On the Bit Communication Complexity of Randomized Rumor Spreading*

Pierre Fraigniaud
CNRS and Univ. Paris Diderot
Paris, France
pierre.fraigniaud@liafa.jussieu.fr

George Giakkoupis
CNRS and Univ. Paris Diderot
Paris, France
ggiak@liafa.jussieu.fr

ABSTRACT

We study the communication complexity of rumor spreading in the random phone-call model. Suppose n players communicate in parallel rounds, where in each round every player calls a randomly selected communication partner. A player u is allowed to exchange messages during a round only with the player that u called, and with all the players that u received calls from, in that round. In every round, a (possibly empty) set of rumors to be distributed among all players is generated, and each of the rumors is initially placed in a subset of the players. Karp *et. al* [16] showed that no rumor-spreading algorithm that spreads a rumor to all players with constant probability can be both time-optimal, taking $O(\lg n)$ rounds, and message-optimal, using $O(n)$ messages per rumor. For address-oblivious algorithms, in particular, they showed that $\Omega(n \lg\lg n)$ messages per rumor are required, and they described an algorithm that matches this bound and takes $O(\lg n)$ rounds.

We investigate the number of communication bits required for rumor spreading. On the lower-bound side, we establish that any address-oblivious algorithm taking $O(\lg n)$ rounds requires $\Omega(n(b+\lg\lg n))$ communication bits to distribute a rumor of size b bits. On the upper-bound side, we propose an address-oblivious algorithm that takes $O(\lg n)$ rounds and uses $O(n(b+\lg\lg n \lg b))$ bits. These results show that, unlike the case for the message complexity, optimality in terms of both the running time and the bit communication complexity is attainable, except for very small rumor sizes $b \ll \lg\lg n \lg\lg\lg n$.

Categories and Subject Descriptors

F.2.2 [**Analysis Of Algorithms And Problem Complexity**]: Nonnumerical Algorithms and Problems; E.4 [**Data**]: Coding And Information Theory—*Data compaction and compression*

*Research supported in part by the ANR projects ALADDIN and PROSE, and by the INRIA project GANG.

SPAA'10, June 13–15, 2010, Thira, Santorini, Greece.

General Terms

Algorithms, Performance, Theory

Keywords

rumor spreading, random phone call, bit communication complexity

1. INTRODUCTION

We study the problem of information spreading in a distributed environment where information is exchanged using randomized communication. Suppose n players communicate in parallel rounds, where in each round every player *calls* a randomly selected communication partner. Each player u is allowed to exchange messages during a round only with the player that u called, and with all the (zero or more) players that called u, in that round. This communication model is often referred to as the *random phone-call model* [16]. In every round, zero or more pieces of information, called *rumors*, are generated, and each rumor is placed to one or more players, the *sources* of the rumor. The goal is that each rumor be distributed among all players within a small number of rounds from the round that the rumor was generated, and by using a small amount of communication between players.[1]

A motivating example for this problem is the maintenance of replicated databases, for instance, on name servers in a large corporate network [4]. In such a system, updates are injected at various nodes and at various times, and these updates must be propagated to all nodes in the network. It is desirable that all databases converge to the same content quickly, and with little communication overhead. The motivation for using a randomized communication model is that such a scheme is simple, scalable, and naturally fault tolerant [4, 11].

A simple rumor-spreading algorithm for the random phone-call model is the so-called *push* algorithm. A rumor r is spread as follows. In each round, starting from the round in which r is generated, every *informed* player u (i.e., every player who knows r) forwards r to the player v that u calls in that round; we say that u *pushes* r to v. The distribution of

[1]A variant of this problem that is often considered in the literature is when each rumor has exactly one source. As we discuss later, with one source per rumor the number of rumors generated per round is essentially bounded by n, since all rumors generated at the same round by the same source can be grouped into a single large rumor; this trick does not work when a rumor can have more than one sources.

r is terminated after $\Theta(\lg n)$ rounds, at which time all players know r with high probability [13, 17]. The runtime of the push algorithm is asymptotically optimal for the random phone-call model, as we will see later. However, the algorithm suffers from high communication overhead, performing $\Theta(n \lg n)$ transmissions of the rumor. Intuitively, the number of informed players roughly doubles in each round, until a constant fraction of the players is informed; and in each subsequent round, the number of non-informed players halves. Thus, in the last $\Theta(\lg n)$ rounds $\Theta(n)$ players push the rumor in each round.

The *push-pull* rumor-spreading algorithm, proposed in [16], has asymptotically optimal runtime as well, but it has a smaller communication overhead than the push algorithm. A rumor r is distributed as follows. In each round from the round when r is generated, every informed player u pushes r to the player that u calls in this round, as in the push algorithm, and, in addition, u forwards r to every player v that calls u in this round; we say that r is *pulled* from u to v. In the basic version of this algorithm, where a rumor is assumed to have a single source, the distribution of r is terminated after $\lg_3 n + \Theta(\lg \lg n)$ rounds. By that time, with high probability, all players know r, and r has been transmitted $\Theta(n \lg \lg n)$ times. The intuition is that the push and pull transmissions roughly triple the number of informed player in each round until a constant fraction of the players is informed, and, from this point on, the pull transmissions shrink the fraction of non-informed players from s_{t-1} to $s_t = s_{t-1}^2$, in each round t. Thus, only $\Theta(\lg \lg n)$ additional rounds are required after a constant fraction of players is informed. Note that when a rumor may have more than one sources, the message complexity per rumor can be as bad as $\Theta(n \lg n)$—e.g., when the rumor has $\Theta(n)$ sources. A variant of the basic push-pull algorithm, also proposed in [16], uses a more robust termination criterion that detects when a large fraction of players is informed. This algorithm takes $O(\lg n)$ rounds and uses $\Theta(n \lg \lg n)$ messages per rumor, regardless of the number of sources per rumor.

On the lower-bound side, it is known that no decentralized rumor-spreading algorithm for the random phone-call model taking $O(\lg n)$ rounds and using $O(n)$ messages per rumor can guarantee that a rumor is spread to all players with constant probability [16]. In other words, it is not possible to achieve simultaneously optimality both in terms of the running time and the message complexity in the random phone-call model.[2] Moreover, for the case of *address-oblivious* algorithm, such as the push and push-pull algorithms above, $\Omega(n \lg \lg n)$ messages are required, regardless of the number of rounds [16]. So, the push-pull protocol is asymptotically optimal among the address-oblivious algorithm in terms of time and message complexity.

In this paper, we investigate the communication complexity of rumor spreading in the random phone-call model, measured in terms of the number of *bits* exchanged between players. The standard approach to measuring the communication complexity has been in terms of messages, counting one message for every quadruplet $\langle r, t, u, v \rangle$ such that information regarding rumor r is exchanged in round t between players u and v. In the rumor-spreading algorithms that have been proposed each such exchange of information typically involves the actual rumor r, plus the values of some small counters, such as the age of the rumor. Arguably, for some applications the volume of information exchanged is at least as relevant as the number of messages, and trying to minimize the number of bits exchanged, in addition to the number of messages, is desirable. This is especially true when a large number of rumors are spread simultaneously, or when rumors are large.

1.1 Our results

As we saw above, no rumor-spreading algorithm in the random phone-call model can be both time-optimal, taking $O(\lg n)$ rounds, and message-optimal, using $O(n)$ messages per rumor. We show that the situation is different when bit communication complexity is considered in place of message complexity. Specifically, we describe an address-oblivious algorithm that uses $O(\lg n)$ rounds and $O(n(b + \lg \lg n \lg b))$ bits of communication to distribute a b-bit rumor among all players with high probability. Also, $O(n \lg \lg n)$ messages per rumor are exchanged. These guarantees hold even when the rumors are generated by an adversary. On the lower-bound side, we establish that any address-oblivious algorithm taking $O(\lg n)$ rounds requires $\Omega(n(b + \lg \lg n))$ communication bits to spread a b-bits rumor to all players with constant probability. These two results imply that, unlike the case for the message complexity, optimality in terms of both the running time and the bit communication complexity is attainable, except for very small rumor sizes $b \ll \lg \lg n \lg \lg \lg n$.

Discussion

Our rumor-spreading algorithm can be described as a push-pull algorithm with "concise" feedback. Note that the original push-pull algorithms proposed in [16] require $O(nb \lg \lg n)$ communication bits per b-bit rumor.[3] So, our algorithm saves a $\lg \lg n$ factor for large b, and a $b/\lg \lg b$ factor for small b.

Informally, the algorithm works as follows. When a player learns a new rumor r, she pushes r in all subsequent rounds, until the 3rd time she pushes the rumor to some player who already knows it (when a rumor is pushed, the recipient informs the sender whether she knew the rumor). These push transmissions guarantee that a constant fraction of the players is informed within roughly $\lg n$ rounds, and that r is pushed no more than $4n$ times. Pull transmissions take place only every $\lg n / \lg \lg n$ rounds—there are $\Theta(\lg \lg n)$ pull rounds during the lifetime of r. Say u calls v in such a round. Ideally, we would like the set of rumors pulled from v to u to consist of exactly those rumors that v knows and u does not know; and this should be achieved without communicating more than roughly $nb/\lg \lg n$ additional bits per b-bit rumor, per pull round. This is a non trivial task, since players do not know the number or size of the rumors currently circulating; an unbounded number of rumors can be generated in each round, and any b-bit string can be a valid rumor, for any b. Also, the fact that a rumor may have more than one sources

[2] Note that if the players knew the complete communication graph and the set of informed players in each round then $\Theta(\lg n)$ rounds and $n - 1$ messages would be required to distribute a rumor started by a single source. The $\Omega(\lg n)$ time bound follows by a simple reachability argument based on the fact that node degrees in the communication graphs are sharply concentrated around their mean value 2.

[3] More precisely, for the basic version this complexity holds for one source per rumor, and for the other version the exact complexity is $O(n(b + \lg \lg \lg n) \lg \lg n)$ bits.

precludes "grouping" into a big rumor all the rumors started at the same time by the same player, which would effectively bound by n the number of rumors generated per round. For these reasons simple solutions such as the use of fingerprints to uniquely describe a rumor with fewer bits do not work. At the core of our rumor-spreading algorithm is a simple data structure for approximate set membership, used to encode the set of rumors that a player knows using roughly $\lg b$ bits per b-bit rumor. This data structure is deterministic, and allows for some false positives. When u calls v in a pull round, u sends to v this data structure of the recent rumors that u knows; and based on that, v decides which rumors to transmit to u.

For the lower bound, note that an $\Omega(n \lg \lg n)$ bound on the number of bits communicated per rumor is immediate from the same bound of [16] on the number of messages. So, we just have to show an $\Omega(nb)$ bound, which seems like a trivial information-theory result. However, a more careful look reveals that this is not the case: Information may be conveyed not just by the content of the messages exchanged, but also by the round in which they are exchanged. Even sending no messages through an established connection also conveys information. In fact, the $\Omega(nb)$ bound no longer holds if we can have more than $O(\lg n)$ rounds. The following (impractical) protocol spreads a b-bit rumor using only $O(n \lg n \lg b)$ bits, within $O(2^b \lg n)$ rounds. We modify the push algorithm such that for each rumor r, the size b of r is pushed instead of r, and also transmissions take place only in rounds t that are equal to r modulo 2^b (where r is viewed as a binary number). So, within $O(2^b \lg n)$ rounds, every player learns b and, thus, r, which is the last b bits of the round in which the player was informed.

We prove the $\Omega(nb)$ bound in two steps. We first establish the bound for large rumors, using essentially a counting argument. Then we reduce the case of smaller rumors into the previous case, by showing that given an algorithm that spreads small rumors using $o(nb)$ bits, we can devise an algorithm that spreads large rumors using $o(nb)$ bits as well. Error-correcting codes are used in this construction. We note that the $\Omega(nb)$ bound holds also for non address-oblivious algorithms.

1.2 Related work

There is a large literature on deterministic rumor spreading and related information dissemination problems in various communication models. For an overview of this volume of work see [12, 14, 15]. The problem of randomized rumor spreading was introduced in [13], where the runtime of the push algorithm in the random phone-call model was analyzed. This result was later refined in [17]. Randomized rumor spreading in the setting where players correspond to nodes in a graph (other than the complete graph), and in each round a player chooses its communication partner at random *among its graph neighbors*, was first studied in [11]. There, bounds on the runtime of the push algorithm in arbitrary graphs were derived, and the runtime of the same algorithm in the hypercube and in random graphs was analyzed. The runtime and message complexity of randomized rumor spreading in random graphs were also studied in [9, 10], where a push-pull algorithm was analyzed, as well as two variations of it where players can remember their recent connections, or they initiate multiple calls per round. Push-pull algorithms have also been proposed and analyzed for random d-regular graphs [1], and for scale-free graphs [8]. In [5], a quasirandom analogue to the random phone-call model was introduced. In this model, each player has a cyclic list of all the players (or of all its neighbors, in case of rumor spreading in a graph). A player initially calls a player at a random position in her list, but from then on she calls her neighbors in the order of the list. It was shown that the push algorithm in the quasirandom model performs asymptotically at least as well as in the random model, for all the cases of graphs studied in [11], even when the lists are determined by an adversary. Rumor spreading in the quasirandom model was further explored in [6].

2. MODEL

In the random phone-call model [16], n players communicate in parallel rounds, in each of which every player u chooses a player v independently and uniformly at random, and u *calls* v. In a given round, u can only communicate with the player that u called, and with the players that called u, in that round. Communication inside each round is assumed to proceed in parallel, that is, any information received in a round cannot be forwarded to another player in the same round. In each call, communication between the caller and the receiver proceeds sequentially: one player sends a message, then the other sends a message back, and so on. No restrictions are imposed on the type or the size of information exchanged.

In each round, an adversary generates a (possibly empty) set of rumors, and places each rumor r to a non-empty subset of players, the *sources* of r. A rumor is just a binary string, and any binary string of any size represents a possible rumor; so, there are exactly 2^b distinct rumors of size b. No limit is imposed on the number of rumors generated in a round. However, we assume that rumors generated in two different rounds t_1, t_2 with $|t_1 - t_2| = O(\lg n)$ are distinct (this assumption is made to simplify the exposition of our algorithm, and can be relaxed). If player u calls player v and rumor r is transmitted from u to v we say that r is *pushed*, while if r is transmitted from v to u we say that r is *pulled*.

We measure the bit communication complexity of rumor spreading, that is, the total number of bits exchanged between players. Specifically, in our rumor-spreading algorithm, each message exchanged is either related to a single rumor, or to a set of rumors of the same size. In the latter case, to count the bits communicated per rumor we divide the size of the message by the size of the set of rumors. For the lower bound, we assume that a set of b-bit rumors are started by a single source at a round t, and that no other rumors are generated. To count the bits communicated per rumor, we count the total number of bits exchanged between players, from round t until the distribution of rumors finishes, and then divide by the number of rumors.

We focus on the class of *address-oblivious* algorithms, that is, when player u calls player v, u and v do not know the id of each other. Of course, they can communicate their ids, but this exchange of information is also counted in the bit communication complexity.

3. UPPER BOUND AND OUR RUMOR-SPREADING ALGORITHM

We establish the following upper bound on the performance of rumor spreading in the random phone-call model.

THEOREM 3.1. *There is some address-oblivious algorithm guaranteeing that, with high probability, any rumor is distributed to all players within* $O(\lg n)$ *rounds using* $O(nb + n \lg\lg n \lg b)$ *bits of communication, where* b *is the rumor's size.*

In Section 3.1, we present a rumor-spreading algorithm, and in Section 3.2, we prove that this algorithm meets the performance guarantees of Theorem 3.1.

3.1 Algorithm description

In this algorithm, the distribution of a rumor is not affected by rumors of different size. So, in our description we focus only on rumors of the same size b. Also, very small rumors are treated slightly differently, as we explain at the end of this section.

The algorithm is a push-pull algorithm with feedback. Consider a b-bit rumor r starting from a set of players S in round t_{start}. The distribution of r continues until round $t_{end} = t_{start} + 6 \lg n - 1$, after which the rumor is considered *cold*. Whenever r is pushed or pulled, its age, that is, the difference between the current round and t_{start}, is also communicated—$O(\lg\lg n)$ bits are required for that. Suppose that a player u learns r in round t; if $u \in S$ we say that u learns r in round $t = t_{start} - 1$. From the next round $t+1$ on, u *pushes* r in every round until the 3rd time that u has pushed r to a player who already knows the rumor from a previous round, or until the round t_{end} is reached. After that time, u does not push r again. A player can tell if in some round she pushed a rumor to someone who already knew the rumor, because in every push transmission the recipient sends back a (constant-size) feedback containing that information.[4]

Pull transmissions occur only every $\lg n / \lg\lg n$ rounds, on rounds that are multiples of $\lg n / \lg\lg n$; these rounds are called *pull rounds*. (During pull rounds, push transmissions take place normally, as in regular rounds.) Suppose that player u calls player v in pull round t. Player u then sends to v a *digest* of the rumors that u has learned recently, and, based on this digest, v decides which rumors should be pulled to u. Next we describe the details of how the digest is created, and which rumors are pulled.

Digest

The digest of u in round t is built from all the non-cold rumors that u knows at the beginning of round t. Let $R_{b,i}$ be the subset of these rumors consisting of the rumors of size b that were generated during the i-th previous epoch, where an *epoch* is the time interval from the beginning of a pull round until the beginning of the next pull round. The digest of u consists of one component $D_{b,i}$ for each non-empty set $R_{b,i}$. Suppose that $R_{b,i} = \{r_1, \ldots, r_\kappa\}$, for some $i, \kappa \geq 1$. The elements of $R_{b,i}$ are indexed such that if the r_k are interpreted as a binary numbers then $r_1 < \cdots < r_\kappa$. The digest $D_{b,i}$ for $R_{b,i}$ consists of two parts:

1. the list $\langle b, i, p_1, p_2, \ldots, p_{\kappa-1}\rangle$, where p_k, for $1 \leq k < \kappa$, is the position of the leftmost bit (most significant bit) where r_k and r_{k+1} differ; and
2. the subset $\{r_{j\ell} : j = 1, \ldots, \lfloor \kappa/\ell \rfloor\}$ of $R_{b,i}$ containing every ℓ-th rumor, where $\ell = \lg n$.

Note that the size of $D_{b,i}$ is $O(\lg\lg\lg n + \kappa \lg b + \kappa b/\ell)$ bits. If $R_{b,i} = \emptyset$, i.e., if u does not know any non-cold rumor of size b generated in the i-th previous epoch, we will write $D_{b,i} = \emptyset$.

We now explain how from digest $D_{b,i} \neq \emptyset$ we obtain information about whether a given b-bit rumor r is a member of $R_{b,i}$. Let $x[j]$ denote the j-th leftmost bit of bit-string x. From the definition of the p_k and the assumption that $r_k < r_{k+1}$, we have that for $1 \leq j < p_k$, $r_k[j] = r_{k+1}[j]$, and

$$r_k[p_k] = 0 \neq 1 = r_{k+1}[p_k].$$

Based on this observation, we describe a simple algorithm that for any given r, computes an index k with $1 \leq k \leq \kappa$ such that

$$r \notin R_{b,i} \setminus \{r_k\}$$

The algorithm does not tell whether $r = r_k$. We denote this index by $\mathrm{ind}(r, D_{b,i})$, and we compute it as follows. We start with the list $1, \ldots, \kappa$ of all indices, and in each step we eliminate one of the first two indices remaining, until there is only one index left; this last index is $\mathrm{ind}(r, D_{b,i})$. For each index $k < \kappa$ in the current list, we maintain the leftmost bit position at which r_k and $r_{k'}$ differ, where k' is the index following k in the current list; so, for the initial list of indices we have the positions $p_1, p_2, \ldots, p_{\kappa-1}$ described in the first component of $D_{b,i}$. Maintaining these positions does not require knowledge of the actual rumors. Let k_1, k_2 be the first two of the indices remaining at the beginning of a step, and let q_1, q_2 be the bit positions currently associated with them. If $r[q_1] = 1$ then $r \neq r_{k_1}$, and, so, k_1 is removed from the list of indices in this step. The bit positions associated with the indices remaining do not change. If $r[q_1] = 0$, instead, then $r \neq r_{k_2}$, and, so, k_2 is removed from the list. Also the bit position associated with k_1 is updated to the leftmost of the positions q_1 and q_2.

Note that the second component $D_2 = \{r_{j\ell} : 1 \leq j \leq \kappa/\ell\}$ of $D_{b,i}$ is not used in computing $\mathrm{ind}(r, D_{b,i})$. The set D_2 is non-empty only when $|R_{b,i}| \geq \ell$, and it can sometimes be used together with $\mathrm{ind}(r, D_{b,i})$ to infer that $r \notin R_{b,i}$. Let $\mathrm{range}(r, D_{b,i})$ denote the set $\{k_1 + 1, \ldots, k_2\}$, where r_{k_1} is the largest element of D_2 such that $r_{k_1} < r$, or $k_1 = 0$ if no such element exits; and r_{k_2} is the smallest element of D_2 such that $r \leq r_{k_2}$, or $k_2 = \kappa$ if no such element exits. Clearly, if $\mathrm{ind}(r, D_{b,i}) = k \notin \mathrm{range}(r, D_{b,i})$ then $r \neq r_k$, and, thus, $r \notin R_{b,i}$.

Based on u's digest, player v determines which rumors should be pulled from v to u as follows. Let $R'_{b,i}$ be the set of non-cold b-bit rumors that v knows, generated in the i-th previous epoch. If $D_{b,i} = \emptyset$ then all the rumors in $R'_{b,i}$ are pulled. Otherwise, for each $r \in R'_{b,i}$, v computes $\mathrm{ind}(r, D_{b,i})$ and $\mathrm{range}(r, D_{b,i})$, as described above, and r is pulled iff at least one of the following two conditions is satisfied:

(1) $\mathrm{ind}(r, D_{b,i}) \notin \mathrm{range}(r, D_{b,i})$;

(2) $\mathrm{ind}(r, D_{b,i}) = \mathrm{ind}(r', D_{b,i})$, for some $r' \in R'_{b,i} \setminus \{r\}$.

Note that if (1) holds then u does not know r; and if (2) holds then u knows at most one of r and r', thus, the bits transmitted are at most twice the bits necessary. Note, however, that

[4]The idea that a player stops pushing a rumor after a fixed number of unnecessary push transmissions was also suggested in [4]. An alternative stopping criterion that would also work is that a player stops pushing a rumor after a constant number of push transmissions (regardless of their outcome). The analysis for this approach is similar to that of Theorem 4.1.3 (Stage A) in [11].

the following bad scenario is possible: u does not know r, but $\mathrm{ind}(r, D_{b,i}) \in \mathrm{range}(r, D_{b,i})$ and $\mathrm{ind}(r', D_{b,i}) \neq \mathrm{ind}(r, D_{b,i})$, for all $r' \in R'_{b,i} \setminus \{r\}$; thus, r is not pulled. Nevertheless, we show that condition (2) ensures that the desired performance guarantees for the distribution of r are still met.

The digest structure employed by our algorithm can be viewed as essentially a data structure for approximate set membership. This problem is traditionally addressed using Bloom filters [2] (see also the survey [3]). Similarly to Bloom filters, our approach allows for false positives, but, unlike them, it is deterministic; and in addition to the information whether an element is a member of the set, it also gives the order of the element in that set. This feature is exploited by our algorithm to tackle the problem of false positives.

The case of very small rumors

In the algorithm above, if $b = (\lg\lg n)^{o(1)}$ then all but a $o(1)$ fraction of the bits used to distribute a single b-bit rumor are used to transmit the age information contained in the digest. We handle this issue by making the following two changes to the algorithm. For every rumor r of size $b = (\lg\lg n)^{o(1)}$, the beginning of the distribution of r is delayed until the next round that is a multiple of $6\lg n$; i.e., if r is generated in round t, its distribution starts in round $t_{start} = \lceil t/6\lg n\rceil \cdot 6\lg n$. (Recall that $t_{end} = t_{start} + 6\lg n - 1$.) Because of this modification, the epoch information for these rumors contained in the digest is no longer useful and is omitted. Apart from these two changes, the protocol remains the same. Note that these changes could also be applied to the other rumor sizes, but the resulting delays and bursty traffic may be undesirable; thus, we use these modifications only for very small rumors.

3.2 Analysis of algorithm

For the analysis, we distinguish two phases in the distribution of a rumor. Roughly speaking, in the first phase the rumor is *pushed* to at least a $\frac{1}{2}+\epsilon$ fraction of the players, and in the second phase the rumor is *pulled* to the remaining players. Below, we bound the duration of each phase, and then we bound the total number of bits communicated in the two phases. We only consider the case of rumor sizes $b = (\lg\lg n)^{\Omega(1)}$; for smaller rumors, the analysis is essentially the same.

3.2.1 Phase I: Pushing the rumor to a $\frac{1}{2}+\epsilon$ fraction of the players

For the analysis of this phase, we focus on a single rumor r of size b. To simplify notation we assume that r is generated in round $t_{start} = 1$. We prove the following lemma.

LEMMA 3.1. *With probability $1 - n^{-3+o(1)}$, at least a 3/4 fraction of the players knows r at the end of round $\tau = \lg n + 3\lg\lg n$.*

We start by introducing some notation. S_t denotes the number of players who know r at the end of round t. A push transmission is called *bad* if the recipient already knows the rumor from a previous round. The number of bad push transmissions of r during the first t rounds is denoted B_t.

CLAIM 3.2. *Let $\tau_1 = \inf\{t : S_t \geq (\lg n)^4\}$. With probability $1 - n^{-3+o(1)}$, $B_{\tau_1} \leq 2$ and $\tau_1 \leq 4\lg\lg n + O(1)$.*

PROOF. Fix some ordering of the set of players, and call a push transmission of r from u to v *good* if it is not bad (i.e., v has not learn r in a previous round), and no player $u' < u$ pushes r to v in this round. The number of good push transmissions of r in the first τ_1 rounds is at most $2(\lg n)^4 - 3$ (at most $(\lg n)^4 - 2$ in the first $\tau_1 - 1$ rounds, and at most $(\lg n)^4 - 1$ in round τ_1). Also, the probability that a given push transmission of r in some of the first τ_1 rounds is good is at least $1 - \frac{2(\lg n)^4-3}{n}$, regardless of the other transmissions in the same round. This is because if we ignore the outcome of this transmission and of the pull transmissions in this round (if it is a pull round) then at most $2(\lg n)^4 - 3$ players know r at the end of the round. So, the probability that 3 or more of the push transmissions of r in the first τ_1 rounds are not good is at most

$$\begin{aligned}\binom{2(\lg n)^4}{3}\left(\frac{2(\lg n)^4-3}{n}\right)^3 &\leq \frac{(2(\lg n)^4)^3}{3!}\left(\frac{2(\lg n)^4-1}{n}\right)^3\\ &= n^{-3+o(1)},\end{aligned}$$

From this, it is immediate that the probability that $B_{\tau_1} \geq 3$ is at most $n^{-3+o(1)}$. Also, if there are at most 2 non-good transmissions of r then no player stops pushing r in the first τ_1 rounds, and it is easy to verify that $\tau_1 \leq 4\lg\lg n + O(1)$. □

Let $H_t \subseteq S_t$ be the number of players who know r at the end of round t, and they have not performed more than 2 bad push transmissions of r by that time. Clearly,

$$S_t - \lfloor B_t/3\rfloor \leq H_t \leq S_t.$$

The players in H_t are precisely the players who push r in round $t+1$, if $t < t_{end}$.

CLAIM 3.3. *For any round $t < t_{end}$, if $H_t \geq (\lg n)^4$ and $S_t \leq \frac{3}{4}n$ then, with probability $1 - n^{-\omega(1)}$,*

$$S_{t+1} \geq S_t + H_t\Big(1 - \frac{S_t}{n}\Big)\Big(1 - \frac{S_t}{2n}\Big)\Big(1 - \frac{1}{\lg n}\Big), \qquad (3.1)$$

and

$$B_{t+1} \leq B_t + \frac{H_t S_t}{n}\Big(1 + \frac{1}{\lg n}\Big) + (\lg n)^3. \qquad (3.2)$$

PROOF. The expected number of players who learn r in round $t+1$ is at least

$$\begin{aligned}(n - S_t)\Big(1 - \Big(1 - \frac{1}{n}\Big)^{H_t}\Big) &\geq (n - S_t)\Big(\frac{H_t}{n} - \frac{H_t^2}{2n^2}\Big)\\ &= H_t\Big(1 - \frac{S_t}{n}\Big)\Big(1 - \frac{H_t}{2n}\Big),\end{aligned}$$

which is in $\Omega((\lg n)^4)$. Since the events: "player u learns r in round $t+1$," for players u who do not know r at the end of round t, are negatively dependent [7], we can apply Chernoff bounds to obtain that the probability that fewer than $H_t\big(1 - \frac{S_t}{n}\big)\big(1 - \frac{H_t}{2n}\big)\big(1 - \frac{1}{\lg n}\big)$ players learn r is $e^{-\Omega((\lg n)^2)} = n^{-\Omega(\lg n)}$. For B_{t+1}, we have that the expected number of bad push transmissions in round $t+1$ is $\frac{H_t S_t}{n}$, and, by Chernoff bounds, we can show that the probability there are more than $\frac{H_t S_t}{n}(1 + \frac{1}{\lg n}) + (\lg n)^3$ bad push transmissions is also $e^{-\Omega((\lg n)^2)}$. □

Let $\mathcal{E}$ denote the event: "$B_{\tau_1} \le 2$ and $\tau_1 \le 4\lg\lg n + \mathrm{O}(1)$ and, for all t with $\tau_1 \le t < t_{end}$ such that $H_t \ge (\lg n)^4$ and $S_t \le \frac{3}{4}n$, inequalities (3.1) and (3.2) hold." By Claims 3.2 and 3.3,

$$\mathbb{P}\mathrm{r}[\mathcal{E}] = 1 - n^{-3+\mathrm{o}(1)}. \tag{3.3}$$

We prove Lemma 3.1 by showing that $\mathcal{E}$ implies $S_\tau \ge 3n/4$.

The claims we describe below assume that n is greater than some appropriate constant.

CLAIM 3.4. *Let $\tau_2 = \inf\{t : S_t \ge n/\lg n\}$. If $\mathcal{E}$ occurs then $B_{\tau_2} \le 4n/(\lg n)^2$ and $\tau_2 \le \lg n + \mathrm{O}(1)$.*

PROOF. We show by induction on $t = \tau_1 + 1, \ldots, \tau_2$ that $B_t \le 4S_{t-1}/\lg n$ and $S_t \ge S_{t-1}(2 - \frac{4}{\lg n})$. From this, it follows that $B_{\tau_2} \le 4S_{\tau_2-1}/\lg n < 4n/(\lg n)^2$, and $\tau_2 \le \frac{\lg n}{\lg(2-4/\lg n)} = \lg n + \mathrm{O}(1)$, as desired. The induction is as follows. For the base case $t = \tau_1 + 1$, by (3.2), we have $B_{\tau_1+1} \le 2 + \frac{S_{\tau_1}^2}{n}(1 + \frac{1}{\lg n}) + (\lg n)^3 \le 2 + \frac{S_{\tau_1}}{\lg n}(1 + \frac{1}{\lg n}) + \frac{S_{\tau_1}}{\lg n} \le \frac{3S_{\tau_1}}{\lg n}$. Also, by (3.1), $S_{\tau_1+1} \ge S_{\tau_1} + S_{\tau_1}(1 - \frac{S_{\tau_1}}{n} - \frac{S_{\tau_1}}{2n} - \frac{1}{\lg n}) \ge S_{\tau_1}(2 - \frac{5}{2\lg n})$. Similarly, for the induction step we have that if $t \ge \tau_1 + 1$ then

$$\begin{aligned} B_{t+1} &\le B_t + \frac{H_t S_t}{n}\Big(1 + \frac{1}{\lg n}\Big) + (\lg n)^3 \\ &\le \frac{4S_{t-1}}{\lg n} + \frac{S_t}{\lg n}\Big(1 + \frac{1}{\lg n}\Big) + \frac{S_{t-1}}{\lg n} \\ &\le \frac{5S_t}{(2 - \frac{4}{\lg n})\lg n} + \frac{S_t}{\lg n}\Big(1 + \frac{1}{\lg n}\Big) \\ &= (3.5 + \mathrm{o}(1))S_t/\lg n, \end{aligned}$$

where the second and third inequalities were obtained using the induction hypothesis. Also,

$$\begin{aligned} S_{t+1} &\ge S_t + \Big(S_t - \frac{B_t}{3}\Big)\Big(1 - \frac{S_t}{n} - \frac{S_t}{2n} - \frac{1}{\lg n}\Big) \\ &\ge S_t + \Big(S_t - \frac{4S_t}{3(2 - \frac{4}{\lg n})\lg n}\Big)\Big(1 - \frac{1}{\lg n} - \frac{1}{2\lg n} - \frac{1}{\lg n}\Big) \\ &= S_t\Big(2 - \frac{19 + \mathrm{o}(1)}{6\lg n}\Big). \quad \square \end{aligned}$$

CLAIM 3.5. *Let $\tau_3 = \inf\{t : S_t \ge n/8\}$. If $\mathcal{E}$ occurs then $B_{\tau_3} \le n/16$ and $\tau_3 \le \tau_2 + 2\lg\lg n$.*

PROOF. It is similar to the proof of Claim 3.4. We show by induction on $t = \tau_2 + 1, \ldots, \tau_3$ that $B_t \le S_{t-1}/2$ and $S_t \ge 3S_{t-1}/2$. From this, it follows that $B_{\tau_3} \le S_{\tau_3-1}/2 < n/16$, and $\tau_3 - \tau_2 \le \frac{\lg\lg n}{\lg(3/2)} \le 2\lg\lg n$, as desired. For the base case $t = \tau_2 + 1$ of the induction, we have that $B_{\tau_2+1} \le \frac{4n}{(\lg n)^2} + \frac{S_{\tau_2}}{8}(1 + \frac{1}{\lg n}) + (\lg n)^3 = S_{\tau_2}(\frac{1}{8} + \mathrm{o}(1))$. Also, $S_{\tau_2+1} \ge S_{\tau_2} + (S_{\tau_2} - \frac{B_{\tau_2}}{3})(1 - \frac{1}{8} - \frac{1}{16} - \frac{1}{\lg n}) = S_{\tau_2}(2 - \frac{3}{16} - \mathrm{o}(1))$. For the induction step, we have that if $t \ge \tau_2 + 1$ then $B_{t+1} \le \frac{S_{t-1}}{2} + \frac{S_t}{8}(1 + \frac{1}{\lg n}) + (\lg n)^3 \le \frac{S_t}{3} + \frac{S_t}{8}(1 + \frac{1}{\lg n}) + (\lg n)^3 = S_t(\frac{11}{24} + \mathrm{o}(1))$, and $S_{t+1} \ge S_t + (S_t - \frac{S_t}{6})(1 - \frac{1}{8} - \frac{1}{16} - \frac{1}{\lg n}) \ge S_t(2 - \frac{17}{46} - \mathrm{o}(1))$. $\square$

CLAIM 3.6. *If $\mathcal{E}$ occurs then $S_{\tau_3+5} \ge 3n/4$.*

PROOF. We compute S_t, for $t = \tau_3 + 1, \tau_3 + 2, \ldots$, under the worst-case assumptions that $S_{\tau_3} = n/8$ and $B_{\tau_3} = n/16$, and also that inequalities (3.1) and (3.2) hold as equalities, and $H_t = S_t - \lfloor B_t/3 \rfloor$. We obtain that $S_{\tau_3+5} \ge 3n/4$, for all n greater than a sufficiently large constant. $\square$

Combining now Equation (3.3) and Claims 3.4–3.6 yields Lemma 3.1.

3.2.2 Phase II: Pulling the rumor to the rest of the players

For this phase, we consider the distribution of all the b-bit rumors generated in the same epoch as r; we denote by R the set of these rumors. To ease comprehension we first study the case of $|R| = \mathrm{O}(n)$, separately. In the analysis of this case, only the first component of the digests for rumors in R is used.

The case of $|R| = \mathrm{O}(n)$

We prove the following result.

LEMMA 3.7. *If at the end of round $\tau' \le 4\lg n$ every rumor in R is known to at least a 3/4 fraction of the players then, with probability $1 - |R| \cdot n^{-3+\mathrm{o}(1)}$, all players know all the rumors in R at the end of round $\tau' + 2\lg n$.*

Intuitively, the proof proceeds by lower-bounding the speed at which the slowest-spreading rumor in R is distributed. A key observation is that if a player u does not know a given rumor $r \in R$, but the digest D for the rumors in R that u knows is non-empty and $\mathrm{ind}(r, D) = k$, then for r to be pulled to u it suffices that u call a player who knows both r and the k-th rumor described in D.

Below, r denotes an arbitrary rumor in R. For $i \ge 1$, t_i is the i-th pull round from round $\tau' + 1$, and $U_{i,r}$ is the number of players who do not know r at the end of round t_i. Also $U_{0,r}$ is the same quantity for round τ'. Finally, $U_i = \max_{r \in R} U_{i,r}$.

CLAIM 3.8. *For any $i \ge 0$ such that $t_{i+1} \le t_{end}$, if $U_i \ge (\lg n)^2\sqrt{n}$ then, with probability $1 - n^{-\omega(1)}$,*

$$U_{i+1,r} \le \frac{2}{n}U_i^2\Big(1 + \frac{1}{\lg n}\Big).$$

PROOF. Consider a player u who does not know r at the beginning of round t_{i+1}, and let D be the digest of u for this round, for the rumors in R that u knows. We distinguish two cases:

If $D = \emptyset$ then r is *not* pulled to u in round t_{i+1} iff u calls a player who does not know r, which happens with probability at most $U_{i,r}/n$.

Otherwise, if $\mathrm{ind}(r, D) = k$ then for r to be pulled to u it suffices that u calls a player who knows both r and r_k, the k-th rumor in D; thus, the probability that r is *not* pulled to u is at most $(U_{i,r} + U_{i,r_k})/n$.

So, in both cases, the probability that r is not pulled to u in round t_{i+1} is at most $\frac{2}{n}U_i$. (This bound holds independently of pull transmissions performed by other players in this round.) Therefore, the expected number of players who do not know r at the beginning of round t_{i+1}, and r is not pulled to them in this round is at most $\frac{2}{n}U_iU_{i,r} \le \frac{2}{n}U_i^2$. And, since $U_i \ge (\lg n)^2\sqrt{n}$, by applying Chernoff bounds, we obtain that the number of these players is at most $\frac{2}{n}U_i^2(1 + \frac{1}{\lg n})$ with probability $1 - e^{-\Omega(\lg^2 n)}$. Hence, the same upper bound applies also to $U_{i+1,r}$. $\square$

CLAIM 3.9. *For any $i \ge 0$ such that $t_{i+7} \le t_{end}$, if $U_i \le (\lg n)^2\sqrt{n}$ then, with probability $1 - n^{-3+\mathrm{o}(1)}$, all players know r at the end of round t_{i+7}.*

PROOF. If player u does not know r at the end of round t_i then u does not learn r by the end of round t_{i+7} only if r is not pushed to u in any of the 7 pull rounds following t_i, which happens with probability at most $\left(\frac{2U_i}{n}\right)^7$—by the same reasoning as in the proof of Claim 3.8. Thus, the probability that all players know r at the end of round t_{i+7} is at least $1 - U_{i,r}\left(\frac{2U_i}{n}\right)^7 = 1 - n^{-3+o(1)}$. □

Lemma 3.7 can now be obtained as follows. If for all $i \geq 0$ such that $U_i \geq (\lg n)^2\sqrt{n}$,

$$U_{i+1,r} \leq \frac{2}{n}U_i^2(1 + \frac{1}{\lg n}) = \frac{aU_i^2}{n},$$

where $a = 2(1 + \frac{1}{\lg n})$, then, for those i,

$$U_i \leq \frac{n}{a}\left(\frac{aU_0}{n}\right)^{2^i} \leq \frac{n(a/4)^{2^i}}{a}.$$

From this and Claim 3.8, it follows that $U_{\lg\lg n} < (\lg n)^2\sqrt{n}$ with probability $1 - |R| \cdot n^{-\omega(1)}$. And if $U_{\lg\lg n} < (\lg n)^2\sqrt{n}$ then, by Claim 3.9, it is $U_{\lg\lg n+7} = 0$ with probability $1 - |R| \cdot n^{-3+o(1)}$. Therefore, with probability $1 - |R| \cdot n^{-3+o(1)}$, we have $U_{\lg\lg n+7} = 0$, which implies that all players know all the rumors in R at the end of round $\tau' + \lg n + 7\lg n/\lg\lg n$.

Now, combining Lemmata 3.1 and 3.7 (the former applied for all $r \in R$, and the latter for $\tau' = \tau$) yields the desired bound on the number of rounds of our algorithm.

The case of $|R| = \omega(n)$

For large sets R the previous approach does not work—note the dependence on $|R|$ of the probabilistic bound of Lemma 3.7. We remove this dependence by utilizing the second component of the digests. This component is used to decouple the progress of the distribution of a rumor r from that of rumors that are further than $O(\ell \lg\lg n)$ from r in the ordered list of the rumors in R.

For the analysis, we consider an extra pull phase, before the pull phase we described in the previous case. During this phase every player learns sufficiently many of the rumors in R that are close to r. More specifically, suppose that $R = \{r_1, \ldots, r_\kappa\}$, where $r_1 < \cdots < r_\kappa$. Fix a rumor $r = r_\rho \in R$, and, for $i = 0, \pm1, \ldots, \pm\lg n$, define

$$R_i = \{r_k : (i-\tfrac{1}{2})\ell\lg n < k-\rho \leq (i+\tfrac{1}{2})\ell\lg n \text{ and } 1 \leq k \leq \kappa\}.$$

The extra pull phase in the distribution of r is completed when every player knows at least a 1/3 fraction of the rumors in each of the sets R_i. The next lemma is used to bound the length of this phase. We say that a set $R' \subseteq R$ is a *contiguous* subset of R if $R' = \{r_k : k_1 \leq k \leq k_2\}$, for some $k_1 \geq 1$ and $k_2 \leq \kappa$.

LEMMA 3.10. *Let R' be a contiguous subset of R with $|R'| = \omega(\ell\lg\lg n)$. If at the end of round $\tau' \leq 4\lg n$ every rumor in R' is known to at least a 3/4 fraction of the players then, with probability $1-n^{-3+o(1)}$, every player knows at least a 1/3 fraction of the rumors in R' at the end of round $\tau' + 2\lg n$.*

PROOF. We start by showing that initially, i.e., at the end of round τ', one out of two players already knows half of the rumors in R'. Let f be the fraction of players who each knows half or more of the rumors at the end of round τ'. The average number of rumors a player knows at that time is bounded from below by $\frac{3|R'|}{4}$, and from above by $f|R'| + (1-f)\frac{|R'|}{2}$. Combining the two yields $f \geq \frac{1}{2}$.

Next we show that if in a pull round player u calls a player v who knows $m = \omega(\ell)$ of the rumors in R' then at the end of the round u knows at least $m - O(\ell)$ of the rumors in R'. Let R'_v be the subset of R' that v knows, and let D_u be the digest of u for rumors in R. Let also $I = \bigcup_{r\in R'_v} \text{range}(r, D_u)$, and $R_{u,I}$ be the subset of $R \cap I$ that u knows. If $|R_{u,I}| < |R'_v| = m$, then it is easy to see that at least $m - |R_{u,I}|$ rumors from $R'_v \setminus R_{u,I}$ will be pulled to u. Note that $R_{u,I}$ contains at most 2ℓ rumors that are not in R', those that correspond to the leftmost and the rightmost of the intervals $\text{range}(r, D_u)$, for $r \in R'_v$. Therefore, at the end of the round, u knows at least $m - 2\ell$ of the rumors in R'.

The lemma now follows similarly to the bound on the number of pull rounds required for the standard push-pull algorithm. The fraction of players who at the end of the i-th pull round from round τ' do not know at least $\frac{|R'|}{2} - O(i\ell)$ of the rumors in R' is roughly the square of the corresponding fraction for the $(i-1)$-th pull round; and $\lg\lg n + O(1)$ pull rounds suffice for all players to learn $\frac{|R'|}{2} - O(\ell\lg\lg n) > \frac{|R'|}{3}$ of the rumors in R', with high probability. □

The next lemma is the analogue of Lemma 3.7.

LEMMA 3.11. *If at the end of round $\tau' \leq 4\lg n$ every player knows at least a 1/3 fraction of the rumors in each of the sets R_i, for $i = 0, \pm1, \ldots, \lg n$, then, with probability $1-n^{-3+o(1)}$, all players know r at the end of round $\tau'+2\lg n$.*

PROOF. It is similar to the proof of Lemma 3.7. The key difference is that now we do not focus on the progress of the distribution of all the rumors in R. Instead, in the i-th pull round from round τ', we focus on the progress of the rumors in $\bigcup_{|j|\leq\lg n-i} R_j$. We drop the two outermost R_j, i.e., $R_{\pm(\lg n-i)}$, after the i-th pull round, because, in the next pull round, the progress of rumors in these sets may be impeded by *external* rumors, i.e., rumors not in $\bigcup_{|j|\leq\lg n-i} R_j$. However, since each player knows at least $\frac{1}{3}\ell\lg n \gg \ell$ of the rumors in each R_j, rumors in the remaining R_j are not affected by external rumors in the next pull round. □

Combining Lemmata 3.1, 3.10, and 3.11 (the first lemma applied for all rumors in $\bigcup_{|j|\leq\lg n} R_j$; the second applied for $\tau' = \tau$ and $R' = R_i$, for $i = 0, \pm1, \ldots, \lg n$; and the third applied for $\tau' = \tau + 2\lg n$) yields the desired bound on the number of rounds of our algorithm.

3.2.3 Number of bits communicated

We now establish an upper bound on the number of communication bits used to distribute rumor $r \in R$. Specifically, we show the following lemma.

LEMMA 3.12. *With probability $1-n^{-3+o(1)}$, the total number of bits communicated for the distribution of r is at most $(6+o(1))nb + 6n\lg\lg n\left(\lg b + \frac{\lg\lg\lg n}{|R|} + O(1)\right)$.*

First we count the overhead induced by unnecessary push transmissions. A push transmission of r from player u to player v in round t is unnecessary if one of the following two conditions applies:

- v already knows r at the beginning of round t; such a push transmission is called *bad*.

- v does not know r, but in round t, r is pulled to v from some player or r is pushed to v form a player u' such that $u' < u$, with respect to some fixed ordering of the players; such a push transmission is called *unlucky*.

Clearly, bad push transmissions of r result in a communication overhead of at most $3nb'$ bits, where

$$b' = b + \Theta(\lg\lg n)$$

is the size of the rumor plus the age counter. The next result bounds the overhead due to unlucky push transmissions.

CLAIM 3.13. *With probability $1-n^{\omega(1)}$, unlucky push transmissions of r result in a communication overhead of at most $(1+\mathrm{o}(1))nb'$ bits.*

PROOF. Let $\mathbf{S}_t$ be the set of players who know r at the end of round t, and let $S_t = |\mathbf{S}_t|$; $\mathbf{S}_0$ is the set of sources of r. Fix the sequence $\{\mathbf{S}_t : t \geq 0\}$. All the probabilistic statements described below will be implicitly conditioned on this sequence. For any round t and any player $u \in \mathbf{S}_t$, let $X_{u,t}$ be the indicator random variable that is 1 iff u performs an unlucky push transmission of r in round t. The expected value of $X_{u,t}$ is at most $\frac{S_{t+1}-S_t}{S_{t+1}}$: this expected value is a non-decreasing value of u, for $u \in \mathbf{S}_t$; and if u is the largest player that pushes r in round t and this transmission is bad or unlucky then the recipient is equally likely to be any of the players in $\mathbf{S}_{t+1}$. So, the expected value of the total number $\sum_t \sum_{u\in\mathbf{S}_t} X_{u,t}$ of unlucky push transmissions of r is at most $n - S_0$. The upper bound above on the expectation of each $X_{u,t}$ holds independently of the values of the other indicator variables, so, we can apply Chernoff bounds to obtain that at most $(1+\mathrm{o}(1))n$ unlucky push transmissions of r occur, with probability $1 - n^{\omega(1)}$. □

Next we count the overhead induced by pull transmissions. The size of the digest for the rumors in R that a player knows is at most $(\lg\lg\lg n + |R|\lg b) + \frac{|R|b}{\ell} + \mathrm{O}(|R|)$. Since there are at most $6\lg\lg n$ pull rounds during which the rumors in R are not cold, the total overhead per rumor because of the digests is at most $6n\lg\lg n(\frac{\lg\lg\lg n}{|R|} + \lg b + \frac{b}{\ell} + \mathrm{O}(1))$. Finally, there are at most $n|R|$ redundant pull transmissions of rumors in R. This is because for every redundant pull transmission of a rumor in R there is at least one useful transmission of another rumor in R (see the second-to-last paragraph in Section 3.1). Note that, because of the way unlucky push transmissions were defined, there are no "unlucky" pull transmissions.

Combining the above we obtain that, with probability $1 - n^{-3+\mathrm{o}(1)}$, the total number of bits communicated for the distribution of r is at most $nb' + 3nb' + (1+\mathrm{o}(1))nb' + 6n(\lg\lg n)\big(\frac{\lg\lg\lg n}{|R|} + \lg b + \mathrm{O}(1)\big) + nb'$, where the first term nb' accounts for the useful transmissions of r. The above expression is equal to the expression in the statement of Lemma 3.12.

4. LOWER BOUND

In this section, we prove the following lower bound on the performance of rumor-spreading algorithms in the random phone-call model.

THEOREM 4.1. *For any $b \geq 1$, no address-oblivious algorithm can guarantee that for any rumor of size b, this rumor is distributed to all players within $\mathrm{O}(\lg n)$ rounds, with constant probability, and $\mathrm{o}(nb + n\lg\lg n)$ bits of communication are used, in expectation.*

Karp *et. al* established a lower bound of $\Omega(n\lg\lg n)$ on the expected number of *messages*, for any address-oblivious algorithm guaranteeing that any one-bit rumor is distributed to all players with constant probability (Theorem 4.1 in [16]). From this result, it is immediate that $\Omega(n\lg\lg n)$ *bits* of communication are required in expectation, for any b. Hence, it remains to prove that the theorem holds for rumor sizes $b = \omega(\lg\lg n)$. We first consider the case $b = \omega(\lg n)$, in Section 4.1, and then we reduce to this case the case of smaller b, in Section 4.2.

4.1 The case of large rumors

Suppose that $b = \omega(\lg n)$. Consider the following setting, which we will refer to as the *single b-bit rumor scenario*: There is only one rumor, which is drawn uniformly at random among all the b-bit rumors. The rumor starts from player s in round 0. The size b, the source s, and the start round of the rumor are known to all players. Also, in each phone call, the two participants know the id of one another; so, the rumor-spreading algorithm can be non address oblivious. Suppose now that an algorithm guarantees that in the above scenario, the rumor is spread to all players within $\rho = \mathrm{O}(\lg n)$ rounds, with at least some constant probability $p > 0$. We show that the algorithm uses an expected number of $\Omega(nb)$ communication bits. The theorem then follows.

We bound the expected number of bits exchanged by a single player. Consider a player $u \neq s$, and let B_u be the total number of bits u exchanges (sends or receives) in the first ρ rounds, i.e., in rounds $0, \ldots, \rho-1$. Define the events:

$\mathcal{E}$: all players know the rumor at the end of round $\rho - 1$;

$\mathcal{E}_u$: u knows the rumor at the end of round $\rho - 1$;

$\mathcal{C}$: u receives at most $2\rho + \lg n$ calls in the first ρ rounds;

$\mathcal{B}_k$: u exchanges at most k bits with other players in the first ρ rounds, i.e., $B_u \leq k$.

We have that for any k,

$$\begin{aligned}\mathbb{E}[B_u] &\geq k\,\mathbb{P}\mathrm{r}[B_u \geq k] \geq k\,\mathbb{P}\mathrm{r}[\mathcal{E}]\cdot\mathbb{P}\mathrm{r}[B_u \geq k \,|\, \mathcal{E}]\\ &\geq kp(1 - \mathbb{P}\mathrm{r}[\mathcal{B}_k \,|\, \mathcal{E}]),\end{aligned}$$

since $\mathbb{P}\mathrm{r}[\mathcal{E}] \geq p$. Also,

$$\begin{aligned}\mathbb{P}\mathrm{r}[\mathcal{B}_k \,|\, \mathcal{E}] &= \mathbb{P}\mathrm{r}[\mathcal{B}_k \wedge \mathcal{E}]/\,\mathbb{P}\mathrm{r}[\mathcal{E}] \leq p^{-1}\,\mathbb{P}\mathrm{r}[\mathcal{B}_k \wedge \mathcal{E}_u]\\ &\leq p^{-1}\left(\mathbb{P}\mathrm{r}[\mathcal{B}_k \wedge \mathcal{E}_u \,|\, \mathcal{C}] + \mathbb{P}\mathrm{r}[\bar{\mathcal{C}}]\right).\end{aligned}$$

Since the expected number of calls that u receives in the first ρ rounds is ρ, using Chernoff bounds we can show that

$$\mathbb{P}\mathrm{r}[\bar{\mathcal{C}}] \leq e^{-(\rho+\lg n)/3} \leq e^{-\lg n/3}.$$

Also, a counting argument yields the following claim.

CLAIM 4.1. $\mathbb{P}\mathrm{r}[\mathcal{B}_k \wedge \mathcal{E}_u \,|\, \mathcal{C}] \leq 6^{3\rho+\lg n+k}/2^b$.

PROOF. We start with two definitions. An *i-call-history* of u specifies the player that u calls, and the set of players that u receives calls from in each of the first i rounds. An *i-history* of u specifies an i-call-history of u, and also the sequence of messages exchanged between u and each of the players that u communicates with in the first i rounds.

For any ρ-call-history of u in which u receives no more than $2\rho + \lg n$ calls, there are at most $6^{\rho+(2\rho+\lg n)+k}$ distinct ρ-histories of u with that ρ-call-history, in which at most k bits are exchanged between u and the rest of the players. This follows from the observation that any such ρ-history can be represented by a string of length at most $\rho + (2\rho + \lg n) + k$ over the alphabet {*end-round*, *begin-call*, *send*-0, *send*-1, *recv*-0, *recv*-1}: the messages that u exchanges during round i are described by the substring between the i-th and the $(i+1)$-th *end-round* symbols of the string; the *begin-call* symbols separate the communication streams of u with different players in the same round; and the sending (receipt) of bit $x = 0, 1$ by u is represented by the symbol *send-x* (*recv-x*).

So, for any ρ-call-history of u in which u receives no more than $2\rho + \lg n$ calls, there are at most $6^{\rho+(2\rho+\lg n)+k}$ rumors that u can distinguish in ρ rounds exchanging at most k bits.[5] Therefore, conditioned on any such ρ-call-history of u, the probability that in the first ρ rounds, u learns the rumor and it exchanges no more than k bits is at most $6^{\rho+(2\rho+\lg n)+k}/2^b$. (Recall that the rumor is chosen at random among the 2^b b-bit tumors.) This implies the claim. □

Combining all the above yields

$$\mathbb{E}[B_u] \geq kp\big(1 - p^{-1}6^{3\rho+\lg n+k}2^{-b} - p^{-1}e^{-\lg n/3}\big),$$

and setting $k = \lfloor (b - 1 + \lg p)/\lg 6 - 3\rho - \lg n \rfloor = \Theta(b)$, we obtain $\mathbb{E}[B_u] \geq kp\big(1/2 - p^{-1}e^{-\lg n/3}\big) = \Theta(b)$. Thus, the expected value of the total number of bits exchanged is at least $\sum_{u \neq s} \mathbb{E}[B_u]/2 = \Omega(nb)$.

4.2 The case of smaller rumors

Suppose now that $\omega(\lg\lg n) \leq b \leq \mathrm{O}(\lg n)$. We show that given an algorithm $\mathcal{A}$ that provides the guarantees described in Theorem 4.1 for that b, we can devise an algorithm $\mathcal{A}'$ that contradicts the result of Section 4.1. That is, for some $b' = \omega(\lg n)$, $\mathcal{A}'$ guarantees that in the single b'-bit rumor scenario, the rumor is distributed to all players within $\mathrm{O}(\lg n)$ rounds, with constant probability, and $\mathrm{o}(nb')$ communication bits are used, in expectation. Roughly speaking, $\mathcal{A}'$ encodes the b'-bit rumor as a collection of b-bit rumors, which are then spread using $\mathcal{A}$.

Suppose that $\mathcal{A}$ ensures that with probability p any b-bit rumor is distributed to all players within ρ rounds. To ease comprehension we consider the case $p = 1$ separately, first. If $p = 1$ then $\mathcal{A}'$ is the following simple algorithm. Let $b' = 2^{b/2-1}b$ (to simplify exposition we assume that b is even). The b'-bit rumor is divided into $2^{b/2}$ substrings $w_0, \ldots, w_{2^{b/2}-1}$ of size $b/2$ each. For each w_i, a b-bit rumor is build consisting of w_i and its index i. All these $2^{b/2}$ b-bit rumors are then spread in parallel (starting at round 0) as in algorithm $\mathcal{A}$. With probability $p = 1$, every player learns all the rumors within ρ rounds, and can easily reconstruct the initial b'-bit rumor. Note that $b' = \omega(\lg n)$ and the expected total number of bits communicated is $2^{b/2} \cdot \mathrm{o}(nb) = \mathrm{o}(nb')$, as desired.

When $p < 1$ the above scheme does not work, because $\mathcal{A}$ does not guarantee that, with constant probability, every player learns *all* the $2^{b/2}$ rumors. (E.g., it may be that with probability 1/2 all players learn the first of these rumors and they do not learn the second, and with probability 1/2 they all learn the second and not the first.) We tackle this problem by employing an error-correction scheme which facilitates reconstruction of the b' bit rumor by just a fraction of the b-bit rumors (such schemes are often called erasure codes). Recall from code theory that a *q-ary (ℓ, M, d)-code* is a set of M *codewords*, where each codeword is a string of length ℓ over an alphabet of size q, and the minimum distance between codewords is d, i.e., any two codewords differ in at least d positions. We employ a q-ary (ℓ, M, d)-code C with $q = \ell = 2^{b/2}$, $M = q^{\ell/4}$, and $d = \ell/2$. By the Gilbert–Varshamov bound (see, e.g., [18]), such a code C exists, because $M < q^\ell / \sum_{i=0}^{d-1} \binom{\ell}{i}(q-1)^i$. Algorithm $\mathcal{A}'$ is then as follows. The rumor size is $b' = 2^{b/2-3}b$. Each b'-bit rumor is mapped to a distinct codewords of C, and this codeword is distributed instead of the actual rumor. Similarly to the case $p = 1$, for each q-ary symbol of the codeword, a b-bit rumor is built consisting of that symbol and its order in the codeword, and the resulting ℓ rumors are spread as in algorithm $\mathcal{A}$. Now, for a player to be able to reconstruct the codeword it suffices to learn $\ell - d + 1$ different b-bit rumors, since any two codewords differ in at least d positions. We can lower-bound the probability that this happens as follows. We assume without loss of generality that $p \geq 1 - 1/e$—if this is not the case, we can achieve that by re-sending the b-bit rumors in the rounds $i\rho$, for $i = 1, \ldots, 1/p$. Let X_i be an indicator random variable that is 1 iff all players learn the i-th b-bit rumor by round ρ, and let $X = \sum_i X_i$. It is $\mathbb{E}[X_i] \geq p$, and, thus,

$$\mathbb{E}[X] \geq p\ell.$$

Also, if p' is the probability that *every* player learns at least $\ell - d + 1$ of the ℓ b-bit rumors by round ρ then

$$\mathbb{E}[X] \leq p'\ell + (1 - p')(\ell - d) \leq p'\ell + \ell - d.$$

Therefore, $p\ell \leq p'\ell + \ell - d$, which yields

$$p' \geq p - 1 + d/\ell \geq 1/2 - 1/e,$$

since $p \geq 1 - 1/e$. We have thus shown that, with constant probability, all players learn enough b-bit rumors in the first ρ rounds to reconstruct the codeword, and learn the b'-bit rumor. As in case $p = 1$, $b' = \omega(\lg n)$ and the total number of bits communicated in expectation is $\ell \cdot \mathrm{o}(nb) = \mathrm{o}(nb')$.

[5] Different executions of the algorithm that have the same ρ-history of u are indistinguishable to u until at least the beginning of round ρ. So, in all these executions, within the first ρ rounds, u learns the same rumor, if any.

5. ACKNOWLEDGMENTS

We thank the anonymous referees of this paper for their helpful comments.

6. REFERENCES

[1] P. Berenbrink, R. Elsässer, and T. Friedetzky. Efficient randomised broadcasting in random regular networks with applications in peer-to-peer systems. In *Proc. 27th ACM Symp. on Principles of Distributed Computing (PODC)*, pages 155–164, 2008.

[2] B. Bloom. Space/time trade-offs in hash coding with allowable errors. *Communications of the ACM*, 13(7):422–426, 1970.

[3] A. Broder and M. Mitzenmacher. Network applications of Bloom filters: A survey. *Internet Mathematics*, 1(4):485–509, 2005.

[4] 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. 6th ACM Symp. on Principles of Distributed Computing (PODC)*, pages 1–12, 1987.

[5] B. Doerr, T. Friedrich, and T. Sauerwald. Quasirandom rumor spreading. In *Proc. 19th ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 773–781, 2008.

[6] B. Doerr, T. Friedrich, and T. Sauerwald. Quasirandom rumor spreading: Expanders, push vs. pull, and robustness. In *Proc. 36th Int. Colloq. on Automata, Languages and Programming (ICALP)*, pages 366–377, 2009.

[7] D. Dubhashi and D. Ranjan. Balls and bins: A study in negative dependence. *Random Struct. Algorithms*, 13(2):99–124, 1998.

[8] R. Elsässer. On randomized broadcasting in power law networks. In *Proc. 20th Int. Symp. on Distributed Computing (DISC)*, pages 370–384, 2006.

[9] R. Elsässer. On the communication complexity of randomized broadcasting in random-like graphs. In *Proc. 18th ACM Symp. on Parallelism in Algorithms and Architectures (SPAA)*, pages 148–157, 2006.

[10] R. Elsässer and T. Sauerwald. The power of memory in randomized broadcasting. In *Proc. 19th ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 218–227, 2008.

[11] U. Feige, D. Peleg, P. Raghavan, and E. Upfal. Randomized broadcast in networks. *Random Structures and Algorithms*, 1(4):447–460, 1990.

[12] P. Fraigniaud and E. Lazard. Methods and problems of communication in usual networks. *Discrete Appl. Math.*, 53(1-3):79–133, 1994.

[13] A. Frieze and G. Grimmett. The shortest-path problem for graphs with random arc-lengths. *Discrete Appl. Math.*, 10:57–77, 1985.

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

[15] J. Hromkovic, R. Klasing, B. Monien, and R. Piene. Dissemination of information in interconnection networks (broadcasting & gossiping). In *Combinatorial Network Theory*, pages 125–212. Springer, 1995.

[16] R. Karp, C. Schindelhauer, S. Shenker, and B. Vocking. Randomized rumor spreading. In *Proc. 41st IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 565–574, 2000.

[17] B. Pittel. On spreading a rumor. *SIAM J. Appl. Math.*, 47(1):213–223, 1987.

[18] J. H. van Lint. *Introduction to Coding Theory*. Springer Verlag, 3rd edition, 1998.