Complexity Results on Election of Multipoint Relays in Wireless Networks

Laurent Viennot

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Complexity Results on Election of Multipoint Relays in Wireless Networks

Laurent Viennot

No 3584

Décembre 1998

THÈME 1

Rapport de recherche

ISSN 0249-6399

Complexity Results on Election of Multipoint Relays in Wireless Networks

Laurent Viennot*

Thème 1 — Réseaux et systèmes
Projet HIPERCOM

Rapport de recherche n ° 3584 — Décembre 1998 — 12 pages

Abstract: The election of multipoint relays allows to decrease the cost of broadcasting in wireless networks. For each source, the fewer elements the set has, the greater the gain is. In this paper, we prove that the computation of a multipoint relay set with minimal size is NP-complete. We also make the analysis of a heuristic proposed by A. Qayyum.

Key-words: mulitpoint relay, wireless network, NP-complete, heuristic.

(Résumé : tsvp)

* Email: `Laurent.Viennot@inria.fr`

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
Téléphone : 01 39 63 55 11 - International : +33 1 39 63 55 11
Télécopie : (33) 01 39 63 53 30 - International : +33 1 39 63 53 30

Quelques analyses de complexité sur l'élection des multipoints relai dans les réseaux sans fil

Résumé : L'élection de multipoints relai permet de diminuer l'utilisation de la bande passante dans les réseaux sans fil. Plus cet ensemble est petit pour chaque source, plus le gain est intéressant. Nous prouvons ici que le calcul d'un ensemble de multipoints relai de taille minimale est NP-complet. Nous faisons de plus l'analyse d'une heurisitique proposée par A. Qayyum.

Mots-clé : multipoint relai, réseau sans fil, NP-complet, heurisitique.

1 Introduction

To support the mechanism of forwarding in broadcast or multicast in wireless networks, there must exist some forwarders who will forward the packets of one node to another node. In such networks, a node has many neighbours (*i.e.* nodes with which it can communicate directly), so that a subset of them have to be elected as forwarders to avoid redundant retransmition of the same message. Such elected nodes are called *multipoint relays* in the HIPERLAN [2] standard and in the drafts of the MANET group of the IETF [4].

To ensure broadcasting, the multipoint relays must be selected such that the message is transmitted to all the nodes within two hops from the source s. In other words, any neighbour of a neighbour of x must be the neighbour of some multipoint relay of x. Ideally, a minimum number of multipoint relays would be selected so that the communication bandwidth is occupied as less as possible.

In this paper, we show that finding an optimal multipoint relay set (*i.e.* a set with a minimal number of multipoint relays ensuring broadcasting) is an NP-complete problem. Moreover, we discuss the quality of an heuristic proposed by Qayyum [4]. Let us first give a formal definition of the problem.

2 Definitions

If x is a node of the network, we denote by $N(x)$ the set of its neighbours, *i.e.* the nodes that can communicate directly with it. $N(x)$ is called the neighbourhood of x. (Here we consider that $x \notin N(x)$.) The neighbours of x are also called one hop nodes and the neighbours of neighbours which are not neighbours of x are called two hop nodes because x has to send his messages to one forwarder to reach them. Let $NN(x)$ denote the two hops from x nodes.

If y is a neighbour of x, we also say that x covers y. Moreover, if S and T are sets of nodes, we say that S covers T iff every node in T is covered by some node in S. A set $S \subseteq N(x)$ is a multipoint relay set for x if S covers $NN(x)$, or equivalently $\cup_{y \in N(x)} N(y) - N(x) \subseteq \cup_{y \in S} N(y)$. A multipoint relay set for a node x is optimal if its number of elements is minimal among all the multipoint relay set for x. We will call this number the optimal multipoint relay number for x. See Figure 1for an example. We can now study the complexity of computing an optimal multipoint relay set.

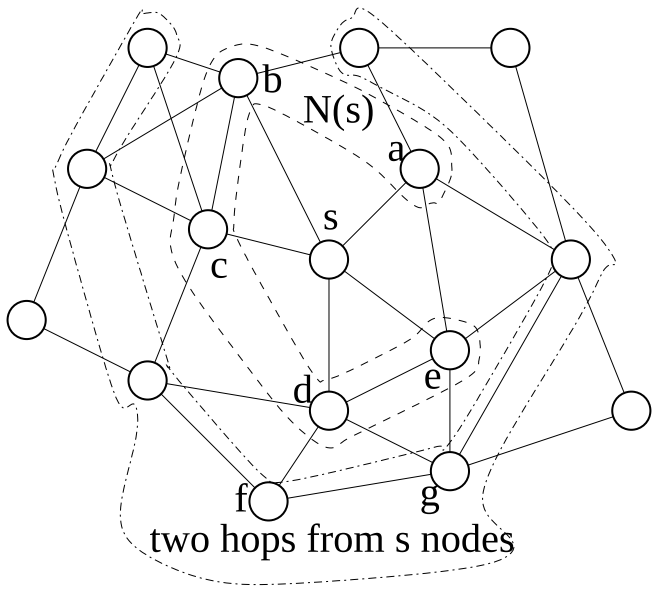

Figure 1: A network where a line between two nodes symbolizes that the two nodes can communicate directly. Imagine for example a radio network where a can reach a, b, c, d, e. $\{a, c, d, e\}$ and $\{a, b, d, e\}$ are multipoint relay set for s. $\{a, c\}$ is not because it does not cover f and g. The optimal multipoint relay number for s is 3. $\{a, c, d\}$ is an optimal multipoint relay set for s.

3 NP-completeness

We prove that the following problem is NP-complete:

Multipoint Relay: *Given a network (i.e. the set of neighbours of each node), a node x of the network and an integer k, is there a multipoint relay set for x of size less than k ?*

First of all, notice that this problem is easier than the problem of finding an optimal multipoint relay set. If an optimal set is known, simply computing its size and comparing it to k allows to answer the question. Let us now show that the Multipoint Relay Problem is NP-complete.

It is obviously in NP since taking a random set in $N(x)$, one can easily check in polynomial time if it is a multipoint realay set and if its size is less than k. To prove that it is NP-complete, we prove that the following Dominating Set Problem which is known to be NP-complete [5] can be reduced to the Multipoint Relay Problem in polynomial time:

Dominating Set Problem: *Given a graph (i.e. a set of nodes and a set of neighbours for each node) and a number k, is there a dominating set of cardinality less than k ? Where a dominating set is a set S of nodes such that any node of the graph is either in S or in the neighbourhood of some node in S.*

Let G be a graph with node set V and let $M(x)$ denote the neighbourhood of any $x \in V$. We construct a reduction as follows. Let us make a copy of V and denote with a prime the copies: x' denotes the copy of x for any $x \in V$ and S' denotes the set of copies of the elements of any set $S \subseteq V$ (V' denotes the set of all the copies). Let s be an element not in V nor in V'. Consider a network where the nodes are $\{s\} \cup V \cup V'$ and where the neighbourhood are the following:

$$\begin{aligned} N(s) &= V \\ N(x) &= \{x'\} \cup M(x)' \text{ for any } x \in V \\ N(x') &= \{x\} \cup M(x) \text{ for any } x \in V \end{aligned}$$

See Figure 2for an example.

Such a data structure can easily be computed in polynomial time. We claim that the answer to the Multipoint Relay Problem for the node s of the computed network with the integer k is valid for the Dominating Set Problem for the considered graph with the same integer k. It is sufficient to prove that any multipoint relay set S for the network is associated to a dominating set of the graph with same cardinality. S is a subset of $N(s) = V$. We show that S itself is a dominating set of the graph. Consider a node $x \in V$ and its copy x'. As S is a multipoint relay set, x' is the neighbour of some node $y \in S$. As $N(y) = \{y'\} \cup M(y)'$ by definition, we have either $x' = y'$ or $x' \in M(y)'$, or equivalently, $x = y$ or $x \in M(y)$. This means that x is in S or is the neighbour of some node in S. S is thus a dominating set and the proof is achieved.

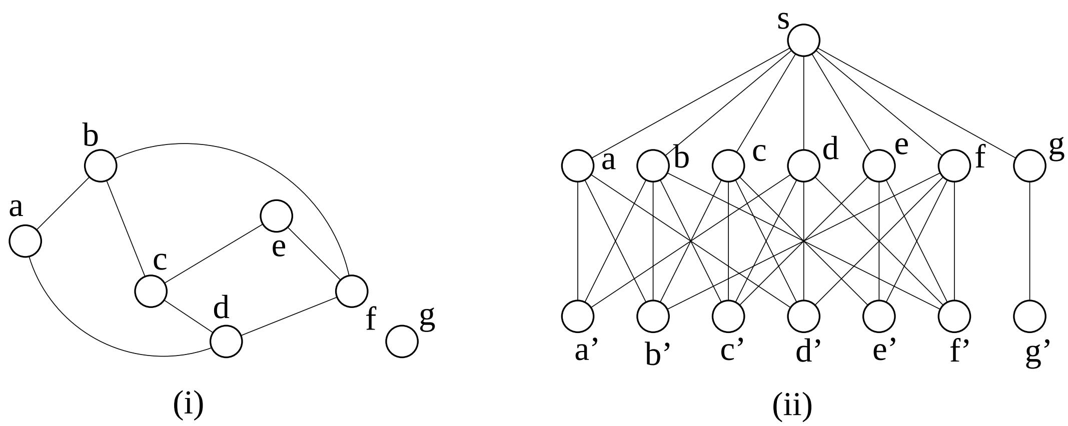

Figure 2: (i) A graph. (ii) The network obtained by the reduction. $\{b, c, g\}$ is a dominating set in (i) and a multipoint relay set for s in (ii).

4 Analysis of the Qayyum Heuristic

A draft by Qayyum [4] in the MANET group of the IETF proposes a heuristic for computing a multipoint relay set for a given node. We prove that it computes a multipoint relay set of cardinality at most $\log n$ times the optimal multipoint relay number where n is the number of nodes in the network. Moreover, we give a sequence of examples for which the heuristic computes a set which is $(\log n - 1)/2$ times bigger than the optimal. Let us first recall what the heuristic is.

4.1 The Qayyum Heuristic

Let s be a node, N its neighbours, and NN be the nodes at two hops from s. Compute a multipoint relay set S as follows [4]:

1. Start with an empty set S of multipoint relays.
2. First select as multipoint relays thoses nodes in N that are the only nighbour of some node $y \in NN$.
3. While there still exists some node in NN which is not covered by the multipoint relay set (*i.e.* no $N(x)$ with $x \in S$ contains it)

- For each $x \in N - S$, compute the number of uncovered nodes in NN that are in $N(x)$.
- Add in S a node for which this number is maximal.

Notice that the step 2 selects the nodes which must be in any multipoint relay set for s. We first show that this heuristic computes a set which is within a factor $\log n$ from the optimal and we will then show that this bound is sharp.

4.2 Upper Bound

We give a proof directly inspired from [3] which is itself inspired from a general proof by Chvátal [1]. The first proof about an analogous heuristic was given in [6].

Let S_1 be the nodes selected in stage 2 of the above algorithm and let $x_1, \ldots, x_k$ be the nodes selected in stage 3 (x_i is the ith added node). Let S^* be a solution with minimal cardinality. First notice that $S_1 \subseteq S^*$ since any node in S_1 is the only neighbour of some node in NN. We will show that $|S - S_1| \leq \log n |S^* - S_1|$ which implies that the computed solution is within a factor $\log n$ from the optimal.

Let NN_1 be the set of nodes in NN that are neighbours of some node in S_1. We set $NN' = NN - NN_1$, $S' = S - S_1$, $S^{*\prime} = S^* - S_1$ and $N'(x) = N(x) \cap NN'$ for each node $x \in N$. We associate a cost c_y to each node $y \in NN'$. For each x_i chosen by the algorithm, a unit cost is equally divided among the nodes newly covered in NN. More formally: if x_i is the first neighbour of y added in S by the algorithm, then we set:

$$c_y = \frac{1}{\left|N'(x_i) - \cup_{j=1}^{i-1} N'(x_j)\right|}$$

The costs are linked with the cardinality of the computed solution in the following way:

$$|S'| = \sum_{y \in NN'} c_y$$

We are going to show that for any node z in $S^{*\prime}$, we have:

$$\sum_{y \in N'(z)} c_y \leq \log |N'(z)| \tag{1}$$

Notice first that this implies immediatly the result. Any node $y \in NN'$ is the neighbour of some $x \in S^{*\prime}$ (remember that no node in S_1 is a neighbour of y by definition). We can thus deduce:

$$\begin{aligned}
|S'| &= \sum_{y \in NN'} c_y \\
&\leq \sum_{z \in S^{*\prime}} \sum_{y \in N'(z)} c_y \\
&\leq \sum_{z \in S^{*\prime}} \log |N'(z)| \\
&\leq |S^{*\prime}| \log n
\end{aligned}$$

We still have to prove Inequation 1 to conclude. Let z be a node in $S^{*\prime}$ and let

$$u_i = \left|N'(z) - \cup_{j=1}^{i} N'(x_j)\right|, \text{ for each } 0 \leq i \leq k \ (u_0 = |N'(z)|)$$

be the number of neighbours of z in NN' which are still not covered after the choice of $x_1, \ldots, x_i$. Let l be the first index such that $u_l = 0$. When x_i is chosen, $u_{i-1} - u_i$ neighbours of z are then covered. We can thus deduce:

$$\sum_{y \in N'(z)} c_y = \sum_{i=1}^{l} (u_{i-1} - u_i) \frac{1}{\left|N'(x_i) - \cup_{j=1}^{i-1} N'(x_j)\right|}$$

We then notice that the choice of x_i by the algorithm implies:

$$\left|N'(x_i) - \cup_{j=1}^{i-1} N'(x_j)\right| \geq \left|N'(z) - \cup_{j=1}^{i-1} N'(x_j)\right| = u_{i-1}$$

This implies:

$$\begin{aligned}
\sum_{y \in N'(z)} c_y &\le \sum_{i=1}^{l} (u_{i-1} - u_i) \frac{1}{u_{i-1}} \\
&\le \int_{u_l}^{u_0} \frac{dt}{t} \\
&\le \log u_0 \\
&\le \log |N'(z)| \\
&\le \log n
\end{aligned}$$

The upper bound on the approximation factor follows. Notice that we can get a sharper bound on the approximation factor: it is bounded by $\log \Delta$ where Δ is the maximum number of two hop nodes a one hop node may cover.

4.3 Lower Bound

We now give a sequence $(L_k),\ k \ge 0$ of examples for which the algorithm gives approximation factors greater than $\frac{\log n - 1}{2}$. This will prove that the algorithm cannot achieve an approximation factor better than $O(\log n)$ in some cases. The upper bound is thus sharp.

L_k is defined as follows. Its nodes are a source s, $k+3$ one hop nodes $a, b, c_0, \ldots, c_k$ and $2(k+1)$ sets $N\!N_0^a, \ldots, N\!N_0^a$ and $N\!N_0^b, \ldots, N\!N_0^b$ of two hop nodes (all this sets are disjoints). Each $N\!N_i^z$ with $0 \le i \le k$ and $z \in \{a, b\}$ is made of 2^i elements. The neighbours of each node are defined as indicated bellow.

$$
\begin{aligned}
N(s) &= \{a, b, c_0, \ldots, c_k\} \\
N(a) &= \{s\} \cup N\!N_0^a \cup \cdots \cup N\!N_k^a \\
N(b) &= \{s\} \cup N\!N_0^b \cup \cdots \cup N\!N_k^b \\
\text{for } 0 \leq i \leq k,\ N(c_i) &= \{s\} \cup N\!N_i^a \cup N\!N_i^b \\
\text{for } 0 \leq i \leq k \text{ and } x \in N\!N_i^a,\ N(x) &= \{a, c_i\} \\
\text{for } 0 \leq i \leq k \text{ and } x \in N\!N_i^b,\ N(x) &= \{b, c_i\}
\end{aligned}
$$

Figure 3gives an illustration of L_3.

There is only one optimal multipoint relay set for s in L_k: it is $\{a, b\}$. We show that the algorithm of Qayyum returns $\{c_k, \ldots, c_0\}$ as a solution, thus yielding an approximation factor greater than $\frac{k+1}{2}$.

First notice that all the two hop nodes have more than one one hop neighbour. Thus Stage 2 of the algorithm does not select any node. We show by recursion on i that the algorithm selects $c_k, \ldots, c_i$ in that order. This is obviously true for $i = k+1$ (no node selected). Suppose it is true for $i \geq 1$, and the algorithm has chosen $S = \{c_k, \ldots, c_i\}$. The uncovered two hop nodes are those in $N\!N_{i-1}^a, \ldots, N\!N_0^a$ and $N\!N_{i-1}^b, \ldots, N\!N_0^b$. Let us count the number $d_i(x)$ of uncovered nodes covered by each one hop node x not in S:

$$
\begin{aligned}
d_i(a) &= |N\!N_{i-1}^a| + \cdots + |N\!N_0^a| = 2^{i-1} + \cdots + 2^0 = 2^i - 1 \\
d_i(b) &= |N\!N_{i-1}^b| + \cdots + |N\!N_0^b| = 2^{i-1} + \cdots + 2^0 = 2^i - 1 \\
\text{for } i > j \geq 0,\ d_i(c_j) &= |N\!N_j^a| + |N\!N_j^b| = 2^{j+1}
\end{aligned}
$$

The algorithm thus chooses c_{i-1} and the property is still true for $i-1$. It is true for any $i \geq 0$ by recursion. The case $i = 0$ is what we wanted to show.

We have shown that the approximation factor is greater than $\frac{k+1}{2}$. Let us express this bound in the number n of two hop nodes from s in L_k or the

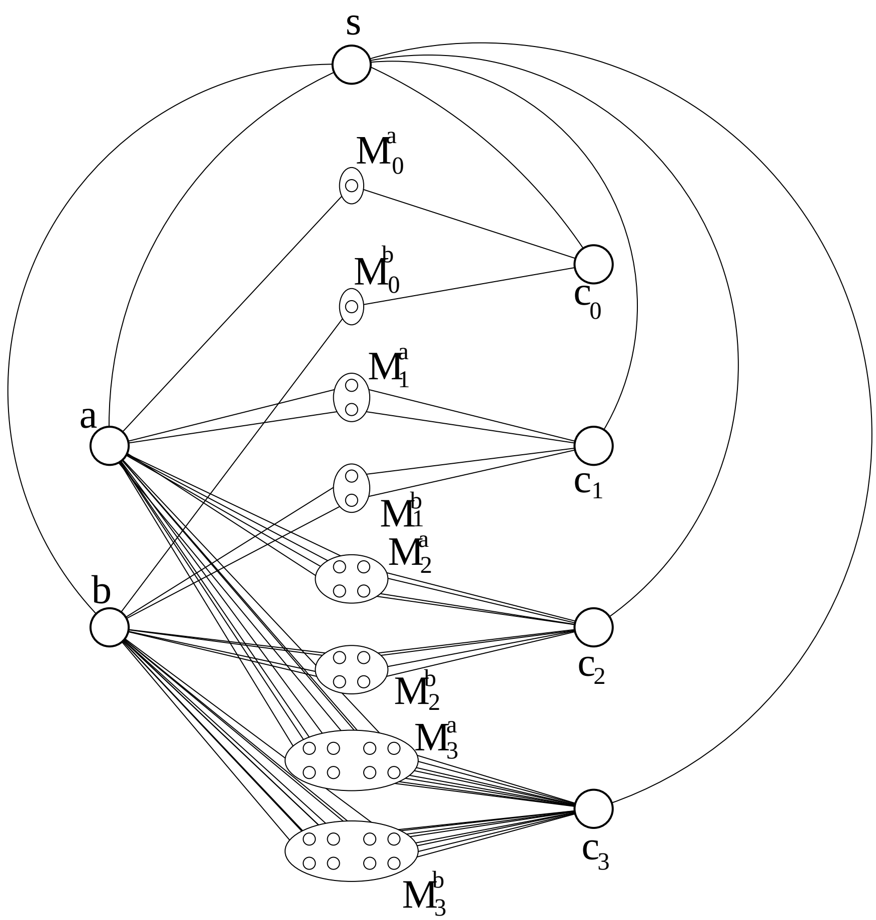

Figure 3: L_3. The Qayyum heuristic will produce $\{c_0, c_1, c_2, c_3\}$ rather than $\{a, b\}$ on L_3.

maximum degree Δ. We have:

$$n = |NN_k^a| + \cdots + |NN_0^a| + |NN_k^b| + \cdots + |NN_0^b| = 2^{k+2} - 2$$
$$\Delta = d_{k+1}(c_k) = 2^{k+1}$$

We have thus proved that the approximation factor is greater or equal than $\frac{k+1}{2} = \frac{\log(n+2)-1}{2} = \frac{\log \Delta}{2}$.

It is possible to generalize the above sequence of examples to get approximation factors close to $\log n$ by increasing the size of the optimal multipoint relay set (*i.e.* by taking α nodes instead of a and b and cutting the neighbourhood of each c_i in α sets instead of 2).

References

[1] V. Chvátal. A greedy heuristic for the set-covering problem. *Mathematics of Operation Research*, 4(3):233–235, 1979.

[2] ETSI STC-RES10 Committee. Radio equipment and systems; high performance radio local area network type 1, functional specifications. ETS 300-652, December 1995.

[3] T. Cormen, C. Leiserson, and R. Rivest. *Introduction to Algorithms*, chapter 37.3. MIT Press, 1990.

[4] M. S. Corson, S. Papademetriou, P. Papadopoulos, V. Park, and A. Qayyum. An internet manet encapsulation protocol (imep) specification. Internet Draft, draft-ietf-manet-imep-spec-01.txt, August 21 1998.

[5] M. R. Garey and D. S. Johnson. *Computers and intractability, a guide to the theory of NP-completeness.* W.H. Freeman, 1979.

[6] S. Sahni and T. Gonzales. P-complete approximation problems. *Journal of the ACM*, 23:555–565, 1976.

Unit´e de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unit´e de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unit´e de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unit´e de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unit´e de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399