega(n)$, which includes gossiping, the protocol is asymptotically optimal.

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 (1991) 290 - 298.
2. R. Bar-Yehuda, O. Goldreich, and A. Itai, On the time complexity of broadcast in radio networks: An exponential gap between determinism and randomization, *Journal of Computer and System Sciences*, 45 (1992) 104 - 126.
3. R. Bar-Yehuda, A. Israeli, and A. Itai, Multiple communication in multi-hop radio networks, *SIAM Journal on Computing*, 22 (1993) 875 - 887.
4. I. Chlamtac, and S. Kutten, On broadcasting in radio networks - problem analysis and protocol design, *IEEE Transactions on Communication*, 33 (1985) 1240 - 1246.
5. B.S. Chlebus, L. Gašieniec, A.M. Gibbons, A. Pelc, and W. Rytter, Deterministic broadcasting in ad hoc radio networks, *Distributed Computing*, 15 (2002) 27 - 38.
6. B.S. Chlebus, L. Gašieniec, A. Lingas, and A. Pagourtzis, Oblivious gossiping in ad-hoc radio networks, in *Proc., 5th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications (DIALM)*, 2001, pp. 44 - 51.
7. M. Christersson, L. Gašieniec, and A. Lingas, Gossiping with bounded size messages in ad-hoc radio networks, in *Proc., 29th International Colloquium on Automata, Languages and Programming (ICALP)*, 2002, pp. 377 - 389.
8. M. Chrobak, L. Gašieniec, and W. Rytter, Fast broadcasting and gossiping in radio networks, *Journal of Algorithms*, 43 (2002) 177 - 189.
9. A.E.F. Clementi, A. Monti, and R. Silvestri, Distributed broadcasting in radio networks of unknown topology, *Theoretical Computer Science*, 302 (2003) 337 - 364.
10. A. Czumaj, and W. Rytter, Broadcasting algorithms in radio networks with unknown topology, in *Proc., 44th IEEE Symposium on Foundations of Computer Science (FOCS)*, 2003, pp. 492 - 501.
11. A. De Bonis, L. Gašieniec, and U. Vaccaro, Optimal two-stage algorithms for group testing problems, *SIAM Journal on Computing*, 34 (2005) 1253 - 1270.
12. R. Elsässer, and L. Gašieniec, Radio communication in random graphs, in *Proc., 17th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, 2005, pp. 309 - 315.
13. L. Gašieniec, E. Kranakis, A. Pelc, and Q.Xin, Deterministic M2M multicast in radio networks, in *Proc., 31st International Colloquium on Automata, Languages and Programming (ICALP)*, 2004, pp. 670 - 682.
14. L. Gašieniec, T. Radzik, and Q. Xin, Faster deterministic gossiping in directed ad-hoc radio networks, in *Proc., 9th Scandinavian Workshop on Algorithm Theory (SWAT)*, 2004, pp. 397 - 407.
15. D.R. Kowalski, and A. Pelc, Time complexity of radio broadcasting: adaptiveness vs. obliviousness and randomization vs. determinism, *Theoretical Computer Science*, 333 (2005) 355 - 371.