Optimal Gossiping with Unit Size Messages in Known Topology Radio Networks*

Fredrik Manne and Qin Xin

Department of Informatics, The University of Bergen, Norway
{fredrikm, xin}@ii.uib.no.

Abstract. Gossiping is a communication primitive where each node of a network possesses a unique message that is to be communicated to all other nodes in the network. We study the gossiping problem in known topology radio networks where the schedule of transmissions is precomputed in advance based on full knowledge about the size and the topology of the network. In addition we consider the case where it is only possible to transmit a unit size message in each time step. This gives a more realistic model than if arbitrary length messages can be sent during each time step, as has been the case in most previous studies of the gossiping problem. In this paper, we propose an optimal randomized schedule that uses $O(n \log n)$ time units to complete the gossiping task with high probability in any radio network of size n. This matches the lower bound of $\Omega(n \log n)$ by Gąsieniec and Potapov in [17] [TCS'02]. Our new gossiping schedule is based on the notion of a gathering spanning tree proposed by Gąsieniec, Peleg and Xin in [19] [PODC'05].

Keywords: Centralized radio networks, gossiping, randomized schedule.

1 Introduction

The two classical problems of information dissemination in computer networks are the *broadcasting* problem and the *gossiping* problem. The broadcasting problem requires distributing a particular message from a distinguished *source* node to all other nodes in the network. In the gossiping problem, each node v in the network initially holds a message m_v, and the aim is to distribute all messages to all nodes. For both problems, one generally considers as the efficiency criterion the minimization of the time needed to complete the task.

This paper concerns the following model of a radio network. A network is an undirected connected graph $G = (V, E)$, where V represents the set of nodes of the network and E contains unordered pairs of distinct nodes, such that $(v, w) \in E$ iff the transmissions of node v can directly reach node w and vice versa (the reachability of transmissions is assumed to be a symmetric relation). In this case, we say that the nodes v and w are *neighbours* in G. One of the particular properties of radio network is that a message transmitted by a node is always sent to all of its neighbours.

* Supported by the Research Council of Norway through the SPECTRUM project.

T. Erlebach (Ed.): CAAN 2006, LNCS 4235, pp. 125–134, 2006.

The number of neighbours of a node w is called its *degree*, and the maximum degree of any node in the network is called the *maximum degree* of the network and is denoted by Δ. The *size of the network* is the number of nodes $n = |V|$.

Communication in the network is synchronous and consists of a sequence of communication steps. During each step, each node v either transmits or listens. If v transmits, then the transmitted message reaches each of its neighbours by the end of this step. However, a node w adjacent to v successfully receives this message iff w is listening during this step and v is the only transmitting node among w's neighbours. If node w is adjacent to a transmitting node but it is not listening, or it is adjacent to more than one transmitting node, then a *collision* occurs and w does not retrieve any message in this step.

The running time of any communication schedule is determined by the number of time steps required to complete the communication task. That is, we do not account for any internal computation within individual nodes.

Most of the work in this field has been done under the assumption that the processors can transmit messages of arbitrary size in a single time step. In particular any node can send all of the information it has received so far in a single (atomic) step of the communication process. Note that this strong assumption is rather unrealistic if the size of the network is very large. In this paper we study the gossiping problem in radio networks where there is a restriction on the size of each message. In particular we investigate the case where each message is of unit size, meaning that it contains information originating from exactly one node of the network.

We focus on algorithms that rely on using complete information about the network topology. This type of topology-wise communication algorithms are useful in radio networks that have a reasonably stable topology/infrastructure. As long as no changes occur in the network topology during the execution of the algorithm, the tasks of broadcasting and gossiping will be completed successfully. Note also that our main goal is the design of time efficient communication procedures. However, it would not be difficult to increase the level of fault-tolerance in our algorithm at the expense of some small extra time consumption.

Communication in radio networks with known topology. The work on communication in known topology radio networks was initiated in the context of the broadcasting problem. In [13], Gaber and Mansour prove that the broadcasting task can be completed in time $O(D \log^2 n)$ where D is the diameter of the network. An $\Omega(\log^2 n)$ time lower bound was proved for the family of graphs of radius 2, see [2] by Alon *et al.* While it was known for quite a while that for every n-node radio network that there exists a deterministic broadcasting schedule of length $O(D \log n + \log^2 n)$, Bar-Yehuda *et al.* [3], an appropriate efficient construction for such a schedule was only recently proposed in [21] by Kowalski and Pelc. Subsequently, an efficient deterministic construction of a broadcasting schedule of length $D + O(\log^4 n)$ was proposed by Elkin and Kortsarz [12]. In this paper, they also present an efficient deterministic construction for a broadcasting schedule of length $D + O(\log^3 n)$ for planar graphs. In [19], Gąsieniec, Peleg and Xin proposed a more efficient deterministic schedule that uses $D + O(\log^3 n)$ time units to complete the broadcasting task in any radio network. This paper also contains an optimal randomized broadcasting schedule of length $D + O(\log^2 n)$ and a new broadcasting schedule using fewer than $3D$ time slots on planar graphs. More recently in [7], Cicalese, Manne and Xin improved the broadcasting time to $D + O(\frac{\log^3 n}{\log \log n})$ in any radio network.

Efficient radio broadcasting algorithms for several types of known network topologies can be found in Diks *et al.* [10]. For general networks, however, it is known that the computation of an optimal (radio) broadcast schedule is NP-hard, even if the underlying graph is embedded in the plane [5,23].

Radio gossiping in networks with known topology was first studied in the context of radio communication with messages of limited size, see [17] by Gąsieniec and Potapov. In this model the authors proposed several optimal or close to optimal $O(n)$-time gossiping procedures for various standard network topologies, including lines, rings, stars and free trees. They also proved that there exists a radio network topology in which the gossiping (with unit size messages) requires $\Omega(n \log n)$ time. The first work on radio gossiping in known topology networks with arbitrarily large messages is [18], where Gąsieniec, Potapov and Xin propose several optimal gossiping schedules for a wide range of radio network topologies. Very recently, Gąsieniec, Peleg and Xin proposed an efficiently computable deterministic schedule that uses $O(D + \Delta \log n)$ time units to complete the gossiping task in any radio network [19]. This improves on the previous best known gossiping schedule [18] with running time $O(D + \sqrt[i+2]{D}\Delta \log^{i+1} n)$, for any network with diameter $D = \Omega(\log^{i+4} n)$,where i is an arbitrary integer constant $i \geq 0$. Subsequently in [7], Cicalese, Manne and Xin improved the gossiping time even further to $O(D + \frac{\Delta \log n}{\log \Delta - \log \log n})$ in radio networks where $\Delta = \Omega(\log n)$.

Our results. In this paper, we study the gossiping problem in known topology radio networks, where during each time step only one unit size message originating from some node of the network can be transmitted successfully. The schedule of transmissions is precomputed in advance based on full knowledge about the size and the topology of the network. We propose an optimal randomized schedule that uses $O(n \log n)$ time units to complete the gossiping task with high probability in any radio network of size n. This matches the lower bound of $\Omega(n \log n)$ by Gąsieniec and Potapov in [17]. Our new gossiping schedule is based on the notion of a gathering spanning tree proposed by Gąsieniec, Peleg and Xin in [19].

2 Centralized Gossiping with Unit Size Messages in Arbitrary Graphs

In this section, we study the time complexity of gossiping in general undirected graphs. We show that radio gossiping with unit size messages in undirected graphs can be performed in time $O(n \log n)$ with high probability. Our gossiping algorithm runs in two stages. In the first stage, we collect all the messages in a distinguished central node c by transporting messages along branches of any BFS spanning tree rooted in c. The second stage is performed through broadcasting of n unit messages from c. These broadcasts are performed in a pipelined fashion along a gathering spanning tree, a structure first proposed by Gąsieniec, Peleg and Xin in [19].

2.1 Gathering Messages in Arbitrary Graphs

Given an arbitrary graph $G = (V, E)$ and a BFS spanning tree T rooted at its central node c, we partition the nodes into consecutive layers $L_i = \{v \mid dist(c, v) = i\}$, for $i = 0, .., r$ where r is a radius of T.

In the following, we will use the standard notions of *parent, children, and descendant* in trees. For simplicity we assume that a node is a descendant of itself.

We say that a node v is unsecured iff v has not delivered all messages stored originally in its descendant to $parent(v)$ in T. The different messages are transmitted toward c in a pipelined fashion, such that only one unsecured node v in L_i of T is allowed to transmit a message to $parent(v)$ in T at any time step t, for $1 \leq i \leq r$. To avoid collisions between different BFS layers, the nodes in L_i are allowed to transmit in time steps t satisfying $t = i \bmod 3$.

The following result follows directly.

Lemma 1. *All different messages can be gathered at the central nodes c of G in time $3(n-1)$.*

2.2 Broadcasting in Arbitrary Graphs

Now that all messages have been gathered in c, we will show how we with high probability can broadcast them to all nodes in G in $O(n \log n)$ time.

We first recall the following recursive ranking procedure of nodes in a tree (see [19]). Leaves have rank 1. Next consider a node v and the set Q of its children and let r_{max} be the maximum rank of the nodes in Q. If there is a unique node in Q of rank r_{max} then set the rank of v to r_{max}, otherwise set the rank of v to $r_{max} + 1$.

Lemma 2. *The largest rank in a tree of size n is bounded by $\lceil \log n \rceil$. (see [19]).*

Given any graph G with central node c, a *gathering spanning tree* of G is a BFS spanning tree T of G rooted at c, such that T is ranked as above and that also satisfies the following condition: every node in L_{j+1} of rank i is at most adjacent to one node in L_j also of rank i, and if all the nodes of rank i in L_j transmit at the same time then the messages will be received by the nodes in L_{j+1} of rank i successfully without any collision.

The following lemma was shown in [19].

Lemma 3. *There exists a polynomial time construction of a gathering spanning tree in any graph G.*

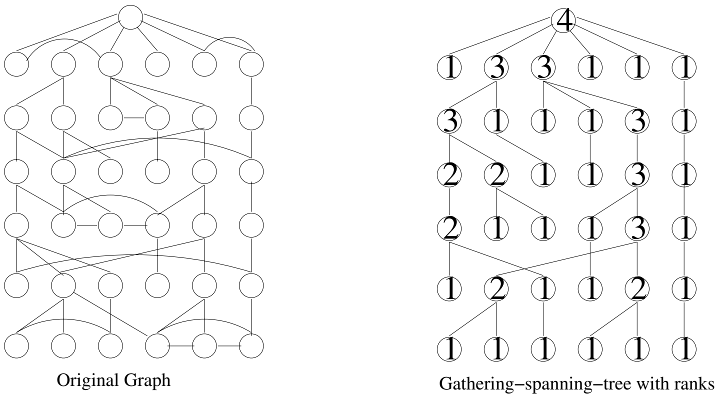

Fig. 1. Creating a gathering spanning tree

Figure 1 shows how a gathering spanning tree can be constructed from a graph G.

For a gathering spanning tree T we say that an edge in T is *fast* if both of its end points have the same rank, and it is *slow* otherwise. Since the largest rank is at most $\lceil \log n \rceil$, there are at most $\lceil \log n \rceil$ slow edges in each path from the root c to any leaf of T.

For a graph G with gathering spanning tree T, we now partition the edges of every path of T from c to a leaf into consecutive edge segments in the following way.

(1) Every maximal connected path of fast edges in T is a *fast* segment.
(2) Every slow edge is a *slow* segment.

Note that a node can belong to both a fast and a slow segment. In the following, the time steps t are divided into *fast blocks* ($t = 0 \bmod 2$) and *slow blocks* ($t = 1 \bmod 2$), such that the communication within the fast segments of T only occur in the fast blocks and similarly, communication within the slow segments of T only occur in the slow blocks. We will not be explicit about this schedule in the future presentation but assume that the time units used for both the fast and slow segments are consecutive.

We now define a graph $\tilde{G} = (\tilde{V}, \tilde{E})$ as follows. Its nodes are the same as in G and $E \subset \tilde{E}$. In addition for every node v in a fast segment of T we add an edge (v, w) where w is the topmost node of the fast segment that v belongs to. See Figure 2 for an example.

Lemma 4. *The graph $\tilde{G}$ has radius at most $2 \log n$.*

Proof. The lemma follows directly from the definition of $\tilde{G}$ and Lemma 2.2.

To simplify our presentation, we also define a modified gathering spanning tree $\tilde{T}$ of $\tilde{G}$ as follows. The central node c is the root of $\tilde{T}$ as well. Every edge of T between two nodes of different rank will be also an edge in $\tilde{T}$. In addition, if w

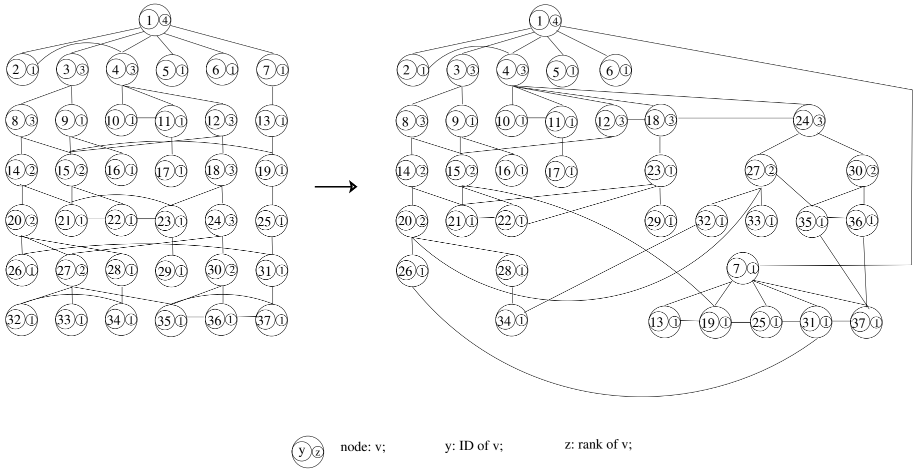

Fig. 2. From the original graph G to $\tilde{G}$

is the topmost node of a fast segment F in T, then w has an adjacent edge in $\tilde{T}$ to each node $v \in F$. We denote each node that belongs to a fast segment in T as fast in $\tilde{T}$. Thus the main change from T to $\tilde{T}$ is that the nodes of every fast segment in T has been collapsed and are now hanging of the topmost node in the segment. For all other nodes their parent relationship in $\tilde{T}$ is the same as in T. The main purpose for defining $\tilde{T}$ is to be able to reason about the time complexity of performing the slow transmissions.

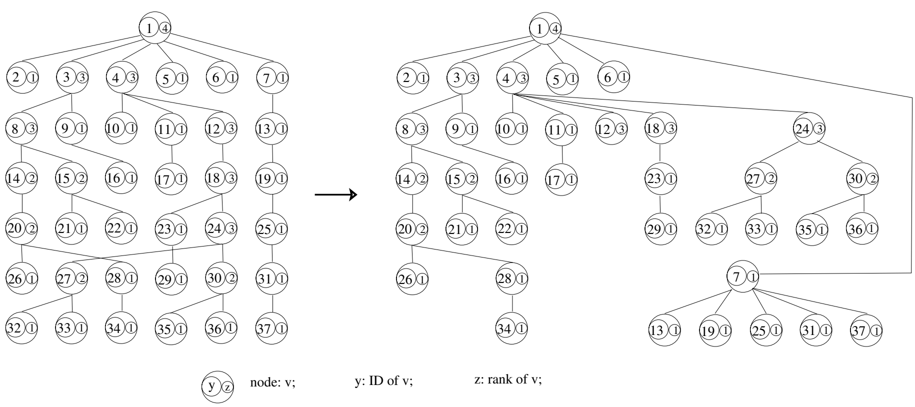

Fig. 3. From a T of G to a $\tilde{T}$ of $\tilde{G}$

Observation 1. *$\tilde{T}$ of $\tilde{G}$ spans G as well.*

We are now ready to describe the gossiping schedule, which consists of *fast steps* and *slow steps* as described below.

(1) **Fast steps:** These are used for transmissions between two fast nodes that are adjacent in T. Let $r_{max} \leq \lceil \log n \rceil$ be the largest rank in T. Consider a node v of rank j, $1 \leq j \leq r_{max}$ on BFS layer L_i in T and which is also the uppermost node of a fast segment F in T. If v receives a new message, then v is set to perform a transmission to its immediate fast child w in F in time step t', where $t' = i + 3j \bmod 3r_{max}$.

Note that since w is in L_{i+1} and has the same rank as v it follows that w will transmit this message to its fast child (if it is not the lowest node in F) in time step $t'' = t' + 1 \bmod 3r_{max}$. Thus it follows that after at most $3r_{max} + |F| - 1$ time steps the message will have reached every node in F.

(2) **Slow steps:** These are used to transmit messages across slow edges. Consider a node v on BFS layer L_i in $\tilde{T}$ that belongs to a slow segment of $\tilde{T}$. If v receives a new message, then v will perform a transmission only in time step t' satisfying $t' = i \bmod 3$. We employ the procedure *RCW* from [19] to perform the transmissions in the slow steps.

Procedure *RCW* allows the movement of one unit messages from one partition of a bipartite graph of size n' (here, an entire BFS layer L_i of $\tilde{T}$) to the other (here, the next layer L_{i+1}) with high probability in time $O(\log n')$. Note that as soon as the message has reached a (slow) node w in layer L_{i+1} then w will start to transmit this to its (slow) descendants (within 3 time units).

The probability that the procedure *RCW* is successful in transmitting a message between two nodes is given by the the following lemma from [19].

Lemma 5. *The probability that a node v in layer L_i will be successful in transmitting a message to its adjacent nodes in L_{i+1} is given by $p \geq 1/(4e)$.*

Note that the described pattern of transmissions separates the transmissions between the fast and slow steps by at least one unit of time. The pattern also ensures that at any time step, transmissions are performed on BFS layers at distances that are multiples of 3 apart. Thus there will be no conflicts between transmissions coming from different BFS layers. There is also no collisions in transmissions between the fast nodes of $\tilde{T}$ with the same rank and within the same BFS layer of T due to the properties of a gathering spanning tree (see [19]).

Corollary 1. *The total time spent on fast transmissions in sending a unit message from c to a leaf v is $O(\log^2 n + r)$ where r is the radius of G.*

Proof. As stated, the time to traverse a fast segment F takes time $3r_{max}+|F|-1 = O(\log n + |F|)$. There are at most $\log n$ fast segments on the path from c to v in T and the sum of the lengths of these segments is at most r, thus the result follows.

Since the construction of a gathering spanning tree T (and consequently of $\tilde{G}$ and $\tilde{T}$) is polynomial in view of Lemma 2.2, we have the following result.

Theorem 2. *For any n-node graph G, there exists a randomized gossiping algorithm with unit size messages that runs in time polynomial in n.*

It remains to bound the probability of success of our gossiping schedule and to estimate the length of the scheme constructed by it.

Theorem 3. *There exists a randomized algorithm that for any known topology radio network of size n, following a polynomial time preprocessing stage, solves the gossiping problem with unit size messages with high probability in time $O(n \log n)$.*

Proof. Consider an arbitrary node v in the graph G, and consider the path along which it is supposed to get the message from the root. This path is divided into "fast segments" and "slow segments" as discussed above. Consider first the fast segments on the path from c to a leaf v and let F be the topmost segment on this path and w the topmost node of F. Then it follows that the messages will be transmitted from w in time steps at most $O(\log n)$ apart. Thus the last message will be transmitted after $O(n \log n)$ time. >From Corollary 1 it follows that this message will spend $O(n + \log^2 n)$ time on fast transmissions before it reaches every descendant including v. Thus the total time spent on fast transmissions is bounded by $O(n \log n + \log^2 n)$.

We now claim that the total number of time steps spent by the messages for the slow steps on its way to v is at most $O(n \log n)$ as well. For each slow transmission we will be activating the RCW procedure $O(1)$ times to ensure that it reaches every node with high probability. Thus the number of times the RCW procedure is activated for a particular message is bounded by the height of $\tilde{T}$ which is $O(\log n)$. It follows that the total time spent on slow transmissions for one particular message is bounded by $O(\log^2 n)$. Since the slow transmissions are performed at intervals that are $O(1)$ apart the last message will start to be transmitted after $O(n \log n)$ time and reach every node after spending $O(\log^2)$ time on slow transmissions. Thus the combined time spent on both fast and slow transmissions is bounded by $O(n \log n)$.

It remains to show that each message will reach every node with high probability. Note first that each participation of a particular message in an activation of the RCW procedure succeeds (i.e., the message crosses from its current node to the next node on the path to v) independently with constant probability $p \geq 1/(4e)$. Then let $\tilde{R}$ be the height of $\tilde{T}$ and X a random variable denoting the number of successes of a particular message on its way from c to v. I.e. X denotes the number of levels of $\tilde{T}$ that the message has crossed successfully. For each level that the message has to cross we will activate RCW a total of $24e$ times. Thus the maximum total number of times the message will participate in the RCW procedure is $24e\tilde{R}$, the expected value of X is $\mu = 6\tilde{R}$. Due to the way $\tilde{T}$ was constructed, we know that $\tilde{R}$ is bounded by $2 \log n$.

Using the Chernoff bound, the probability $P_{fail}(v)$ that the message will not reach v after $24e\tilde{R}$ participations in RCW can be bounded from above by

$$P_{fail}(v) \leq P(X < \tilde{R}) = P(X < (1 - 5/6)\mu) < \exp\left(-\frac{1}{2}\left(\frac{5}{6}\right)^2 \mu\right) < n^{-1} .$$

Subsequently, the probability that the message will require more than $24e\tilde{R}$ participations in the RCW procedure before it reaches v, is smaller than $1/n$.

The presented algorithm requires the possibility to store n messages in the center node c. This happens just after finishing the gathering stage before the start of the broadcast stage. We note that it is possible to get around this requirement by interleaving the gathering and broadcasting stage. In this way no node would need more than $O(1)$ extra space. In fact, if y denotes the maximal number of simultaneous messages allowed in a receive or send buffer on any node then it is possible to modify the presented algorithm so that it solves the gossiping problem with high probability in time $O((\frac{n}{y} + r) \log n)$ where r is the radius of the network. Due to space limitation we defer more details to the full version of the paper.

3 Conclusion

We have proposed a new efficient (polynomial time) randomized schedule that performs the gossiping task with unit size messages in radio networks with high probability in optimal time $O(n \log n)$. The evident open problem regarding gossiping is whether there exists a deterministic gossiping schedule of time $O(n \log n)$ for every n-node graph G.

References

1. B. Awerbuch and D. Peleg. Sparse partitions. *Proc. 31st Symp. on Foundations of Computer Science*, 1990, pp. 503-513.
2. N. Alon, A. Bar-Noy, N. Linial and D. Peleg. A lower bound for radio broadcast. *J. Computer and System Sciences* 43, (1991), 290 - 298.
3. R. Bar-Yehuda, O. Goldreich and A. Itai. On the time complexity of broadcasting in radio networks: an exponential gap between determinism and randomization. *Proc. 5th Symp. on Principles of Distributed Computing*, 1986, 98 - 107.
4. M. Chrobak, L. Gąsieniec and W. Rytter, Fast broadcasting and gossiping in radio networks. *J. of Algorithms* 43(2), (2002), pp. 177-189.
5. I. Chlamtac and S. Kutten. On broadcasting in radio networks-problem analysis and protocol design. *IEEE Trans. on Communications* 33, (1985), pp. 1240-1246.
6. T.H. Cormen, C.E. Leiserson and R.L. Rivest. *Introduction to Algorithms*. MIT Press, 1990.
7. F. Cicalese F. Manne and Q. Xin. Faster centralized communication in radio networks. Manuscript, 2006.
8. A. Czumaj and W. Rytter. Broadcasting algorithms in radio networks with unknown topology. *Proc. 44th Symp. on Foundations of Computer Science*, 2003, pp. 492-501.
9. I. Chlamtac and O. Weinstein. The wave expansion approach to broadcasting in multihop radio networks. Proc. *Proc. INFOCOM*, 1987.
10. K. Diks, E. Kranakis and A. Pelc. The impact of knowledge on broadcasting time in radio networks. *Proc. 7th European Symp. on Algorithms*, 1999, pp. 41-52.
11. M. Elkin and G. Kortsarz. Polylogarithmic inapproximability of the radio broadcast problem. *Proc. APPROX*, 2004, LNCS 3122.

12. M. Elkin and G. Kortsarz. Improved broadcast schedule for radio networks. *Proc. 16th ACM-SIAM Symp. on Discrete Algorithms*, 2005.
13. I. Gaber and Y. Mansour. Broadcast in radio networks. *Proc. 6th ACM-SIAM Symp. on Discrete Algorithms*, 1995, pp. 577-585.
14. M. Christersson, L. Gąsieniec and A. Lingas. Gossiping with bounded size messages in ad-hoc radio networks. *Proc. 29th Int. Colloq. on Automata, Languages and Programming*, 2002, pp. 377-389.
15. L. Gąsieniec and A. Lingas. On adaptive deterministic gossiping in ad hoc radio networks, *Information Processing Letters* 2(83), 2002, pp. 89-94.
16. L. Gąsieniec, E. Kranakis, A. Pelc and Q. Xin. Deterministic M2M multicast in radio networks. *Proc. 31st Int. Colloq. on Automata, Languages and Programming*, 2004, LNCS 3142, pp. 670-682.
17. L. Gąsieniec and I. Potapov, Gossiping with unit messages in known radio networks. *Proc. 2nd IFIP Int. Conference on Theoretical Computer Science*, 2002, pp. 193-205.
18. L. Gąsieniec, I. Potapov and Q. Xin. Efficient gossiping in known radio networks. *Proc. 11th Int. Colloq. on Structural Information and Communication Complexity*, 2004, LNCS 3104, pp. 173-184.
19. L. Gąsieniec, D. Peleg and Q. Xin. Faster communication in known topology radio networks. *Proc. 24th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing*, PODC'2005, pp. 129-137.
20. L. Gąsieniec, T. Radzik and Q. Xin. Faster deterministic gossiping in ad-hoc radio networks. *Proc. 9th Scandinavian Workshop on Algorithm Theory*, 2004, LNCS 3111, pp. 397-407.
21. D. Kowalski and A. Pelc. Centralized deterministic broadcasting in undirected multi-hop radio networks. *Proc. APPROX*, 2004, LNCS 3122, pp. 171-182.
22. D. Liu and M. Prabhakaran. On randomized broadcasting and gossiping in radio networks. *Proc. 8th Int. Conf. on Computing and Combinatorics*, 2002, pp. 340-349.
23. A. Sen and M.L. Huson. A new model for scheduling packet radio networks. *Proc. 15th Joint Conf. of IEEE Computer and Communication Societies*, 1996, pp. 1116-1124.
24. P.J. Slater, E.J.Cockayne and S.T. Hedetniemi. Information dissemination in trees. *SIAM J. on Computing* 10, (1981), pp. 892–701.
25. A.N. Strahler. Hypsometric (area-altitude) analysis of erosional topology. *Bull. Geol. Soc. Amer.* 63, (1952), pp. 117–1142.
26. X.G. Viennot. A Strahler bijection between Dyck paths and planar trees. *Discrete Mathematics* 246, (2002), pp. 317–329.
27. Y. Xu. An $O(n^{1.5})$ deterministic gossiping algorithm for radio networks. *Algorithmica*, 36(1), (2003), pp. 93–96.