Hardness and approximation of Gathering in static radio networks

J.-C. Bermond*, J. Galtier†, R. Klasing‡, N. Morales*, S. Perennes*
* MASCOTTE project I3S-CNRS/INRIA/Université de Nice-Sophia Antipolis.
emails : bermond,rklasing,nmorales,speren@sophia.inria.fr
† France Telecom Research and Development
email: jerome.galtier@francetelecom.com
‡ LaBRI - Université Bordeaux 1, 351 cours de la Libération, 33405 Talence cedex, France

***Abstract*— In this paper, we address the problem of gathering information in a central node of a radio network, where interference constraints are present. We take into account the fact that, when a node transmits, it produces interference in an area bigger than the area in which its message can actually be received. The network is modeled by a graph; a node is able to transmit one unit of information to the set of vertices at distance at most d_T in the graph, but when doing so it generates interference that does not allow nodes at distance up to d_I ($d_I \geq d_T$) to listen to other transmissions. Time is synchronous and divided into time-steps in each of which a round (set of non-interfering radio transmissions) is performed. We give a general lower bound on the number of rounds required to gather on any graph, and present an algorithm working on any graph, with an approximation factor of 4. We also show that the problem of finding an optimal strategy for gathering (one that uses a minimum number of time-steps) does not admit a Fully Polynomial Time Approximation Scheme if $d_I > d_T$, unless P=NP, and in the case $d_I = d_T$ the problem is NP-hard.**

I. Introduction

A. Background and motivation

In radio networks a set of radio devices communicate by using radio transmissions which, depending on the technology used, are subject to different interference constraints (see for instance [1]–[3]). This means that only certain transmissions can be performed simultaneously, therefore the devices have to act in a cooperative manner in order to achieve an effective flow of information in the network. In this context, we study a problem proposed by France Telecom, about "how to provide Internet to villages" (see [4]).

The houses of the village are equipped with radio devices and they want to access the *rest of the world* via Internet. For that purpose they have to send (and receive) information via a gateway where there is a *central* antenna. This creates a special *many-to-one* information flow demand in which the access to the gateway must be provided. Therefore, we will consider a specific traffic pattern, similar to a single commodity flow with a distinguished node representing the gateway, called *sink* and denoted t.

Unlike in wired networks, when a node u transmits a message it does not use a resource as simple as some capacity on a link; instead it produces a signal that may prevent other transmissions to occur. The set of possible concurrent transmissions follows from a complex n-ary interference relation which properly models the idea that the noise intensity must be small enough compared to the signal intensity. In order to get tractable models, a widely used simplification consists of associating to each node a *transmission area* in which it can transmit a message and an *interference area* in which it produces a strong noise. Then, the communication from a node u to a node v is possible if v is in the transmission area of u, and no third node transmitting has v in its interference area. Note that, by doing so, we replace the n-ary relation with a binary relation: two (possible) transmissions (that we will denote calls) can be performed concurrently when they do not interfere.

This work was supported by the European project AEOLUS and is a part of the CRC CORSO with France Telecom R&D. N. Morales is funded by a CONICYT/INRIA doctoral grant.

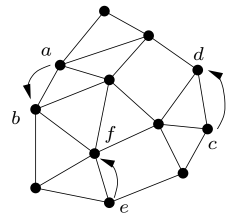

Fig. 1. Interfering/compatible calls.

B. Modeling aspects

One possible way of modeling would be to represent the houses (radio devices) as nodes in the plane with Euclidean distance (the areas of transmission and interference being disks). Here, we choose to model the network by an undirected graph $G = (V, E)$, where V is the set of devices in the network and to use as distance the distance between nodes in the graph. Firstly, it simplifies the analysis and enables us to give tractable gathering algorithms. Secondly, for some graphs like grids or hexagonal grids the distance in the graphs is a good approximation for the Euclidean distance. Finally, some nodes which are close to each other in the plane might not be able to communicate due to different reasons like obstacles, hills, social relations, security. So, there is an edge if two houses are neighbors and able to communicate.

We model the transmission area and the interference area as balls in the graph by introducing two parameters: d_T, the transmission radius and d_I, the interference radius and we suppose that $d_I \geq d_T$. The transmission area (resp. interference area) is then the ball of radius d_T (resp. d_I).

The information transmitted by a node becomes available to all the nodes that are in its transmission area if they are listening, and if they are not in the interference area of a third node. We will denote the fact that node s (like sender) is transmitting a message to node r (like receiver) by saying there is a *call* (s, r). We will say that two calls (s, r) and (s', r') with $s \neq s'$ are *compatible* if s does not interfere with r' and s' does not interfere with r. Indeed, as we are considering the gathering problem, we can assume that when a sender transmits, it does so to a unique receiver.

Figure 1 shows a set of 3 calls, which are represented by the arrows over the edges of the graph. If $d_I = d_T = 1$, all these calls are compatible. However, if $d_I = 2, d_T = 1$, vertex b is under the interference of vertex e, and vertex f is under the interference of vertices a and c. In this case, a round could either consist of one single call (e, f), or of the two calls $(a, b), (c, d)$.

Under this model, the problem raised by France Telecom consists of gathering information from each node of the network into the central node (the sink t). We will suppose that each node has to transmit an integer (≥ 0) number of units of information.

Our measure of efficiency is the time (i.e., the number of rounds) needed to achieve gathering, hence our objective is to study the *minimum time gathering problem.* Figure 2 shows an optimal gathering protocol using 18 rounds for a path with 7 vertices (each having one piece of information), with $d_T = 1$ and $d_I = 2$.

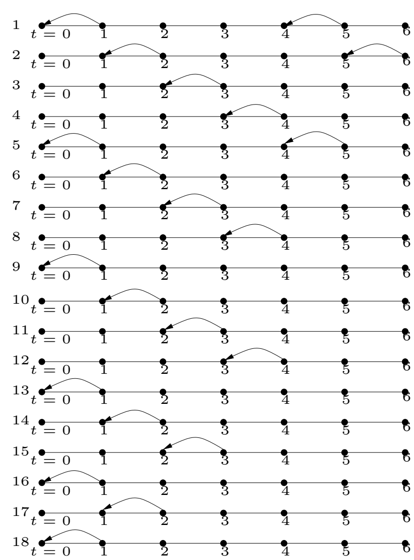

Fig. 2. A gathering protocol in the path when $d_T = 1, d_I = 2$ and every vertex has one message to send to the sink.

Note that we may as well study the converse problem (personalized broadcast) for which we need to send personalized information from the central node to each node. But since personalized broadcast and gathering are almost identical (inverse) problems, we focus on gathering. However, unlike in many other communication models, we cannot simply reverse the time or the communication steps and state that gathering and personalized broadcast are formally identical or equivalent. For example, in Figure 1 (for $d_I = 2, d_T = 1$) the 2 calls $(a, b), (c, d)$ are compatible, but the reversed calls $(b, a), (d, c)$ interfere. Despite the lack of perfect equivalence, all the results (algorithms, complexity, bounds) that we give are also valid for personalized broadcast.

C. Related work

Basic communication problems for the dissemination of information (like gathering, broadcasting, gossiping) have been widely studied in classical interconnection networks (see the book [5]). The fact that a node cannot both send and receive in the same round is known as the *half-duplex* 1*-port* model and the unit message constraint is studied for example in [6]–[8].

The broadcasting and gossiping problems in radio networks with $d_T = d_I = 1$ are studied in [9]–[11] and [12]–[14] respectively. Note that in a broadcast the same information has to be transmitted to all the other nodes and therefore flooding techniques can be used.

With respect to the gathering problem, the uniform case with $d_T = 1$ and any d_I is studied in depth for the case of the path in [15] and the two-dimensional square grid, for which optimal solutions are provided in [16], whereas in [17], the case $d_T \geq 2$ and any $d_I \geq d_T$ is studied for the same topologies. Another related model can be found in [18], where the authors study the case in which steady-state flow demands $f(u, v)$ between each pair of nodes (u, v) have to be satisfied.

D. Results & Structure of the paper

The paper is organized as follows. In Section II, the model and the gathering problem are formalized through a number of definitions. In Section III, we provide a general lower bound and a protocol (valid for any graph and any quantity of information) which allows us to prove that our protocol achieves an approximation ratio of 4, independently of d_I and d_T.

Hardness results are given in Section IV: we show that gathering in general radio networks is NP-HARD for any value of d_I, d_T and that, as soon as $d_I > d_T$, it does not admit an FPTAS which means that to find an approximation with a quality $(1 + \varepsilon)$ requires time which is not polynomial in ε^{-1}, unless P=NP.

We note that most of the proofs are omitted in this paper.

II. The model: definitions and notation

In the whole paper, we are given a graph $G = (V, E)$ with n vertices and with a distinguished vertex $t \in V$, called the *sink*, and two integers $d_I, d_T \in \mathbb{N}$, such that $d_I \geq d_T > 0$, where d_I is the *interference distance* and d_T is the *transmission distance*.

The distance between two vertices u and v is the length of the shortest path from u to v and is denoted $d_G(u, v)$. For $u \in V$ and $h \in \mathbb{N}$, we define the h-neighborhood of u as $\Gamma_G^h(u) = \{v \in V : d_G(u, v) \leq h\}$. When the context is clear we will omit the index G.

In the *gathering problem*, every node $u \in V$ has $w(u)$ pieces of information (called shortly *messages*) which have to reach the sink t; where $w(u)$ is a positive integer.

A *call* is a couple (s, r), $s, r \in V$ with $0 < d(s, r) \leq d_T$ where s is the *sender* and r the *receiver*. The call (s, r) *interferes* with the call (s', r') if $d(s, r') \leq d_I$. We say that the two calls (s, r) and (s', r') are *compatible* if they do not interfere, that is both $d(s, r') > d_I$ and $d(s', r) > d_I$.

A *round* is a set of compatible calls. During a round, a sender transmits a new message if there is one available.

A *gathering protocol* is an ordered sequence of rounds that allows to gather the information of the nodes in the sink.

We will only consider protocols that manipulate the original messages. We can show that there always exists an optimal protocol having this property.

Finally, we will often specify protocols by giving simply the sequence of rounds, without specifying which message is sent, indeed that is irrelevant as long as each vertex can forward something new.

Our objective is to find gathering protocols minimizing the number of rounds needed to gather all the messages into the sink. The minimum number of rounds will be called the *gathering number* and denoted shortly $g(G, t)$ (although it formally depends on d_T and d_I and the function w and should be denoted $g^w_{d_I,d_T}(G, t)$).

Note that in any gathering protocol there is a bottleneck near the sink as there is a critical section, where during one round only one message near the sink can move towards the sink. First, let us rule out a trivial case.

Trivial case : When V itself is a critical section, that is when any two calls in V interfere. Hence, in that case there is at most one call per round and to transmit a message of u to the sink t we need at least $\left\lceil \frac{d(u,t)}{d_T} \right\rceil$ rounds and so in that case

$g(G,t) = \sum_{u \in V} w(u) \left\lceil \frac{d(u,t)}{d_T} \right\rceil$.

In what follows, we will suppose that we are not in the trivial case. We define a *critical section of the sink* t as an h-neighborhood of t, $\Gamma^h(t)$, such that any two vertices in $\Gamma^h(t)$ cannot receive in the same round; said otherwise, there cannot exist two compatible calls (s,r) and (s',r') with both r and r' in $\Gamma^h(t)$. We define the *critical radius* $r_{\mathcal{C}} = r_{\mathcal{C}}(G,t)$ as the greatest integer h such that $\Gamma^h(t)$ is a critical section.

Example: Consider a path P_n with n vertices $0, 1, \ldots, n-1$. If $n \leq d_I + 2$, we are in the trivial case where V is a critical section. So we suppose that $n \geq d_I + 3$. The computation of $r_{\mathcal{C}}$ will depend on the position of the sink. If the sink is at one end, say vertex 0, we have $r_{\mathcal{C}} = d_I + 1$; indeed the h-neighborhood of the sink 0 consists of the vertices $0, 1, \ldots, h$ and so if $h \leq d_I + 1$ it is a critical section (as the sink is not a sender); but for $h > d_I + 1$ both 0 and $d_I + 2$ can receive, the two calls $(1,0)$ and (d_I+1, d_I+2) for example being compatible. If $n > d_I + d_T + 1$ and if the sink is $\lceil \frac{d_I + d_T}{2} \rceil$, the two calls $(0, d_T)$ and $(d_I + d_T + 1, d_I + 1)$ are compatible and therefore a simple computation shows that $r_{\mathcal{C}}(P_n, \lceil \frac{d_I+d_T}{2} \rceil) \leq \frac{d_I - d_T}{2}$. The next lemma shows that there is equality.

Lemma 1:

$$\left\lfloor \frac{d_I - d_T}{2} \right\rfloor \leq r_{\mathcal{C}}(G,t) \leq d_I + 1.$$

Proof: For the first inequality, let $(s,r), (s',r')$ be two calls such that $r, r' \in \Gamma^{\lfloor \frac{d_I - d_T}{2} \rfloor}(t)$. Then $d(s,r') \leq d(s,r) + d(r,r') \leq d_T + 2 \lfloor \frac{d_I - d_T}{2} \rfloor \leq d_T + d_I - d_T = d_I$. Therefore these calls interfere. For the second inequality, suppose that $r_{\mathcal{C}}(G,t) \geq d_I + 2$, then the two calls (s,t) with $d(s,t) = 1$ and (s',r') with $s'r'$ an edge of G and $d(s',t) = d(r',t) = d_I + 1$ are compatible. Note that the bounds are attained as shown by the example of the path. ■

III. Constant approximation algorithm for arbitrary graphs

A. Lower bounds

Recall that we suppose we are not in the trivial case. For a vertex u, let us denote $p(u,t)$ the minimum number of calls, with their receiver inside the critical section that are necessary to bring a message originated at u to the sink.

Lemma 2: If $u \in \Gamma^{rc}(t), p(u,t) = \left\lceil \frac{d(u,t)}{d_T} \right\rceil$ and if $u \notin \Gamma^{rc}(t)$: $p(u,t) = \left\lceil \frac{1 + r_{\mathcal{C}}(G,t)}{d_T} \right\rceil$.

Proof: For a given message the distance to the sink of the vertex containing this message can decrease during a call by at most d_T. Thus, if u is in the critical section, we need $\left\lceil \frac{d(u,t)}{d_T} \right\rceil$ calls (rounds) in order that one message of u reaches the sink. If u is outside the critical section, the first call with a receiver inside the critical section has its sender at distance at least $r_{\mathcal{C}}(G,t) + 1$ from the sink, and we need at least $\left\lceil \frac{1 + r_{\mathcal{C}}(G,t)}{d_T} \right\rceil$ calls (rounds) to take a message from u to the sink. ■

Corollary 1:

$$g(G,t) \geq \sum_{u \in V(G)} w(u) p(u,t)$$

Proof: For any vertex u, the minimum number of calls having their receiver in the critical section and needed to transmit a message of u is $p(u,t)w(u)$. Hence $\sum_{u \in V(G)} w(u)p(u,t)$ such calls have to be performed. By definition of the critical section, all these calls have to be done in different rounds. ■

Corollary 2: Let $\delta = \max[d(u,t), w(u) > 0$. For any $1 + r_{\mathcal{C}} + d_T \leq a \leq \delta$,

$$g(G,t) \geq \left\lceil \frac{a - (d_T + r_{\mathcal{C}})}{d_T} \right\rceil + \sum_{u \in V(G), d(u,t) \geq a} w(u)p(u,t)$$

Proof: It follows from the observation that no message from a vertex at a distance greater than a can reach $\Gamma^{rc + d_T}(t)$ before $\left\lceil \frac{a - (d_T + r_{\mathcal{C}})}{d_T} \right\rceil$ rounds and that $\Gamma^{rc}(t) \subset \Gamma^a(t)$. ■

B. A general protocol

We can derive a protocol that matches the above lower bound up to a factor of 4.

Theorem 1: There exists a 4-approximation for the gathering problem.

Sketch of proof: The idea is to pipeline the messages towards the sink according to their distance. We partition the vertices into sets B_i where B_0 contains the messages at distance in $[1, K_0 d_T]$ and B_i contains the messages at distance in $[K_0 d_T + 1 + (i-1) K d_T, K_0 d_T + i K d_T]$. ($K_0$ depends on the distribution of the messages near the sink, but in many cases is of order $\left\lceil \frac{1 + r_{\mathcal{C}}}{d_T} \right\rceil$. $K = \left\lceil \frac{d_I + d_T + 1}{d_T} \right\rceil$.)

Every K rounds, one message from B_i (if any) is transmitted to interval B_{i-1} and in general one message reaches the sink. The detailed protocol and the analysis (involved when some vertices have no messages) is given in the full version. The total number of rounds is roughly $K \sum_u w(u)$, while the lower bound is about

$K' \sum_u w(u)$ rounds, where $K' = \left\lceil \frac{1+r_C}{d_T} \right\rceil$. The ratio 4 follows from the fact that $K/K' \leq 4$. ∎

The exact value of the approximation ratio depends on d_I and d_T, with 4 being an upper bound, independent of d_I and d_T. This ratio goes to 2 when $d_T/d_I \to 0$. Furthermore, 4 is the best value that we can obtain with the above protocol.

IV. Hardness results

If $d_I = d_T$, we show that 3SAT can be reduced to gathering, and henceforth the problem is NP-Hard in this case. When $d_I > d_T$, we show that there exists no FPTAS (an FPTAS is a method that ensures a $(1+\varepsilon)$-approximation in time polynomial in ε^{-1} and the size of the problem; see [19], [20] for a definition) for gathering in general graphs. The result follows from a reduction of Minimum Vertex Coloring and Minimum Independent Set to Minimum Gathering Time. Our reductions are such that the gap introduced on gathering time is small (it is not a constant factor of the optimal) but not exponentially small.

Theorem 2: (i) When $d_I > d_T$ there exists no FPTAS for Minimum Gathering Time unless P=NP. (ii) For $d_I = d_T$ the problem is NP-hard.

V. Conclusions

In this paper, we investigated whether the radio bandwidth can be shared optimally to carry traffic over a wireless network. We proved that the problem is NP-Hard if $d_I = d_T$ and it does not admit an FPTAS when $d_I > d_T$. We proposed a constant approximation algorithm. Some complexity issues remain open: Does there exist a PTAS or a $(1+\varepsilon)$-approximation algorithm for general graphs? For particular topologies, like trees or paths, we can find an approximation close to 1 (for example in the case of paths it is possible to give approximations up to an additive constant depending on d_I, d_T [17]), but it is unclear if the problem is polynomial or not. A more practical question would be to study more dynamic cases (e.g. using online algorithms) or to derive algorithms that would not assume a global control but rely on local decisions (distributed algorithms).

References

[1] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal of Selected Areas of Communication*, vol. 18, pp. 535–547, 2000.

[2] J. Galtier, "Optimizing the IEEE 802.11b performance using slow congestion window decrease," in *Proc. 16th ITC Specialist Seminar on Performance Evaluation of Wireless and Mobile Systems*, 2004, pp. 165–176.

[3] P. Mühlethaler, *802.11 et les réseaux sans fil.* Eyrolles, 2002.

[4] P. Bertin, J.-F. Bresse, and B. Sage, "Accès haut débit en zone rurale: une solution "ad hoc"," *France Telecom R&D*, vol. 22, pp. 16–18, 2005.

[5] J. Hromkovic, R. Klasing, A. Pelc, P. Ruzicka, and W. Unger, *Dissemination of Information in Communication Networks: Broadcasting, Gossiping, Leader Election, and Fault-Tolerance.* Springer-Verlag, 2005.

[6] J.-C. Bermond, L. Gargano, and S. Perennes, "Optimal sequential gossiping by short messages," *Discrete Applied Mathematics*, vol. 86, pp. 145–155, 1998.

[7] J.-C. Bermond, L. Gargano, A. Rescigno, and U. Vaccaro, "Fast gossiping by short messages," *SIAM Journal on Computing*, vol. 27, no. 4, pp. 917–941, 1998.

[8] J.-C. Bermond, T. Kodate, and S. Perennes, "Gossiping in Cayley graphs by packets," in *Proceedings of 8-th Franco-Japanese*, ser. LNCS, vol. 1120. Brest, France: Springer, 1995, pp. 301–315.

[9] I. Chlamtac and O. Weinstein, "The wave expansion approach to broadcasting in multihop radio networks," *IEEE Transaction on Communications*, vol. 39, no. 3, pp. 426–433, 1991.

[10] M. Elkin and G. Kortsarz, "Logarithmic inapproximability of the radio broadcast problem," *Journal of Algorithms*, vol. 52, no. 1, pp. 8–25, 2004.

[11] L. Gasieniec and I. Potapov, "Gossiping with unit messages in known radio networks," in *Proceedings of the IFIP 17th World Computer Congress.* Kluwer, B.V., 2002, pp. 193–205.

[12] M. Chrobak, L. Gasieniec, and W. Rytter, "Fast broadcasting and gossiping in radio networks," *Journal of Algorithms*, vol. 43, no. 2, pp. 177–189, 2002.

[13] I. Gaber and Y. Mansour, "Centralized broadcast in multihop radio networks," *Journal of Algorithms*, vol. 46, no. 1, pp. 1–20, 2003.

[14] M. Christersson, L. Gasieniec, and A. Lingas, "Gossiping with bounded size messages in ad-hoc radio networks," in *Proceedings of 29th ICALP'02*, ser. LNCS, vol. 2380. Springer-Verlag, 2002, pp. 377–389.

[15] J.-C. Bermond, R. Corrêa, and J. Yu, "Gathering algorithms on paths under interference constraints," 2005, manuscript.

[16] J.-C. Bermond and J. Peters, "Efficient gathering in radio grids with interference," in *AlgoTel'05*, Presqu'île de Giens, May 2005, pp. 103–106.

[17] J.-C. Bermond, R. Klasing, N. Morales, and S. Pérennes, "Nearly optimal protocols for gathering in specific radio networks," 2005, manuscript.

[18] R. Klasing, N. Morales, and S. Perennes, "On the complexity of bandwidth allocation in radio networks with steady traffic demands," 2004.

[19] M. Garey and D. Johnson, *Computers and Intractability - A Guide to the Theory of NP–Completeness.* W.H. Freeman and Company, 1979.

[20] J. Hromkovic, *Algorithmics for Hard Problems: introduction to combinatorial optimization, randomization, approximation, and heuristics: 2nd edition.* Springer Verlag, 2003.