# HAL
## archives-ouvertes.fr

# Computing connected dominated sets with multipoint relays

Cédric Adjih, Philippe Jacquet, Laurent Viennot

## ▶ To cite this version:

Cédric Adjih, Philippe Jacquet, Laurent Viennot. Computing connected dominated sets with multipoint relays. [Research Report] RR-4597, INRIA. 2002. inria-00071988

HAL Id: inria-00071988
https://hal.inria.fr/inria-00071988

Submitted on 23 May 2006

# *INRIA*

# *Computing connected dominated sets with multipoint relays*

Cedric Adjih  — Philippe Jacquet  — Laurent Viennot

**N° 4597**

Octobre 2002

THÈME 1

*Rapport de recherche*

# Computing connected dominated sets with multipoint relays

Cedric Adjih[*] , Philippe Jacquet[*] , Laurent Viennot[*]

**Abstract:** Multipoint relays offer an optimized way of flooding packets in a radio network. However, this technique requires the last hop knowledge: to decide wether or not a flooding packet is retransmitted, a node needs to know from which node the packet was received. When considering broadcasting at IP level, this information may be difficult to obtain. We thus propose a scheme for computing an optimized connected dominating set from multipoint relays. This set allows to efficiently broadcast packets without the last hop information with performances close to multipoint relay flooding.

**Key-words:**  multipoint relays, connected dominating set, ad hoc network

* INRIA Rocquencourt, Domaine de Voluceau, 78153 Le Chesnay, France, `{Cedric.Adjih, Philippe.Jacquet, Laurent.Viennot}@inria.fr`

# Calcul d'ensemble dominants connexes par les relais multipoints

**Résumé :** Ce papier présente un algorithme pour calculer un ensemble dominant connexe d'un graphe à partir d'ensembles de relais multipoint. Cet ensemble peut être utilisé pour diffuser efficacement une information à tous les nœuds d'un réseau radio.

**Mots-clés :** relais multipoint, diffusion, réseau ad-hoc, ensemble connexe dominant

# 1   Introduction

Ad hoc networks [2] offer new paradigms for routing. Unicast routing has been widely studied. On the other hand, broadcasting in an ad hoc network is still an open issue. The main focus here will be on optimizing the bandwidth utilization for broadcasting a packet. Wireless interfaces have an inherent capacity for broadcasting, *i.e.* with one emission, a node can reach all the nearby nodes. However, it is hard in practice to benefit from this capacity in order to avoid redundant retransmissions.

Another problem is to avoid loops when broadcasting. We will assume in the following that some mechanism allows a node to know if it has already received a packet or not. This can usually be achieved by using sequence numbers or computing signatures of packets.

The simplest way of broadcasting a packet to all nodes in an ad hoc network is pure flooding.

> **Pure flooding forwarding rule:** a node retransmits the packet only once after having received the packet.

Clearly all nodes reachable from the source will receive the packet. However, every node will retransmit the packet when it is possible, as we will see, to greatly reduce the number of retransmissions.

Several techniques have been proposed to optimize flooding with two hop neighborhood knowledge (*i.e.* a node knows its neighbors and the neighbors of its neighbors). Link state routing protocols usually provide this information. It is usually obtained by regularly emitting so-called Hello packets containing lists of neighbors.

From graph theory point of view, the set of nodes that will retransmit a given broadcast packet (including the source) must form a connected dominating set. A dominating set is a set of nodes such that any node in the network is neighbor of some element of the set. It is connected if the subgraph formed by this set is connected. The connectedness of the dominating set insures that all nodes of the connected dominating set will receive the packet and will thus be able to retransmit it. The dominating set property insures that all nodes will receive the packet (assuming no transmission error).

Computing a minimum size connected dominating set is NP hard. Moreover, the selection of the connected dominating set must be distributed. Based on two hop neighborhood knowledge, a node has to decide wether or not it is in the dominating set. Some heuristics based on neighborhood inclusions have been proposed [11, 4, 9, 5, 10]. If a connected dominating set has been elected, then the forwarding rule becomes:

> **Connected dominating set forwarding rule:** a node retransmits the packet only once after having received the packet and if it belongs to the connected dominating set.

On the other hand, the multipoint relay technique has been proposed to optimize flooding when the last hop information is given [6, 1, 7]. The idea behind this technique is to compute

some kind of local dominating sets. Each node computes a multipoint relay set with the following properties:

- the multipoint relay set is included in the neighborhood of the node, the element of the multipoint relay set are called *multipoint relays* (or MPR for short) of the node,

- each two hop neighbor of the node has a neighbor in the multipoint relay set (we say that some multipoint relay *covers* the two hop neighbor).

The multipoint relay set plus the node forms a dominating set of the two hop neighborhood of the node.

More formally, let $\mathcal{N}(E)$ denote the set of all nodes that are in a given set $E$ or have a neighbor in $E$. We say that $E$ *covers* a set $F$ when $F \subset \mathcal{N}(E)$. Let $\mathcal{N}_1(E) = \mathcal{N}(E) - E$ denote the nodes at distance 1 from $E$ and $\mathcal{N}_2(E) = \mathcal{N}(\mathcal{N}(E)) - \mathcal{N}(E)$ denote the nodes at distance 2 from $E$. When $E = \{x\}$ contains only one node, $\mathcal{N}_1(\{x\})$ is the neighborhood of $x$ and $\mathcal{N}_2(\{x\})$ is the two hop neighborhood of $x$. A MPR set $M$ of a node $x$ is thus any subset of $\mathcal{N}_1(\{x\})$ such that $\mathcal{N}_2(\{x\}) \subset \mathcal{N}(M)$. We say that $M$ is a subset of neighbors that covers the two hop neighborhood of $x$ or equivalently $M$ which is a dominating set of the subgraph induced by $\mathcal{N}(\mathcal{N}(\{x\}))$.

We call multipoint relay selector of a node $A$ a node which has selected node $A$ as multipoint relay. It can then be proven that the following forwarding rule allows to reach all the nodes in the network (assuming no transmission error):

> **Multipoint relay flooding forwarding rule:** a node retransmits the packet only once if it has received the packet the first time from a multipoint relay selector.

The smallest the multipoint relay sets are, the fewer retransmissions will occur. It is NP hard to compute a multipoint relay set with minimum size [7] but there exists good heuristics based on preferring neighbors with large degree as multipoint relays.

In this paper, we propose algorithms for computing MPR sets and a connected dominating set based on multipoint relays. The only knowledge assumed for a given node to determine its MPR set is its two hop neighborhood, and to determine the dominating set, the list of neighbors that have selected the node as multipoint relay (such neighbors are called multipoint relay selectors). These information can be contained in hello packets that nodes periodically broadcast to their neighbor in orderto monitor links validity. The algorithms does not need any distributed knowledge of the global network topology. These assumptions make the algorithm very attractive for mobile ad hoc networks since it needs just local updates at each detected topology change.

## 2   Computing MPR sets

In practice the following greedy algorithm works very well.

We start with an empty MPR set.

- Step 1: Find the two-hop neighbors that are neighbor to a single neighbor of $A$'s. Put in the MPR set these latter neighbors (called single connector neighbors).

- Step 2: Add in the MPR set the neighbor node that covers the largest number of two-hop neighbors of $A$ that are not yet covered by the current MPR set. Repeat this step until all two-hop neighborhood is covered.

It has been shown [7] that the greedy algorithm provides an MPR set whose size is at most $\log m$ larger than the optimal MPR set, where $m$ is the size of the neighborhood. The number of step required by the greedy algorithm is $O(m^2)$.

There is a suboptimal algorithm called the min-id algorithm that consists into selecting the nodes in the increasing order of their ID's (or any arbitrary increasing order).

Start with an empty MPR set.

- Check the neighbor nodes in the increasing order of their identifiers. If the current node covers a two-hop neighbor which was not yet covered by the current MPR set, then add the current node to the MPR set.

Although the min-id algorithm is far from optimal (there is no bound with the optimal MPR set size), it has the advantage to work in $O(m)$ step. It also has the advantage that the node can detect by itself whether or not it belongs to the MPR set of a neighbor [8], but in this case the algorithm requires $O(m^4)$. We call the set of nodes that select a given node $A$ as MPR set, the MPR selector set of node $A$. The reverse MPR selection algorithm for node $A$ operates as follow:

start with an empty MPR selector set.

- Check for each pair of neighbor nodes $B$ and $C$. If node $A$ has the smallest identifier of all the nodes that are neighbor to both $B$ and $C$, then $B$ and $C$ are added to the MPR selector set.

# 3   Connected dominating set algorithm

The first dominating set one can imagine is the set of all MPR nodes. A node belongs to the dominating set if it is MPR of at least one node in the network. Unfortunately, as we will see in the performance evaluation section this dominating set is not really optimal since it contains too many nodes. In order to get a much smaller dominating set we must eliminate some of the MPR nodes. We introduce an algorithm for this purpose.

To allow coordination between nodes, our algorithm requires the knowledge of a total order of the nodes. One can possibly use the smallest IP address of a node as an ID, and we will say that a node is smaller than an other if it has a smaller ID. Any total order on which all nodes agree can be used.

A node then decides that it is in the connected dominating set if:

- Rule 1: the node has a smaller ID than all its neighbors,

- Rule 2: or it is multipoint relay of its neighbor with the smallest ID.

# 4    Proof of correctness

Let us call $D$ the set of all nodes that have decided to be in the connected dominating set. The smallest node of the network is clearly in $D$ by rule 1. Let $C$ be the connected component of the smallest node in the subgraph induced by $D$. We are going to show that $C$ is a dominating set for the network (assuming that the network is connected of course). This will prove in particular that any node in $D$ has a neighbor in $C$, implying that $C$ indeed equals $D$. This will thus prove that $D$ is connected on the one hand and that $D$ is a dominating set on the other hand.

Assume by contradiction that $C$ is not a dominating set. There must exist some nodes that are not in $\mathcal{N}(C)$. Consider the set $V$ of nodes connecting some node in $C$ to some node in $\overline{\mathcal{N}(C)}$, the complementary of $\mathcal{N}(C)$. $V$ is the set of nodes which have at least one neighbor in $C$ and at least one neighbor in $\overline{\mathcal{N}(C)}$. As the network is connected, our assumption implies that $V$ is not empty. Notice that $V \cap C = \emptyset$ by construction. We now consider the smallest node $m$ in $\mathcal{N}(V)$.

- Either $m$ is in $\overline{\mathcal{N}(C)}$. As $m \in \mathcal{N}(V)$, there exists a neighbor $v$ of $m$ in $V$. Let $c$ be a neighbor of $v$ in $C$. Consider the multipoint relay set of $m$: as $c$ is a two hop neighbor of $m$, there must exist some multipoint relay $r$ of $m$ which is a neighbor of $c$. Notice that $r$ must be in $V$. As the smallest neighbor of $r$ is $m$, $r$ should have elected itself in $D$ by rule 2, contradicting $V \cap C = \emptyset$.

- Either $m$ is in $\mathcal{N}(C)$ which can be partitioned in $V$ and $\mathcal{N}(C) - V$:

  - If $m$ is in $V$, $m$ should have elected itself as being in $D$ since all its neighbors are greater. This is again a contradiction with $V \cap C = \emptyset$.

  - On the other hand if $m$ is not in $V$, it cannot have any neighbor in $\overline{\mathcal{N}(C)}$. Let $v$ be a neighbor of $m$ in $V$, and let $x$ be a neighbor of $v$ in $\overline{\mathcal{N}(C)}$. As $m$ has no neighbor in $\overline{\mathcal{N}(C)}$, $x$ is not a neighbor of $m$ and some multipoint relay $r$ of $m$ must be connected to $x$. As $m$ cannot have any neighbor in $\overline{\mathcal{N}(C)}$, $r$ must be in $\mathcal{N}(C)$. As $r$ has a neighbor in $\overline{\mathcal{N}(C)}$, it is in $V$. The smallest neighbor of $r$ is thus $m$ implying that $r$ should be in $D$ by rule 2, contradicting again $V \cap C = \emptyset$.

In all cases we get a contradiction. $C$ thus have to be a dominating set. As mentioned before, this implies that $D = C$ is a connected dominating set.

# 5    Comparison with multipoint relay flooding

When a multipoint relay flooding is carried out, the set of nodes that retransmit the packet (including the source) forms a connected dominating set. We recall the proof of that fact for the paper to be self contained.

The nodes that retransmit are clearly connected. We are going to show that any node at distance $d$ from the source is neighbor of some re-transmitter at distance $d - 1$ from the

source by recurrence on $d$. The nodes at distance 1 from the source are neighbors of the source. Assume $d > 1$ and the property verified for $d' < d$. Consider a node $x$ at distance $d$ from the source. Let $V$ be the neighbors of $x$ at distance $d-1$ from the source ($V$ is clearly not empty). Our recurrence hypothesis implies that any node $y \in V$ has some re-transmitter neighbor at distance $d-2$ from the source. Among all those re-transmitters let $z$ be the one that emitted first. Some multipoint relay $m$ of $z$ must be a neighbor of $x$ since $x$ is a two hop neighbor of $z$. $m$ is thus in $V$. As $z$ emitted first among the neighbors $m$, and $m$ is multipoint relay of $z$, $m$ must have emitted according to the multipoint relay forwarding rule. $m$ is thus a re-transmitter which a neighbor of $x$. This achieves our recurrence proof.

We are going to show that the connected dominating set obtained by a multipoint relay flooding can be obtained by our algorithm if the right order on the nodes is taken *a posteriori*.

Consider one multipoint relay from a source $s$ and let $R$ denote the set of re-transmitters including $s$. We define a total order on nodes as follows:

- $s$ is smaller than any other node,

- any node in $R$ is smaller than any node not in $R$,

- for $x, y \in R$, $x < y$ if and only if $x$ has emitted before $y$,

- the nodes not in $R$ are ordered arbitrarily.

Let us show that our algorithm computes $R$ as connected dominating set. Let $D$ denote the set computed by our algorithm. We have to show $D = R$. First notice that $s \in D$ by rule 1. Consider any node $x \in R - \{s\}$. Since $x$ has emitted, the first time it received the packet from some neighbor $y$, he was multipoint relay from $y$ and decided to retransmit. $y$ is clearly the smallest neighbor of $x$, implying $x \in D$ by rule 2. This proves $R \subset D$.

Now consider some node $x \in D$, and let $m$ be the smallest node in $\mathcal{N}(\{x\})$. As proven before, $R$ is a dominating set. $m$ is thus in $R$. Either $m = x$ proving $x \in R$ or $m \neq x$, implying that $x$ was elected by rule 2: $x$ is a multipoint relay of $m$. As $m$ was the first neighbor of $x$ re-transmitting the packet, $x$ must have decided to retransmit, and again $x \in R$. This proves $D \subset R$, achieving the proof of $D = R$.

The multipoint relay flooding forwarding rule can be seen as a particular case of our connected dominating set algorithm where the only node elected by rule 1 is the source. This is thus not surprising that multipoint relay flooding gives better results when nodes are ordered arbitrarily since a node is elected by rule 1 with probability $1/(\delta)$ (where $\delta$ is its degree). The average number of nodes elected by rule 1, is then approximately $N/\Delta$ when all nodes have a degree close to the average degree $\Delta$.

Based on the above remark, we can propose a better order than random for computing the connected dominating set when all nodes know their distance from the source:

- for two nodes $x, x'$ at distances $d, d'$ from the source, we have $x < x'$ if and only if $d < d'$ or $d = d'$ and the ID of $x$ is smaller than the ID of $x'$.

Such an ordering would result possibly in a smaller connected dominating set since the only node elected by rule 1 would be the source. Here the forwarding rule would use the source of the broadcast packet to decide wether the node is or not in the connected dominating set.

# 6   Performance simulations

We take as model network of 2,000 nodes randomly and uniformly in a square field $L \times L$. A pair of node is in range of each other when the distance is smaller than a fixed radius $R$. We call density $m$ the quantity $2000\frac{\pi R^2}{L^2}$, that is the average neighborhood size of a node in the center of the square field. We run different simulation with different value of $R$ in order to let $m$ vary. The average neighborhood size for any node is smaller than $m$ since some nodes have their neighborhood cut by the border of the field. Figure 1 shows the actual average neighbor size with respect to density $m$. The border effect becomes apparent when $m > 100$.

In figure 2 we display the average number of MPR per node using greedy algorithm and min-id algorithm. It turns out that the number of MPR nodes is much smaller than the average neighborhood size, illustrating the performance of the MPR selection. Anyhow the greedy algorithm is more performant than the min-id algorithm since it eliminates non optimal nodes that would be inserted in the min-id algorithm.

Figure 3 shows the average fraction of nodes in the network that retransmit in an MPR flooding. Notice that the performance ratio between greedy algorithm and min-id algorithm is similar than the ratio between average MPR set sizes.

Figure 4 shows the total number of nodes in the network that are MPR of at least one node in the network. Notice the surprising results that greedy algorithm and min-id algorithm provides very similar results although the average numbers of MPR per node greatly differ. The reason is that the min-id algorithm counterbalances its low performance by forcing nodes to be MPRs of several nodes. However the dominating set made by all the MPRs is not performant (too many nodes).

Figure 5 shows the fraction of nodes in the network belonging to the MPR dominating set with rules 1 and 2. Notice that the MPR dominating set greatly outperforms the dominating set made of all MPRs. The greedy algorithm outperforms the min-id algorithm within the same ratio as the average MPR set size ratio.

Figure 6 compares the fraction of nodes belonging to the MPR dominating set with the fraction of nodes implied in an MPR flooding (greedy algorithm in both case). Notice that the MPR flooding outperforms the MPR dominating set but not significantly.

Figure 7 shows the fraction of nodes in the network which belong to the MPR dominating set with respect to rule 2. Notice that this fraction is similar to the fraction of node implied in an MPR flooding when the density is important.
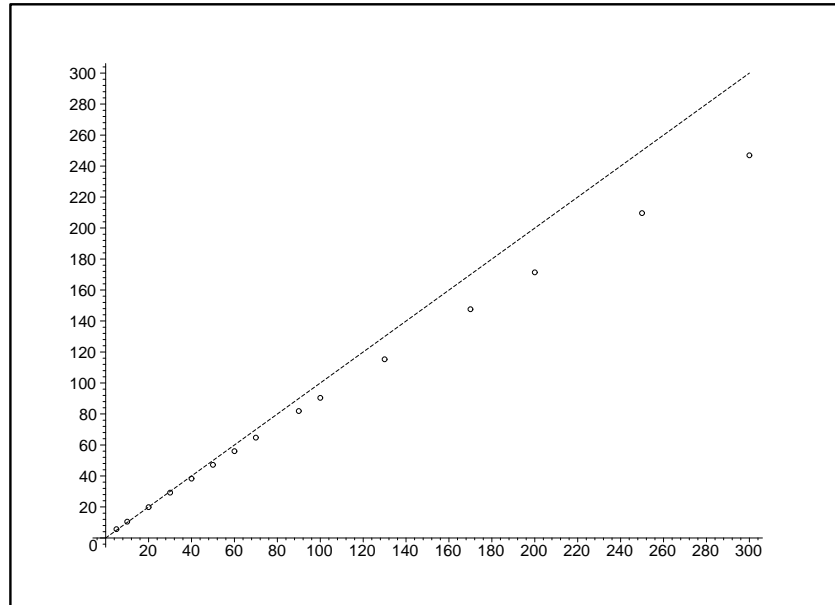
Figure 1: average neighbor size (circles) and average neighborhood size of a central node (dahed) versus density.
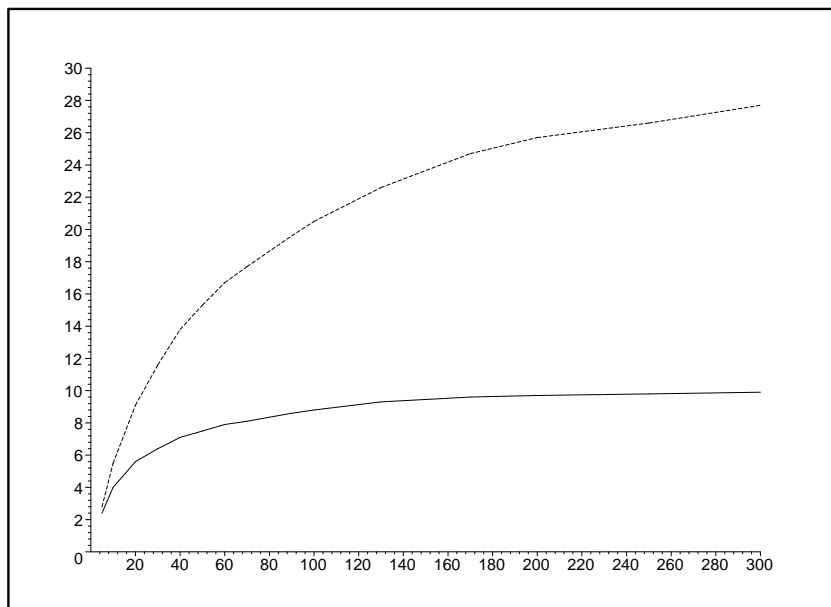
Figure 2: average MPR size with greedy algorithm (plain) and with in-id algorithm (dashed) versus density.
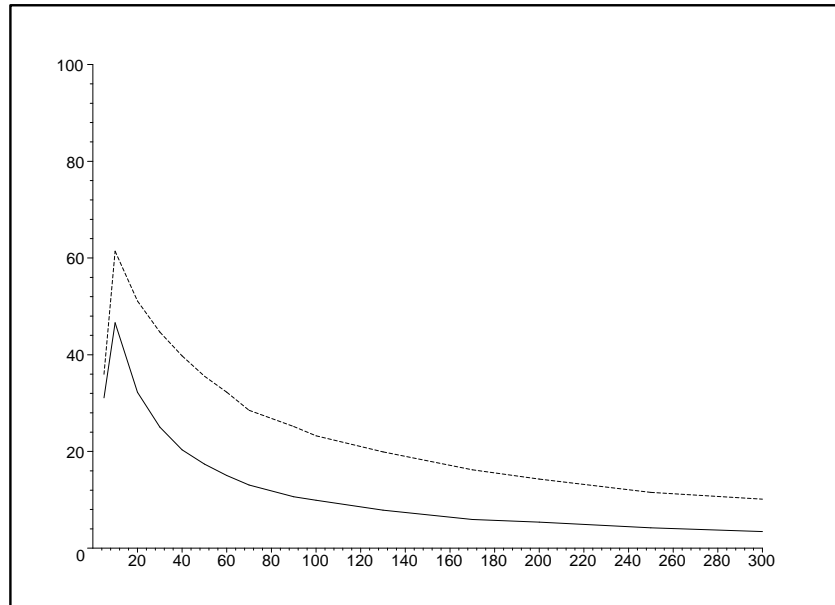
Figure 3: average fraction of nodes in the network implied in an MPR flooding, with greedy algorithm (plain) and with in-id algorithm (dashed), versus density.
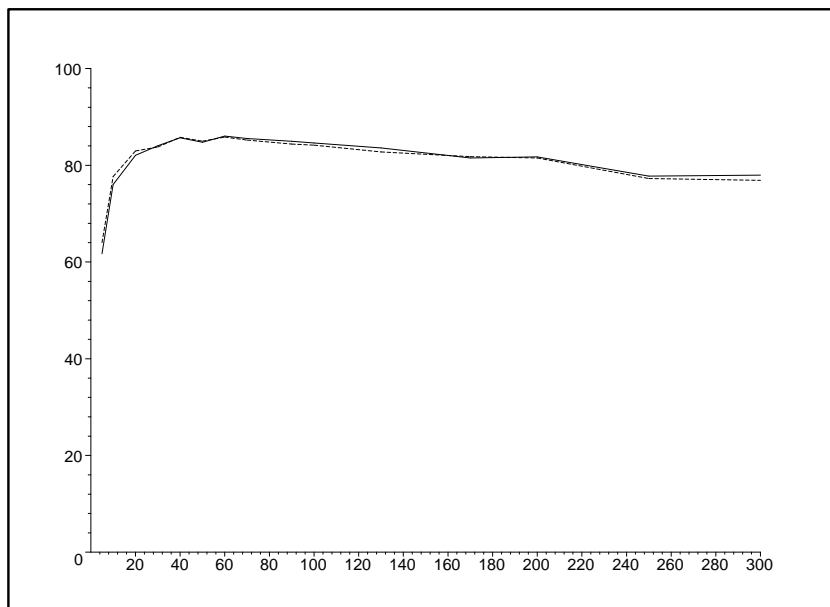
Figure 4: average fraction of nodes which are MPR of at least one node in the network, with greedy algorithm (plain) and with in-id algorithm (dashed), versus density.
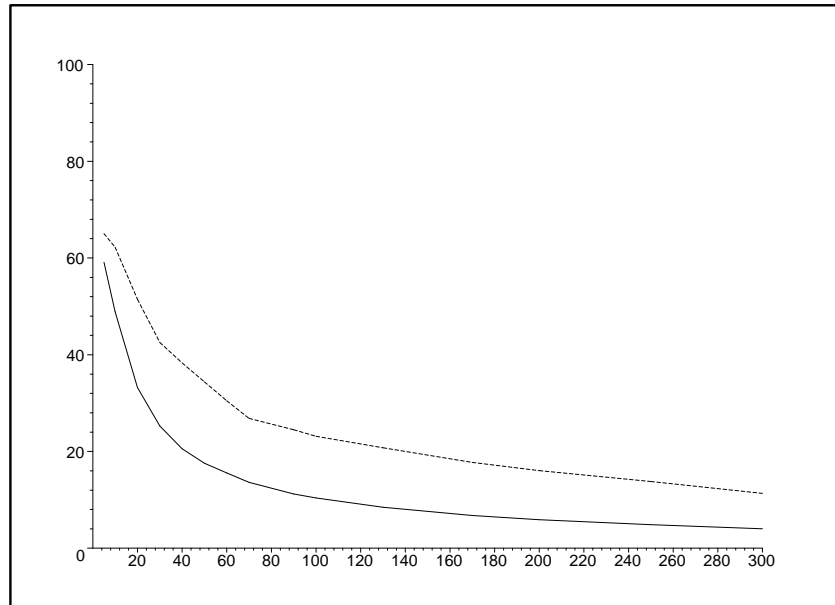
Figure 5: average fraction of nodes which are in the MPR dominating set, with greedy algorithm (plain) and with in-id algorithm (dashed), versus density.
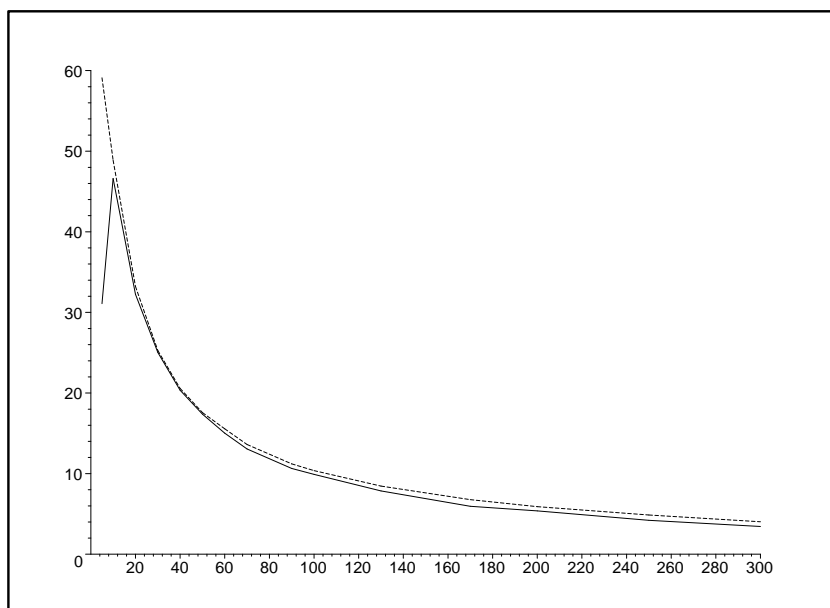
Figure 6: average fraction of nodes implied in an MPR flooding (plain) and average fraction of nodes belonging to the MPR dominating set (dashed), versus density.
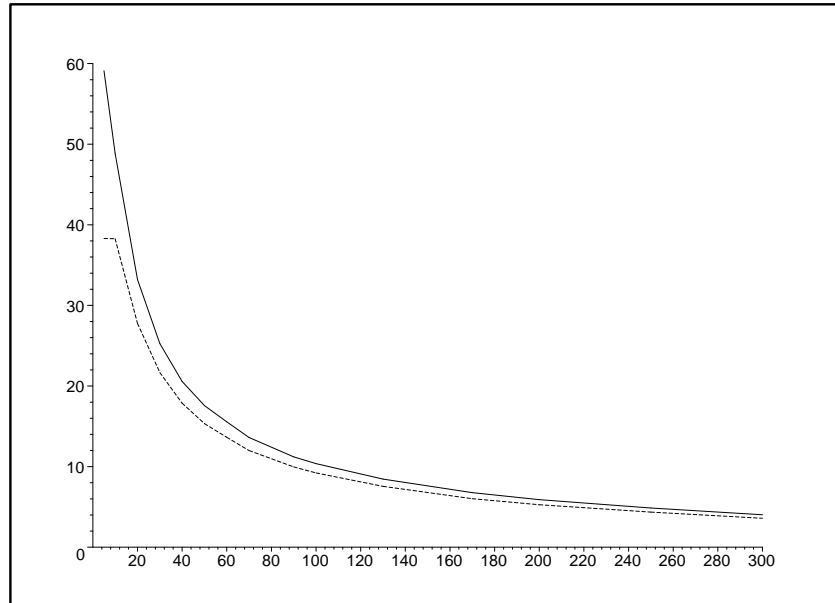
Figure 7: average fraction of nodes belonging to the MPR dominating set (plain), fraction of nodes belonging to the dominating set with rule 2 (dashed), versus density.

# 7    Conclusion

We have presented a very performant algorithm for computing a connected dominating set. Contrary to previous works on these subject the algorithm only requires two-hop topology informations and no synchronization and iterations. These assumptions make the algorithm very attractive for mobile ad hoc networks since it needs just local updates at each detected topology change. Our future work will be to make quantitative comparisons with other connected dominating set heuristics [11, 3]

# References

[1] ETSI STC-RES10 Committee. Radio equipment and systems: High performance radio local area network (hiperlan) type 1, functional specifications, June 1996. ETS 300-652.

[2] M. S. Corson and J. Macker. Mobile ad hoc networking (MANET): Routing protocol performance issues and evaluation considerations. RFC 2501, January 1999.

[3] S. Datta, I. Stojmenovic, and J. Wu. Internal node and shortcut based routing with guaranteed deliv ery in wireless networks. In *Proc. IEEE Int. Conf. on Distributed Computing and Systems*, April 16-19 2001. Phoenix, AR.

[4] S. Datta, I. Stojmenovic, and J. Wu. Internal node and shortcut based routing with guaranteed delivery in wireless networks. In *Proc. IEEE Int. Conf. on Distributed Computing and Systems*, April 16-19 2001. Phoenix, AR.

[5] Sudipto Guha and Samir Khuller. Approximation algorithms for connected dominating sets. In *European Symposium on Algorithms*, pages 179–193, 1996.

[6] Philippe Jacquet, Pascale Minet, Paul Muhlethaler, and Nicolas Rivierre. Increasing reliability in cable-free radio LANs: Low level forwarding in HIPERLAN. *Wireless Personal Communications*, 4(1):65–80, January 1997.

[7] A. Laouiti, A. Qayyum, and L. Viennot. Multipoint relaying: An efficient technique for flooding in mobile wireless networks. In *35th Annual Hawaii International Conference on System Sciences (HICSS'2001)*. IEEE Computer Society, 2001.

[8] R. Ogier, F. Templin, B. Bellur, and M. Lewis. Topology broadcast based on reverse-path forwarding. Internet Draft, draft-ietf-manet-tbrpf-05.txt, March 7 2002, Work in progress.

[9] I. Stojmenovic, M. Seddigh, and J. Zunic. Internal node based broadcasting algorithms in wireless networks. In *Proc. IEEE Hawaii Int. Conf. on System Sciences*, January 2001.

[10] P.-J. Wan, K. M. Alzoubi, and O. Frieder. Connected dominating set in wireless ad hoc networks. In *IEEE INFOCOM*, June 2002. New York.

[11] J. Wu and H. Li. On calculating connected dominating set for efficient routing in ad hoc wireless networks. In *Proc. DIAL M*, pages 7–14, August 1999. Seatle, USA.