Available online at www.sciencedirect.com
ScienceDirect
Ad Hoc Networks 5 (2007) 504–530
Ad Hoc Networks
www.elsevier.com/locate/adhoc

# Bounded-distance multi-clusterhead formation in wireless ad hoc networks

Marco Aurélio Spohn [a,*], J.J. Garcia-Luna-Aceves [b,c]

[a] *Universidade Federal de Campina Grande, Departamento de Sistemas e Computação, Campina Grande, PB, Brazil*
[b] *University of California at Santa Cruz, Computer Engineering Department, Santa Cruz, CA 95064, United States*
[c] *Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304, United States*

Received 15 June 2005; received in revised form 9 December 2005; accepted 2 January 2006
Available online 6 April 2006

**Abstract**

We present a clustering technique addressing redundancy for bounded-distance clusters, which means being able to determine the minimum number of *cluster-heads* per node, and the maximum distance from nodes to their cluster-heads. This problem is similar to computing a $(k,r)$-*dominating set*, $(k,r)$-DS, of the network. $(k,r)$-DS is defined as the problem of selecting a minimum cardinality vertex set $D$ of the network such that every vertex $u$ not in $D$ is at a distance smaller than or equal to $r$ from at least $k$ vertices in $D$. In mobile ad hoc networks (MANETs), clusters should be computed distributively, because the topology may change frequently. We present *the first centralized and distributed solutions* to the $(k,r)$-DS problem for arbitrary topologies. The centralized algorithm computes a $(k \cdot \ln \Delta)$-approximation, where $\Delta$ is the largest cardinality among all $r$-hop neighborhoods in the network. The distributed approach is extended for *clustering* applications, while the centralized is used as a lower bound for comparison purposes. Extensive simulations are used to compare the distributed solution with the centralized one. As a case study, we propose a novel multi-core multicast protocol that applies the distributed solution for the election of cores. The new protocol is compared against PUMA, one of the best performing multicast protocols for MANETS. Simulation results show that the new protocol outperforms PUMA on the context of static networks.


*Keywords:* Ad hoc networks; Clustering; Domination in graphs; Distance domination; Multiple domination; $(k,r)$-dominating sets; Multicast protocols; Multi-core multicast protocols

## 1. Introduction

In mobile ad hoc networks (MANETs), hierarchical architectures can be used to prolong the network's lifetime [1–3], attain load balancing [4], and increase network scalability [5,6]. *Clustering* is the problem of building a hierarchy among nodes [7]. The substructures that are collapsed in higher levels are called *clusters*. In each cluster, at least one node may represent the cluster, and this node is usually called a *cluster-head*. The network can then be abstracted such that any cluster-head connects to

* Corresponding author. Tel.: +1 831 459 5436.
*E-mail addresses:* maspohn@cse.ucsc.edu (M.A. Spohn), jj@cse.ucsc.edu (J.J. Garcia-Luna-Aceves).

1570-8705/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.adhoc.2006.01.005

another cluster-head whenever there is at least one node in each cluster directly connected to each other.

Clustering usually entails the computation of a *dominating set* (DS) of the network. The domination problem seeks to determine a minimum number of nodes $D$ (called *dominating nodes* or *cluster-heads*), such that any node $i$ not in $D$ is adjacent to at least one node in $D$. The computation of a DS of minimum cardinality for arbitrary graphs is known to be NP-complete [8].

A variety of conditions may be imposed on the dominating set [9]. The $(k,r)$-DS problem [10] has been defined as the problem of selecting a minimum cardinality vertex set $D$ of a graph $G=(V,E)$, such that every vertex $u$ not in $D$ is at a distance smaller than or equal to $r$ (*distance domination*) from at least $k$ (*multiple domination*) vertices in $D$. The problem of computing a $(k,r)$-DS of minimum cardinality for arbitrary graphs is also NP-complete [10].

When selecting dominating nodes, redundancy is achieved by choosing a value for the parameter $k$ greater than one. At the same time, the distance parameter $r$ allows increasing local availability by reducing the distance to the dominating nodes. Depending on the requirements, problems that require the computation of a DS can be solved by setting the two dominating parameters appropriately.

Section 2 discusses the related work. Surprisingly, to date, there is no clustering technique addressing redundancy for bounded-distance clusters, which means being able to determine the minimum number of *cluster-heads* per node, and the maximum distance from nodes to their cluster-heads.

The main contribution of this work consists of providing a framework for building flexible hierarchies for clustering with the support for fault-tolerant applications. The $(k,r)$-DS could be applied to solve a large variety of problems, of which we mention two below.

*Hierarchical routing:* By grouping clusters into super-clusters, and so on, an $m$-level hierarchical clustering [11] structure can be built. Some approaches [12] target wired networks assuming that the predefined hierarchical address of each node reflects its position within the hierarchy. During the early days of *packet radio networks* (PRNET's), hierarchical routing had been considered for reducing the routing cost and improve the performance of the network [13]. In MANETs, group mobility is usually assumed when deriving hierarchical clusters [14,15]. However, none of the existing solutions address redundancy within the hierarchical structures. Instead of having just one cluster-head representing a group of nodes, hierarchies could account for node failures by deploying multiple cluster-heads within each sub-structure (i.e., domain). Nodes within the same domain would have alternate access points when accessing nodes outside their own domain, and adjacent domains could be connected among each other through alternate paths. Furthermore, load balancing could be explored by routing packets via alternate paths connecting any pair of nodes.

*Core placement in shared-tree multicasting:* Instead of deploying just one core per multicast group (e.g., like in *core based trees* [16]), multiple cores can be selected within regions of variable radius in the network. That is, multicast members have the choice of joining multiple cores within a maximum distance. More than one core provides a certain degree of fault-tolerance, and a maximum distance to the cores can reduce the average end-to-end delay.

*Demand-driven applications in sensor networks [17]:* Multiple sinks can be distributed in a sensor network to provide some degree of fault-tolerance, and a maximum distance from nodes to sinks could support a bounded report delay. Sinks could also be organized in a multi-level hierarchy, where sinks aggregate data collected from their domain, and transmit it to a higher level sink.

We propose the first centralized and distributed solutions to the $(k,r)$-DS problem for arbitrary network topologies. The centralized solution, presented in Section 3, provides an approximation to the optimal solution, and is used as a lower bound when evaluating the performance of the distributed solution. Section 4 presents our distributed solution, which is applicable to ad hoc networks, given that it relies on information limited to the neighborhoods of nodes. Section 5 addresses the performance of our distributed solution compared to the centralized one. Section 6 presents a novel multi-core multicast protocol that applies the distributed solution for the election of cores. Section 7 concludes this work. In Appendix A, we present the analysis of both algorithms.

## 2. Related work

Clustering based on domination in graphs has been explored extensively. Chen et al [18] explore

*independent dominating sets* (IDS) for computing cluster such that cluster-heads are at least $k+1$ hops from each other. However, the use of IDS for clustering is not recommended when the topology changes, because cluster-head changes may propagate throughout the network, an effect that has been called the *chain reaction* [19]. Hence, DSs are preferred for clustering in ad hoc wireless networks.

The simplest form of dominating sets, $(1,1)$-DS, have been explored extensively for enhancing broadcasting and routing in wireless ad hoc networks [2,20–33]. The $(1,2)$-CDS (two-hop connected dominating set) has also been applied to improve the network discovery process in on-demand routing protocols [34].

*Max–Min* [35] is an election-based distributed solution for the $(1,r)$-DS problem, which takes $2r$ rounds to complete, where a round is defined as the reliable exchange of a message between a node and its neighbors. Cluster-heads are computed during the first $r$ rounds, and nodes decide which dominating nodes are going to be their cluster-heads during the subsequent $r$ rounds. The authors also show that the problem of computing the minimum $r$-hop dominating set is NP-complete for unit-disk graphs [36].

Liang and Hass [37] proposed a distributed algorithm to compute $(1,r)$-DS. The algorithm is a distributed version of *greedy set cover* (GSC), producing dominating sets with the same cardinality as the centralized solution for this problem. However, their solution requires the $2r$-hop neighborhood information.

A natural greedy solution to the $(1,r)$-DS problem has also been applied for *core* placement in multicast trees with multiple cores [38]. This solution is intended for wired networks, and requires knowledge of the entire topology.

Belding-Royer [39] proposed using one-level and multi-level clustering for scalable ad hoc hierarchical routing. One-level clusters are computed using a distributed DS algorithm, and the multi-level approach is similar to the $d$-hop DS problem, where each level-$n$ cluster-head is a cluster-head for some level-$(n-1)$ cluster-head. Some approaches (e.g., the *landmark hierarchy* [12] for wired networks) assume that the predefined hierarchical address of each node reflects its position within the hierarchy. LANMAR [15] borrows the notion of landmark for selecting a subset of nodes to keep track of logical subnets, and it is shown to improve routing performance in MANETs in which groups (subnets) are likely to move as a group. *Hierarchical state routing* [14] (HSR) is another approach that builds hierarchies assuming group mobility in MANETs. Hierarchically organized networks [40] have also been proposed for providing quality-of-service in MANETs.

Dai and Wu [41] presented a distributed solution for computing a *k-connected k-dominating set* as a backbone of wireless networks. Their approach combines multiple domination and the $k$-vertex connected property, which guarantees that a CDS remains vertex connected even when removing up to $k-1$ nodes from the graph.

Shi and Srimani [42] proposed a distributed solution for computing a *d-hop connected d-hop dominating set*. For any graph $G$, let $G_d$ denote the *d-closure* of $G$, which means the graph built on vertices of $G$ such that any pair of vertices in $G$ are connected by an edge in $G_d$ if the vertices are within distance $d$ of each other in $G$. A $d$-hop DS is $d$-hop connected if it is connected in $G_d$.

Fernandess and Malkhi [43] presented $k$-clustering as a framework in which the wireless network is divided into non-overlapping clusters where every two nodes in the same cluster are at most $k$ hops apart.

Krishna et al. [44] proposed a routing approach based on 1-clusters, which builds overlapping clusters such that nodes in the same cluster are adjacent to each other. This is why it is called a 1-cluster, which can be further extended to $k$-clusters in which nodes are at most $k$ hops apart in their clusters. 1-Clusters can be seen as a DS of the network in case one node per cluster is elected as a cluster-head.

Kim et al. [45] introduced a clustering approach where any node can start a cluster such that it includes all nodes within distance $k$ from the node starting the cluster. However, the resulting structure does not constitute a $k$-hop dominating set of the network.

Joshi et al. [10] have provided centralized solutions for solving the $(k,r)$-DS problem in *interval graphs* (IG). A graph $G$ is said to be an interval graph if there is a one-to-one correspondence between a finite set of closed intervals of the real line and the vertex set $V$, and two vertices $u$ and $v$ are said to be connected if and only if their corresponding intervals have a non-empty intersection. Even though the solutions presented by Joshi et al. [10] are optimal, IGs are limited to very simple network topologies.

## 3. $(k,r)$-Dominating sets: centralized solution in arbitrary graphs

The centralized solution presented in this section, KR, requires that the entire network topology be known. Hence, if the solution were used in a network, the network topology would have to be broadcast, and the dominating set computed by each node in the network would be the same for any node.

### 3.1. Description

Any node $i$ is said to be $(k,r)$-dominated (or simply dominated) if node $i$ has at least $k$ neighbors within distance $r$ in $D$ (refer to Table 1 for notation).

For the computation of a DS with parameters $k$ and $r$, the *Domin* value of node $i$, $i.Domin$, is $k$ minus the number of nodes in $D$ within distance $r$ from $i$ if node $i$ is not covered (or dominated). Once node $i$ has been dominated, or node $i$ is selected as dominating, its *Domin* is set equal to zero, and no longer changes its value. The *total* value of node $i$, $i.total$, is given by $i.total = \sum_{k \in \Re_{r,i}} k.Domin$.

A natural greedy solution to the $(k,r)$-DS problem would be to repeatedly select as *dominating* the node that covers the most number of uncovered nodes (i.e., nodes not yet $(k,r)$-dominated), until all nodes are $(k,r)$-dominated. However, KR applies a different greedy approach, and repeatedly selects as *dominating* the node with the current largest *total* value. That is, KR selects as *dominating* the node that covers the most number of nodes with fewer dominating nodes (it could be the node with the most number of nodes not yet covered), which is quantified by the *total* parameter of any node. This way, any selected node $i$ potentially affects the *total* value of any node within distance $2r$ from node $i$. On the other hand, in the natural greedy approach, any selected node $i$ only affects the coverage of nodes when some node in node $i$'s $r$-hop neighborhood gets $(k,r)$-covered. That is, any selected node $i$ reduces by one the *Domin* value of any node $n$ not yet covered in node $i$'s $r$-hop neighborhood, but it does not reduce by more than one unit node $n$'s coverage if no node gets $(k,r)$-covered other than node $i$ itself. For *multiple domination* one (i.e., $k = 1$), KR is equivalent to the natural greedy approach (i.e., the first time any node $i$ gets covered, its *Domin* value becomes zero).

Table 1
Notation

| | |
|---|---|
| $G = (V,E)$ | A graph, $G$, where $V$ is the set of nodes (vertices), and $E$ is the set of links (edges) |
| $N_r^i$ | The set of $r$-hop neighbors of node $i$ (assuming $N_0^i = \{i\}$) |
| $\Re_{r,i}$ | The $r$-hop neighborhood of node $i$; i.e., $\Re_{r,i} = \bigcup_{k=0}^{r} N_k^i$ |
| $\Delta$ | The largest cardinality among all $r$-hop neighborhoods in the network. That is, $\Delta = \max(\lvert\Re_{r,k}\rvert),\ \forall k \in V$ |
| $D$ | The dominating set |

Fig. 1 presents a pseudo-code for KR. Initially, *Domin* is set equal to $k$ for all nodes, and *total* depends on the number of nodes in the $r$-hop neighborhood of each node. Nodes are inserted in a *Heap* structure to make the selection of the node with the largest *total* value easier. At the beginning, all nodes are in the *Heap*, and while there are nodes in the *Heap*, there are nodes yet to be covered. The node with the largest *total* is selected as the next dominating node. Once node $s$ is selected, all nodes in its $r$-hop neighborhood that are not yet covered (i.e., $Domin > 0$), must have their *Domin* value recalculated to reflect the selection of node $s$. Nodes that become covered (i.e., $Domin = 0$) are removed from the *Heap*. After this update, nodes in node's $s$ **2r**-hop neighborhood remaining in the *Heap* must have their *total* value recalculated. Because nodes within distance $r$ from node $s$ may have their *Domin* value changed, it implies that nodes within distance $2r$ from $s$ may have their *total* value affected as well. After these updates, the *Heap* must be sorted, so that the node with the largest *total* can be selected next (ties are broken choosing the node with lowest ID). This process repeats, until the *Heap* is empty.

### 3.2. Example

Fig. 2 depicts an example of KR computing a $(2,2)$-DS of the network. Initially nodes have their *Domin* parameter set to 2, and the *total* parameter is computed depending on each node's two-hop neighborhood (Fig. 2(A)). Nodes $\{4,5,7,8\}$ have the same largest *total*, but recall that in KR ties are broken lexicographically; hence, node 4 is selected as dominating. Fig. 2(B) shows the network reflecting the selection of node 4, and once again there are several nodes with the same largest *total* (i.e., nodes $\{0,1,2,7,8,9\}$). In this case, node 0 is selected. After its selection (Fig. 2(C)), nodes $\{6,8\}$ have the same largest total. Node 6 is then selected, which makes all nodes $(2,2)$-covered

```
Data: G = (V, E), k, r
Result: D, the (k,r)-DS
begin
    D = ∅
    /* Initialize parameters Domin and total                    */
    foreach i ∈ V do
        i.total = k · |𝔑_{r,i}|
        i.Domin = k
        /* Insert i in the HEAP (sorted according to the total
           parameter). Ties are broken choosing the node with the
           lowest ID.                                           */
        Insert(HEAP,i)
    /* Select dominating nodes                                  */
    while HEAP ≠ ∅ do
        /* Node with largest total is the next dominating node  */
        s = First(HEAP)
        s.Domin = 0
        D = D ∪ {s}
        /* Update Domin value for each covered node             */
        foreach i ∈ 𝔑_{r,s} | i ∈ HEAP do
            − − i.Domin
            /* If a node becomes dominated remove it from heap  */
            if i.Domin == 0 then
                Remove(HEAP, i)
        /* Update total of nodes affected by the current selection */
        foreach i ∈ HEAP | i ∈ 𝔑_{2r,s} do
            i.total = 0
            foreach j ∈ 𝔑_{r,i} do
                i.total = i.total + j.Domin
        Sort(HEAP)
end
```

Fig. 1. KR: pseudo-code.

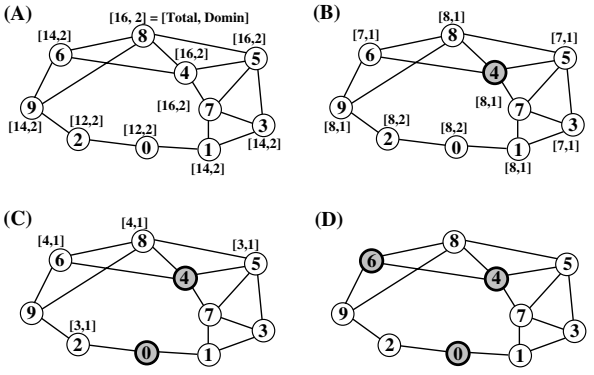


Fig. 2. Computing a (2,2)-DS of the network with KR.

(Fig. 2(D)). If the natural greedy approach were applied to the same network, then the order of nodes selected would be $\{4,5,0,2\}$, increasing the cardinality of the DS by one node.

## 4. Distributed clustering using $(k,r)$-dominating sets

As discussed previously, the *distance* and the *multiple* domination parameters can be used to define the degree of redundancy for bounded-distance clusters. That is, the two dominating parameters define the maximum distance from nodes to their cluster-heads, and the minimum number of cluster-heads per node, respectively.

We propose DKR, which is a distributed algorithm for clustering using $(k,r)$-DS. DKR is well suited for both synchronous and asynchronous networks. In the synchronous network model,

nodes exchange messages in synchronous rounds. In the asynchronous network model, nodes take steps at arbitrary times. Even though there are no rounds in the asynchronous model, it is possible to simulate rounds [46]. In order to do that, a node tags the message with its round number $x$. The recipient waits to receive round $x$ messages from all its neighbors before transitioning to the next round.

### 4.1. Summary description

We assume that nodes have unique identifiers (IDs), and that nodes know who their neighbors are. The latter can be implemented by means of a neighbor protocol with which nodes exchange hello messages [47], as part of the MAC protocol, or using periodic *HELLO* messages as part of the protocol itself.

Associated with any node $i$ in the network, there is a *process* that consists of the following components: A set of *states*, which is used for describing the current state of node $i$. A message-generation function that specifies any messages that node $i$ should send and to whom it should send them, depending on the current state of the system. Optionally, a list of *events*, each of them scheduled to happen at a specific time. A state-transition function specifying the new state to which node $i$ should transition for each possible state and messages received.

The *status* of a node reflects its role during the *clustering* process. Initially, there is no established hierarchy among nodes, and the nodes assume an *unknown* status. As the nodes organize themselves, their status change to reflect their role in a cluster, which can be one of the following:

- *Dominating*, the node is a *cluster-head*.
- *Pending dominating*, the node may become a *cluster-head*.
- *Dominated*, the node has *at least* $k$ cluster-heads within distance $r$.
- *Gateway*, in addition to being *dominated*, the node connects other nodes to their cluster-heads.

A **round** of messages is defined as the successful transmission of a message **m** by any node $n$ to all its one-hop neighbors. If rounds are numbered, a round $x$ is deemed complete only after all nodes have sent the messages for round $x$. DKR has two phases:

- *Phase one* (election phase): Each node elects $k$ nodes with smaller IDs (possibly including the node itself) within distance $r$. Elected nodes are just *candidates to be cluster-heads*. Because each node has its own set of $k$ elected nodes within distance $r$, the sets of elected nodes *dominate* all non-elected nodes in the network.
- *Phase two*: During this phase *cluster-heads* are assigned, and nodes are affiliated to their cluster-heads.

Clearly, there must be at least $k$ nodes in every node's $r$-hop neighborhood for the required *multiple domination* to be satisfied. In the subsequent description of DKR, we assume that multiple domination can be satisfied at each node.

It is possible that not all nodes elected during *phase one* become *cluster-heads*, because some redundant candidates are identified, and pruned. The rationale for choosing node IDs over node degree for the election process is that elections based on node degree can result on high turnover of dominating nodes when the topology changes, because the degree of a node is much more likely to change than the node ID relative to its neighborhood [48].

#### 4.1.1. Phase one

This phase takes $r$ rounds to complete in a static topology. A pseudo-code for *phase one* is presented in Fig. 3. For asynchronous networks, rounds are simulated as described previously. At the beginning of a new round, a node advertises its list of $K \leqslant k$ smaller ID nodes. After a number of rounds ($r$ rounds in a static topology), a node $i$ in the network learns the set of $k$ nodes with smaller IDs (possibly including the node itself) within distance $r$ from it. We denote such a set by $D_i'$.

An elected node can elect itself or be elected by other nodes. A node that elects itself is called *properly elected* if the node is not elected by any other node, and is called *self-elected* if the node is elected by at least one other node. A node that does not elect itself and is elected by other nodes that are not elected is called *neighbor-elected*. After the election, any node $i$ in the network changes its status as follows: If node $i$ is properly elected or self-elected, node $i$ changes status to pending dominating. Otherwise, node $i$ has status dominated.

Note that a *properly elected* node must become dominating, because there are at most $k-1$ other

**states consists of:**

| | |
|---|---|
| $i$ | the node ID |
| $k, r$ | the dominating set parameters |
| $D_i'$ | set of tuples, with each tuple specifying a node ID, the node advertising it (i.e., next-hop toward the node), and the distance to the node. **Initially** contains only the tuple $\{i, i, 0\}$ |
| $A$ | set formed from set $D_i'$, with each tuple specifying a node ID and its known distance (i.e., first and third parameter of each tuple from set $D_i'$) |
| $H$ | set of all nodes known during the election process |
| *status* | {*unknown*, *pending_dominating*, *dominating*, *dominated*}, **initially** *unknown* |
| *rounds* | an integer, initially 0 |
| $NA$ | set of nodes requiring *Neighborhood Advertisement* (**initially** $NA = \emptyset$). To be used during *Phase Two* |

**messages:**

**if** $rounds < r$ **then**
  /* send current list of known $K \leq k$ known smallest node identifiers */
  send $A$ to all neighbors

**transitions:**

$rounds = rounds + 1$
$M = \{$set of tuples received from neighbors$\}$
**foreach** $t = \langle domin, dist \rangle \in M$ **do**
  **if** $\exists\ \langle a, b, c \rangle \in D_i' \mid a == domin,\ (dist + 1) < c$ **then**
    /* found shorter path to some node in set $D_i'$ */
    $b = node\ sending\ tuple\ t$
    $c = dist + 1$
    update correspondent entry in set $H$
    $M = M \setminus \{t\}$

/* keep in set $H$ info about all nodes known during the election process */
$H = H \cup M$
$P = D_i' \cup M$
$D_i' = \{min(k, |P|)$ smaller IDs in $P\}$
update $A$ with new set $D_i'$
**if** $rounds == r$ **then**
  /* Election is over! */
  **if** $i \in D_i'\ OR\ |D_i'| < k$ **then**
    **if** $|D_i'| < k$ **then**
      /* Not satisfiable (i.e., fewer $r$-hop neighbors than the required *multiple* domination parameter); hence, must become *dominating* */
      $status = dominating$
      /* Must send an NA message during *phase two*, advertising I am a *cluster-head* */
      $NA = \{i\}$
    **else**
      $status = pending\_dominating$
    $D_i' = \emptyset$
  **else**
    $status = dominated$
  /* go to *Phase Two* */
  DKR: *Phase Two*

Fig. 3. DKR: *phase one* (election phase).

elected nodes in node $i$'s $r$-hop neighborhood. Because identifying properly elected nodes would incur extra overhead, they are implicitly notified of their dominating status after not hearing from enough dominating nodes within a given period of time.

A *neighbor-elected* node $i$ is elected by at least one node, call it $n$, which is not elected and for which node $i$ is strictly required. That is, there is no self-elected node in node $n$'s $r$-hop neighborhood that could possibly replace node $i$; otherwise, node $n$ would have elected that self-elected node. Even though in some cases a properly elected node could replace node $i$, initially DKR chooses to select all neighbor-elected nodes as cluster-heads.

#### 4.1.2. Phase two

During *phase two*, some or all nodes elected during *phase one* become cluster-heads. In addition, the rest of the nodes are affiliated to their cluster-heads. A pseudo-code for phase two is presented in Figs. 4 and 5.

**states consists of (in addition to states from** *phase one*):

$N$ set of nodes requiring Notification (**initially** $N = \emptyset$)

$LA$ **initially** *True*

$J$ set of nodes requiring a *Join* message (**initially** $NA = \emptyset$)

**messages:**

**if** $LA ==$ *True* **then**
- /* only *dominated* nodes send Local Advertisement */
- **if** *status* $==$ *dominated* **then**
  - Broadcast $\langle$LocalAdvertisement, $D_i^{'}\rangle$
- **else**
  - /* Schedule event for checking dependencies */
  - ScheduleEvent(CheckStatus, *Wait Period*)
- LA = False;

**if** $N \neq \emptyset$ **then**
- **foreach** $\langle target, next\_hop\rangle \in N$ **do**
  - Unicast to *next_hop* $\langle$Notification, *target*$\rangle$;
- $N = \emptyset$

**if** $NA \neq \emptyset$ **then**
- /* a node may be initiating or relaying a neighborhood advertisement */
- **foreach** $n \in NA$ **do**
  - *Broadcast* $\langle$NeighborhoodAdvertisement, $n\rangle$
- $NA = \emptyset$

**if** $J \neq \emptyset$ **then**
- **foreach** $n \in J$ **do**
  - $l =$ get next hop to $n$ from $H$
  - Unicast to $l$ $\langle$Join, $n\rangle$;
- $J = \emptyset$

**events:**

**case** *CheckStatus*
- $L = \{$validated nodes in set $D_i^{'}\}$
- **if** *status* $==$ *pending_dominating* **then**
  - **if** $|L| < k$ **then**
    - /* not enough dominating nodes within distance $r$ */
    - *status* = *dominating*
  - **else**
    - **if** *any Notification relayed by this node* **then**
      - *status* = *gateway*
    - **else**
      - *status* = *dominated*
- **if** *status* $\neq$ *dominating* **then**
  - /* send a *Join* message to all my dominating nodes */
  - **if** $|L| < k$ **then**
    - $F = D_i^{'} \cap L$
    - /* Node still need to choose $k - |L|$ dominating nodes. Choose the closest ones! */
    - $L^{'} = \{k - |L|$ closest nodes in $F\}$
    - $J = L \cup L^{'}$
    - Validate entries in $L^{'}$
  - **else**
    - $J = L$

Fig. 4. DKR *phase two*: **states**, **messages**, and **events**.

The messages used during this phase are

- *Local advertisement* (LA): A message having the list of nodes elected by the sender, and the respective next-hop to each one of the elected nodes.
- *Neighborhood advertisement* (NA): A message advertising a cluster-head.
- *Notification*: A message sent to notify a node that must become cluster-head.
- *Join*: A message sent to notify, or to connect to a given cluster-head.

Because neighbor-elected nodes are not aware of their election, a notification mechanism is needed to notify them. Depending on the coverage provided by neighbor-elected nodes, some self-elected nodes may be ruled out as cluster-heads.

At the beginning of *phase two*, dominated nodes send to their one-hop neighbors an LA message

```
transitions:
  M = {message(s) from neighbor(s)}
  foreach m ∈ M do
    Sender = node that sent message m
    switch m do
      case LocalAdvertisement
        /* local advertisements consist of tuples ⟨a, b, c⟩, where a=dominating node,
           b=next hop, and c=distance                                              */
        foreach ⟨a, b, c⟩ ∈ m do
          if a == i AND status≠ {dominating OR pending_dominating} then
            status=dominating
            NA = NA ∪ {i}
          else
            if status==dominated AND b == i then
              status=gateway
            /* If a is not one of my selections, and I am on the path to node a,
               send a notification to a                                           */
            if a ∋ D'_i AND b == i then
              n = get next hop to a from set H
              if n ≠ Sender then
                N = N ∪ {⟨a, n⟩}
      case Notification
        if target == i then
          /* I am being notified                                                  */
          if status ≠dominating then
            status=dominating
            NA = NA ∪ {i}
            D'_i = ∅
        else
          if target ∋ D'_i then
            n = get next hop to target from set H
            if n ≠ Sender then
              if status == dominated then
                status=gateway
              if Notification not previously sent to target then
                N = N ∪ {⟨target, n⟩}
      case Join
        if target ≠ i then
          if Join OR Notification not yet sent to target then
            if status == dominated then
              status=gateway
            J = J ∪ {n}
      case NeighborhoodAdvertisement
        a = advertised node
        if status ≠ dominating then
          if a ∋ D'_i then
            D'_i = D'_i ∪ {a}
        if (distance to a) < r then
          /* relay advertisement                                                  */
          NA = NA ∪ {a}
```

Fig. 5. DKR *phase two*: **transitions**.

containing their elected nodes. Any dominated node $i$ proceeds as follows upon receiving an LA message:

- If node $i$ is listed in the advertisement, node $i$ changes its status to dominating, triggering an NA message announcing node $i$ as cluster-head to all its $r$-hop neighbors. This is accomplished by broadcasting the NA message using restricted *blind-flooding* with the *time-to-live* (TTL) field set equal to $r$.
- If node $i$ is not listed in the LA message but is listed as a next hop to any advertised node, then node $i$ changes its status to gateway.

- For any advertised node $n \in \mathrm{LA}$ that is not among the nodes elected by node $i$ (i.e., $n \ni D'_i$) it sends a *Notification* message to node $n$. Upon receiving the notification, if the notified node is not yet dominating, the node advertises itself via an NA message.

**Definition 1.** For any node $i$, and for all $n \in D'_i$, node $n$ is deemed **validated** only upon the reception of the respective NA message advertising node $n$; otherwise, node $n$ is not yet **validated**.

Any neighbor-elected node $n$ eventually changes its status to dominating, by either receiving a *Notification* message originated at some node $k$ within distance $1 < d < r$, or by receiving an LA message from some one-hop neighbor. In any case, once node $n$ becomes dominating, it sends an NA message.

Because an NA message is sent only when a node changes its status to dominating, nodes receiving an NA message advertising node $n$ know that it is a cluster-head. When processing an NA message for node $n$, any node $i$ inserts an entry for node $n$ in $D'_i$.

A *pending dominating* node $i$ does not become a *cluster-head* if (a) node $i$ has *at least* $k$ validated dominating nodes within distance $r$, and (b) every node $n$, which elected node $i$ during *phase one*, is also covered by a set of validated dominating nodes.

**Definition 2. Wait Period** is the minimum time required for reaching an agreement in *phase two*.

In the worst case, a *Notification* for node $n$ is initiated by some neighbor $i$ located $r-1$ hops away from node $n$. The correspondent NA message initiated by node $n$ reaches the most distant neighbors (i.e., $r$ hops away from node $n$) after $r$ successfully transmissions of message NA. Therefore, **Wait Period** should be larger than the time required for $2r$ transmissions of a message. After a period of time equivalent to *Wait Period* any node $i$ in the network checks its coverage. If node $i$ is *pending dominating*, and it does not have enough validated entries in $D'_i$, then node $i$ changes status to *dominating*, and sends an NA message. This means that node $i$ does not have enough information for ensuring its own coverage (i.e., node $i$ does not know if it has at least $k$ dominating nodes within distance $r$). Otherwise, any non-*dominating* node $i$ sends a *Join* message to $k$ nodes from $D'_i$ (including the nodes already validated). If there are more than $k$ validated entries in $D'_i$, node $i$ chooses the closest ones (ties are broken choosing the node with smaller ID).

Like a *Notification*, which also serve for assigning *gateways* while the message is being routed to its destination, *Join* messages also serve to notify any *pending dominating* node that is still required as cluster-head. That is, even though some *pending dominating* node $i$ finds itself covered, there might be some node $j$ that still needs node $i$ (i.e., without node $i$, node $j$ does not have the required number of dominating nodes). While a *Join* message is being routed to destination $n$, a node $i$ processing the message does not need to relay it if a *Notification*, or another *Join* message, had already been sent to node $n$. So, we assume that every node keeps track of recent *Notification* and *Join* messages sent by the node. After *Join* messages reach their destinations, all regular nodes are connected to their cluster-heads.

### 4.2. Examples

Consider the example presented in Fig. 6, where nodes are computing a $(2,3)$-DS of the network. During *phase one* nodes elect the two nodes with smaller IDs in their 3-hop neighborhood (Fig. 6(A)). Nodes 10 and 15 are *self-elected*, and nodes 18 and 20 (both elected by node 90) are *neighbor-elected*. *Self-elected* nodes assume status *pending dominating*. The remaining nodes assume status *dominated*, and send an LA message advertising their list of elected nodes. After receiving the LA message from neighbor 90, node 18 changes status to *dominating*, and sends an NA message that eventually reaches all nodes within distance 3 from node 18. Fig. 6(B) show the status of nodes, and their corresponding *validated* dominating entries. Besides sending the NA message, node 18 also sends a notification to node 20. Fig. 6(C) presents the status of the network after the notification has reached node 20. The notification makes node 20 change its status to *dominating*, triggering an NA message that eventually reaches all neighbors within distance 3 from node 20. After all affected nodes process this NA message, we notice that all dominated nodes are satisfied, because each of them have two validated entries in their $D'$ lists. *Wait Period* should be set appropriately, so that by the time the event *CheckStatus* happens all NA messages have already been delivered and processed. Fig. 6(D) shows the status of the network after all dominated nodes have sent out their *Join* messages to their

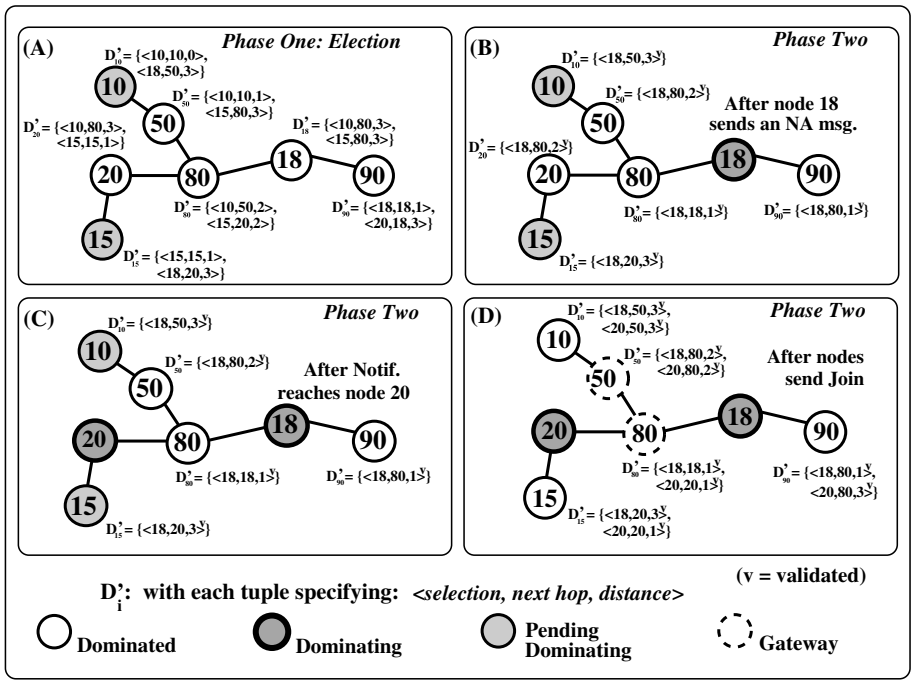


Fig. 6. Computing a (2,3)-DS of the network.

dominating nodes (assume that *Join* messages are grouped together whenever different dominating nodes are reachable through the same node). In this case, because nodes 50 and 80 have either relayed a *Notification*, or a *Join* message, they serve as *gateways* for other dominated nodes.

Fig. 7 presents an example comparing DKR (*assuming maximum ID nodes are elected, rather than the default minimum ID*) to *Max–Min* [35], when computing a (1,3)-DS of the network. This is the same network example presented in [35]. While *Max–Min* produces four cluster-heads, DKR produces three. For DKR, node 100 is a *self-elected* node. The other elected nodes (i.e., 65, 73, and 85) are eventually notified, and announce themselves by sending NA messages. All nodes that have elected node 100 (including node 100 itself) receive at least one of these NA messages, validating the respective advertised node. This is an indication that the election process can be improved, producing smaller dominating sets. Simulation results presented latter corroborate this.

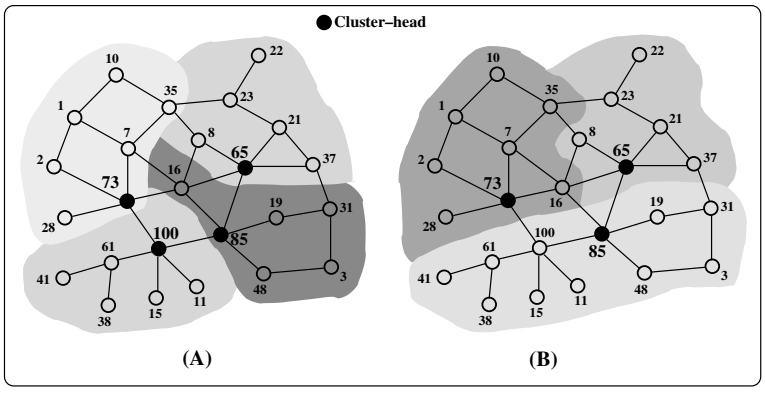


Fig. 7. Computing a (1,3)-DS of the network (same network example presented in [35]): (a) Max–Min, and (b) DKR (*assuming maximum ID nodes are elected, rather than the default minimum ID*).

## 5. Performance

Even though KR and DKR compute a $(k,r)$-DS of any arbitrary topology, the topologies considered for simulations are those of wireless networks modeled using *unit disk graphs* [36]. Given that the $(k,r)$-DS problem is NP-Complete and that KR is the first known approximation, we use it as a lower bound to assess the performance of the distributed solution.

To focus on the efficiency of the heuristics themselves, we use a customized simulator for ad hoc networks, and assume an ideal MAC protocol with which no collisions can occur. This is the same approach adopted in all prior work [20,49,23,28] to compare the efficacy of heuristics. As discussed previously, DKR works in both synchronous and asynchronous networks. However, for the simulations we assume synchronous networks.

Experiments are repeated for 100 trials with different network topologies, varying the number of nodes and terrain size. Nodes are randomly placed over the terrain, and connectivity is tested to ensure that the network is connected. The radio range is set to 250 m. The results represent the average over the 100 different networks . The network size is varied from 100 nodes to 500 nodes, with increments of 50 nodes. For the same number of nodes, we vary the terrain size according to two configurations so that we can test the algorithms for different node density (see Table 2 for terrain dimensions). Configuration 1 has a node density of 50 nodes/km$^2$, and Configuration 2 has 100 nodes/km$^2$.

To give an idea of the characteristics of the networks being evaluated, Table 3 presents the values for the network diameter (i.e., the largest distance between any pair of nodes), and the average node degree for all network sizes. These results show that as the network size increases, so does the network diameter. But it also shows that, in each configuration, the average node degree is similar for all network sizes. In Configuration 2, nodes have almost twice as many neighbors compared to the networks from Configuration 1. On the other hand, the network diameter is smaller for networks in Configuration 2, because there are in average twice as many nodes spread over the same area compared to Configuration 1.

Table 2
Terrain size (in meters)

| # of nodes | Configuration 1 | Configuration 2 |
|---|---|---|
| 100 | $1414 \times 1414$ | $1000 \times 1000$ |
| 150 | $1732 \times 1732$ | $1225 \times 1225$ |
| 200 | $2000 \times 2000$ | $1414 \times 1414$ |
| 250 | $2236 \times 2236$ | $1581 \times 1581$ |
| 300 | $2449 \times 2449$ | $1732 \times 1732$ |
| 350 | $2645 \times 2645$ | $1871 \times 1871$ |
| 400 | $2828 \times 2828$ | $2000 \times 2000$ |
| 450 | $3000 \times 3000$ | $2121 \times 2121$ |
| 500 | $3162 \times 3162$ | $2236 \times 2236$ |

Table 3
Network diameter and node degree (results represent the average $\pm$ std over 100 samples)

| # of nodes | Configuration 1 | | Configuration 2 | |
|---|---|---|---|---|
| | Diameter | Degree | Diameter | Degree |
| 100 | $10.4 \pm 0.1$ | $7.8 \pm 0.1$ | $6.2 \pm 0.1$ | $14.9 \pm 0.1$ |
| 150 | $12.8 \pm 0.2$ | $8.1 \pm 0.1$ | $7.9 \pm 0.1$ | $15.5 \pm 0.1$ |
| 200 | $14.7 \pm 0.1$ | $8.2 \pm 0.1$ | $9.2 \pm 0.1$ | $16.2 \pm 0.1$ |
| 250 | $16.5 \pm 0.1$ | $8.3 \pm 0.1$ | $10.2 \pm 0.1$ | $16.5 \pm 0.1$ |
| 300 | $18.1 \pm 0.1$ | $8.3 \pm 0.1$ | $11.3 \pm 0.1$ | $16.7 \pm 0.1$ |
| 350 | $19.4 \pm 0.1$ | $8.5 \pm 0.1$ | $12.2 \pm 0.1$ | $16.9 \pm 0.1$ |
| 400 | $20.9 \pm 0.2$ | $8.7 \pm 0.1$ | $13.0 \pm 0.1$ | $17.1 \pm 0.1$ |
| 450 | $21.9 \pm 0.1$ | $8.7 \pm 0.1$ | $13.8 \pm 0.1$ | $17.2 \pm 0.1$ |
| 500 | $23.4 \pm 0.1$ | $8.7 \pm 0.1$ | $14.6 \pm 0.1$ | $17.3 \pm 0.1$ |

Three performance metrics are evaluated:

- *Signaling overhead*, which consists of the total number of control packets exchanged for gathering topology information in a centralized algorithm, or for the execution of the two phases in the distributed algorithm.
- Total number of *cluster-heads* (i.e., *dominating* nodes), which is the number of $(k,r)$-dominating nodes for different configurations.
- *Cluster-head sparseness*, which gives a measure of the efficacy of distributing cluster-heads over the network. It gives the average number of cluster-heads, within the *distance* parameter, per regular node. Results closer to the *multiple domination* parameter means better distribution of cluster-heads. Ideally, regular cluster nodes should have **exactly** $k$ cluster-heads within distance $r$, but this is not a requirement when solving the $(k,r)$-DS problem (which stipulate *at least* $k$ dominating nodes).

To apply KR, complete topology information must be gathered using reliable flooding of link-state updates. In the distributed algorithm, HELLO messages are used for obtaining the one-hop neighbor information, and control messages are reliably transmitted during the two phases of DKR.

Fig. 8 presents the *signaling overhead* for Configuration 1 (results for Configuration 2 are omitted because they are similar to Configuration 1). The control overhead incurred by the centralized algorithm is due only to the dissemination of topology information. After the nodes have complete topology information, they can compute clusters locally for any parameters. For the distributed algorithm, the signaling overhead varies mostly due to the distance parameter. The number of control packets increases as the number of rounds increases; that is, larger the distance parameter, larger the number of advertisements. To show how the signaling overhead varies with the distance parameter, results are presented for parameter $r$ varying from 1 to 4 when computing a $(1,r)$-DS of the network.

Because DKR discards self-elected nodes whenever possible, in the worst case all elected nodes become *cluster-heads*. However, in *Max–Min* [35] all nodes elected at the end of the first $r$ rounds become cluster-heads; but, it is only during the second set of $r$ rounds that some of the elected nodes find out about their *dominating* status. Besides that, in certain scenarios *Max–Min* generates cluster-heads that are on the path between a node and their elected cluster-heads. In that case, only during the *convergecast* (which is used to connect regular nodes to their cluster-heads) that nodes adjust their selections to the closest cluster-head. To show how DKR reduces the number of cluster-heads compared to *Max–Min*, when computing $(1,r)$-DS (recall that *Max–Min* computes only $r$-hop dominating sets), we present simulations for different values of the distance parameter.

Fig. 9 presents the results for the total number of cluster-heads for Configuration 1, varying the distance parameter from 2 to 4. And Fig. 10, presents the results for Configuration 2. For both configurations DKR always selects fewer cluster-heads compared to *Max–Min*, meaning that usually some *self-elected* nodes are ruled out as cluster-heads.

Fig. 11 presents the *cluster-head sparseness* results for DKR and *Max–Min* for the networks in Configuration 1. DKR not only reduces the total number of *cluster-heads*, but it also distribute them more evenly over the network. Recall that, in DKR, some *pending dominating* nodes do not become cluster-heads, and regular nodes join the closest elected nodes.

Fig. 12 and 13 present the results in terms of total number of *cluster-heads* for Configuration 1 when computing $(k,2)$-DS, $(k,3)$-DS, and $(k,4)$-DS using KR and DKR. As expected, DKR produces more cluster-heads. For dominating distance two (i.e., $r=2$), both algorithms behave similarly, presenting a large difference between one dominating node (i.e., $k=1$) and two-dominating nodes (i.e., $k=2$). As we increase the distance parameter, fewer nodes are necessary for the dominating set. In KR, each time a dominating node is selected, it spans (i.e., dominates) a larger set of nodes. In DKR, nodes with smaller IDs get elected by more nodes farther

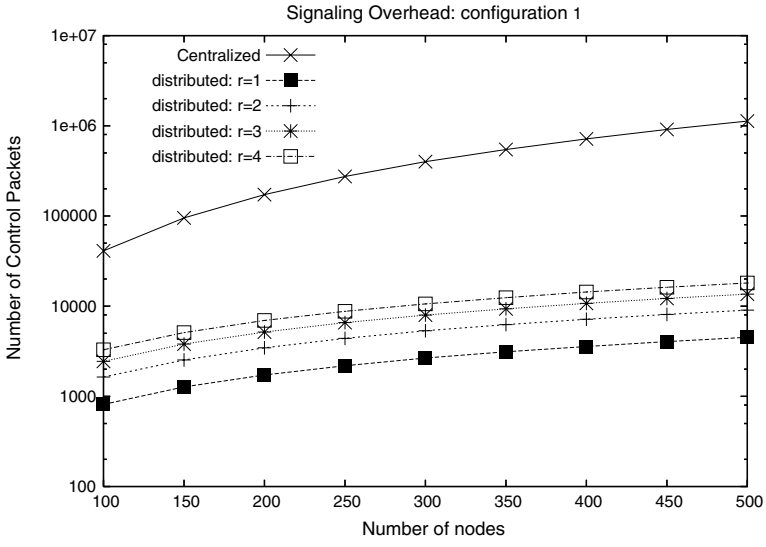


Fig. 8. Centralized versus distributed: *signaling overhead.*

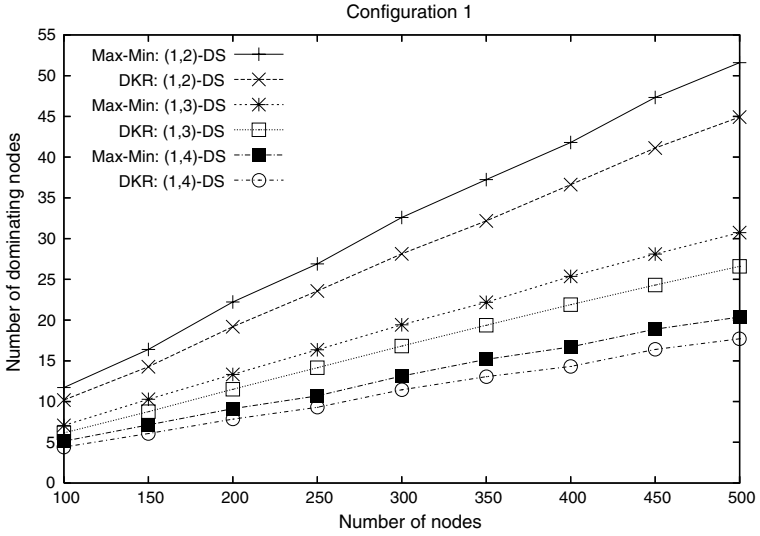


Fig. 9. DKR versus *Max–Min*, Configuration 1: # of *cluster-heads*.

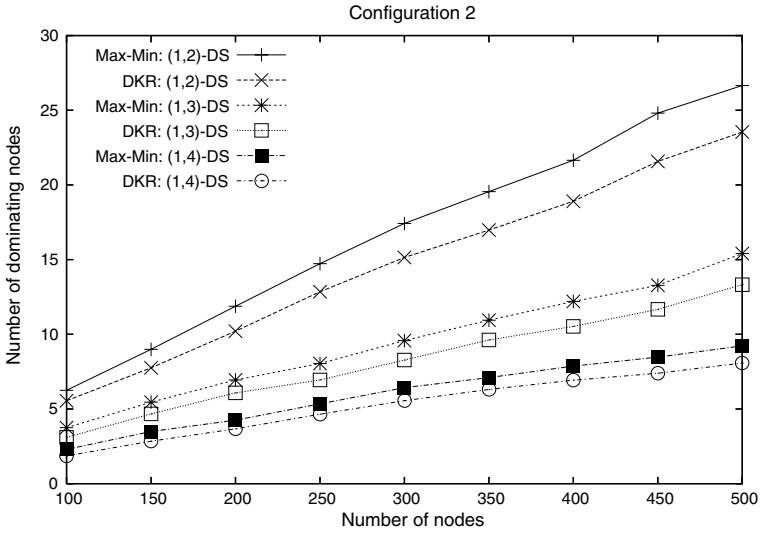


Fig. 10. DKR versus *Max–Min*, Configuration 2: # of *cluster-heads*.

away as the radius of the election increases. DKR produces similar results for (1,2)-DS and (2,3)-DS, because the election of one cluster-head in the two-hop neighborhood of any node increases the chances that, at the end, more elected nodes exist in the three-hop neighborhood of any node. Results for KR computing (3,3)-DS are similar to those for DKR computing (2,3)-DS. Similarly, results for KR computing (3,4)-DS are close to those for DKR computing (2,4)-DS (Fig. 13). In both cases, it shows that while the election of nodes is a simple and economic (in terms of overhead) solution, it is shown to be not as efficient as the centralized solution, because the election does not take into account the coverage of nodes when selecting dominating nodes. In case any coverage information should ever be a requirement for the election, this extra information would certainly increase signaling overhead.

Fig. 14 presents the results in terms of total number of *cluster-heads* for Configuration 2 using KR and DKR, when computing $(k,2)$-DS and $(k,3)$-DS of the networks. Both approaches produce in

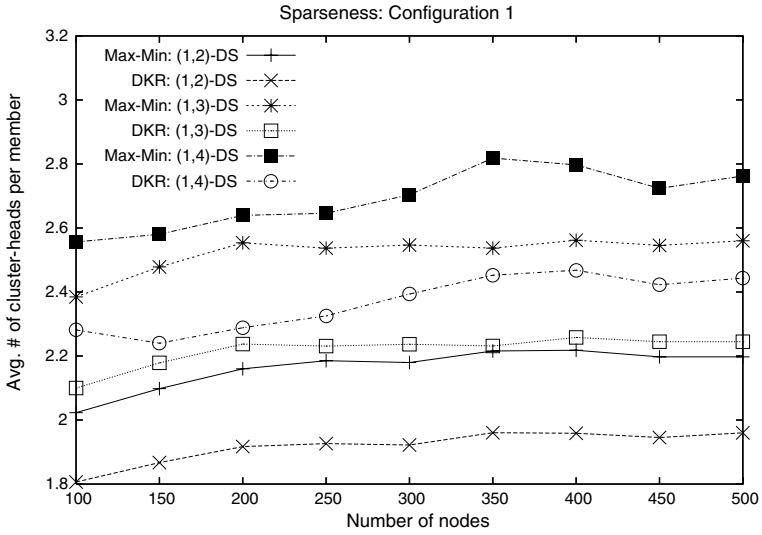


Fig. 11. DKR versus *Max–Min*: *cluster-head sparseness*.

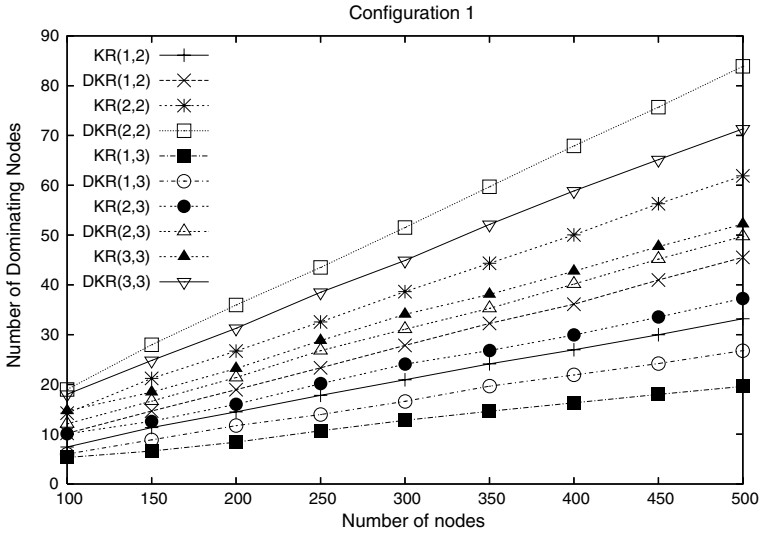


Fig. 12. Configuration 1, computing $(k,2)$-DS and $(k,3)$-DS: # of *cluster-heads*.

average half the number of cluster-heads compared to the results for Configuration 1. This is expected, because in Configuration 2 the networks have smaller diameter, and nodes have almost twice as many neighbors. However, the network diameter does not grow as fast as in the networks from Configuration 1. The diameter increases in average less than one unit, as the network size is incremented by 50 nodes (with the exception of the initial jump from approximately 6.2–7.9, see Table 3). Because the network diameter grows slower, so does the cardinality of the DS computed.

Fig. 15 presents the results when computing $(k,4)$-DS for the networks in Configuration 2. For networks with 200–350 nodes, the distributed algorithm performs well compared to the centralized. Because the distance parameter is larger compared to those from the previous scenarios, and the network diameter does not grow too much for larger networks, it seems that, for relatively small diameter

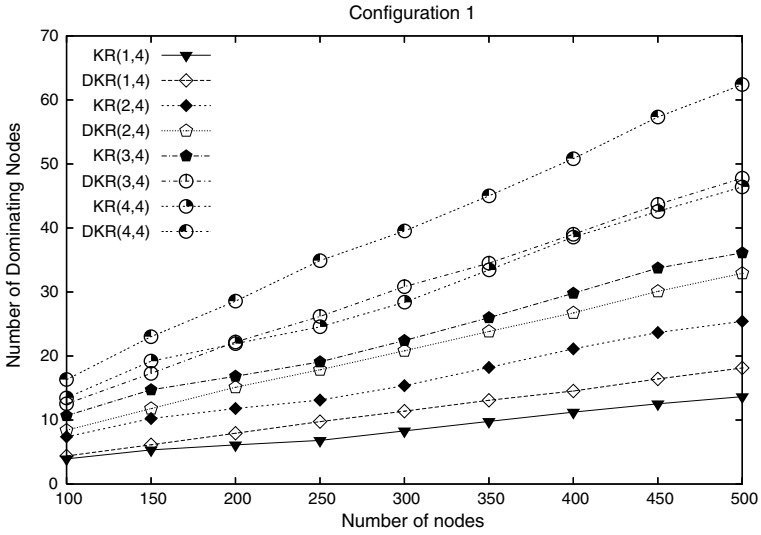


Fig. 13. Configuration 1, computing $(k,4)$-DS: # of *cluster-heads*.

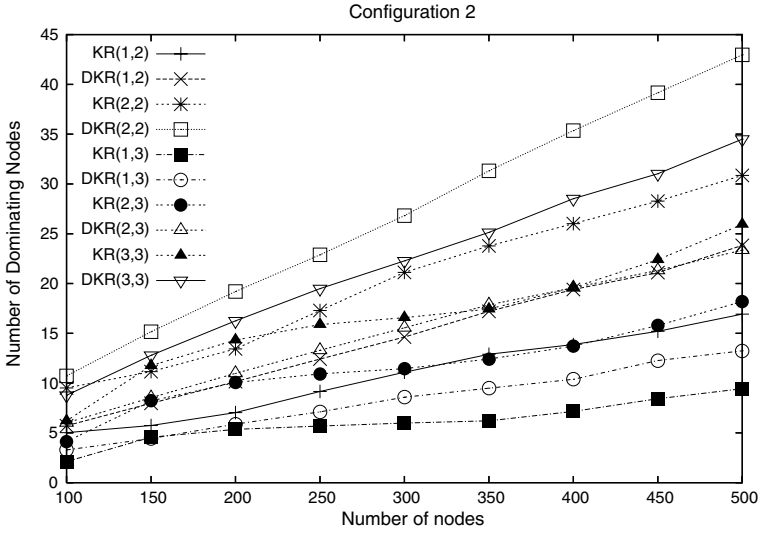


Fig. 14. Configuration 2, computing $(k,2)$-DS and $(k,3)$-DS: # of *cluster-heads*.

networks, as the radius of the election process increases, much more nodes select the same set of dominating nodes, hence reducing the total number of *cluster-heads*. In addition to that, *self-elected* nodes are more likely to be ruled out as cluster-heads.

There is a clear trade-off between efficiency, and communication cost. For MANETs, it pays-off to increase the average number of cluster-heads per cluster, considering that keeping an accurate view of the entire network topology is not possible, and redundancy is desirable.

Figs. 16 and 17 present the results for the *cluster-head sparseness* metric, when computing $(k,4)$-DS for the networks in Configurations 1 and 2. Lower values for this metric indicate a better distribution of cluster-heads over the network. Previously, for the same scenarios, it was shown that the networks in Configuration 1 produce more cluster-heads than the networks in Configuration 2. As expected, the *cluster-head sparseness* improves when fewer nodes are selected as cluster-heads. In addition to that, we must compare these results to the corresponding *multiple domination* parameter. That is, as the

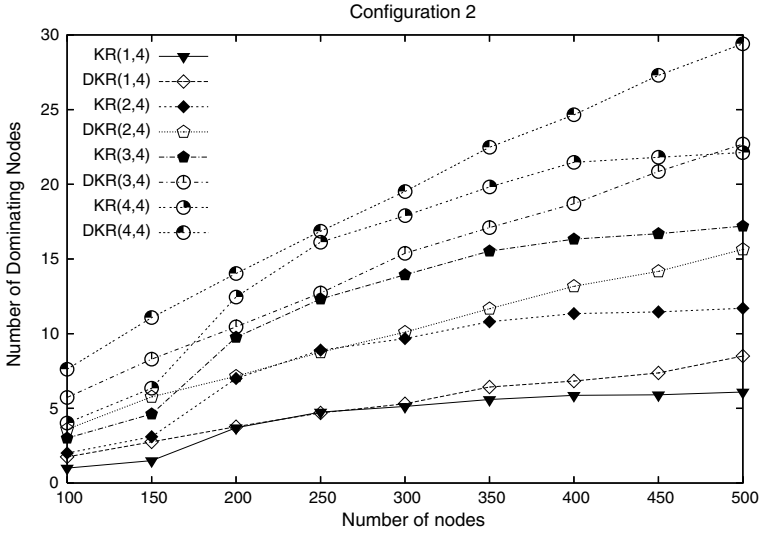


Fig. 15. Configuration 2, computing $(k,4)$-DS: # of *cluster-heads*.

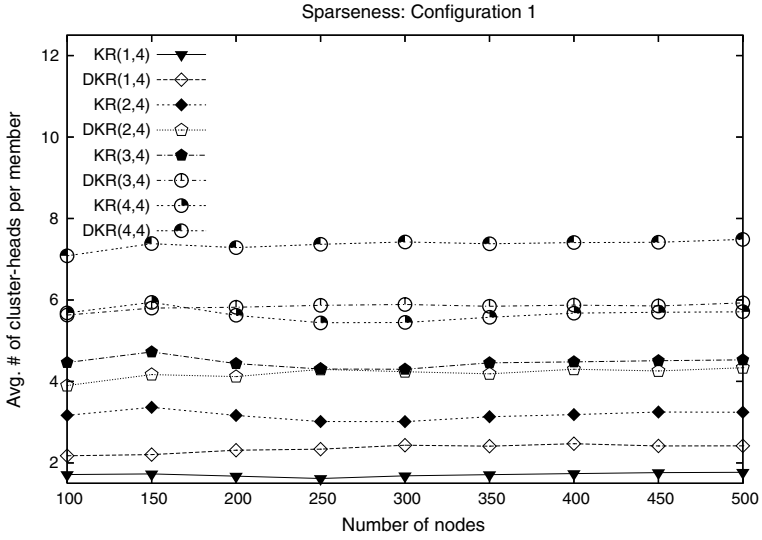


Fig. 16. Configuration 1, computing $(k,4)$-DS: *cluster-head sparseness*.

*cluster-head sparseness* metric approaches the *multiple domination* parameter, better the distribution of cluster-heads. Ideally, each node should have **exactly** $k$ cluster-heads within distance $r$, but this is not a requirement for solving the $(k,r)$-DS problem, which requires *at least k* cluster-heads.

Because more nodes exist in the $r$-hop neighborhoods of the networks in Configuration 2, the fewer cluster-heads selected are expected to affect a larger subset of nodes compared to Configuration 1. KR outperforms DKR in any situation under consideration, presenting results for the *cluster-head sparseness* metric closer to the corresponding *multiple domination* parameter. However, in Configuration 2 (with denser networks, and fewer cluster-heads) the two algorithms present similar results when the *multiple domination* parameter assume values 1 and 2. For larger values of this parameter, the difference increases, because in Configuration 2 the networks have smaller diameter (as shown in Table 3).

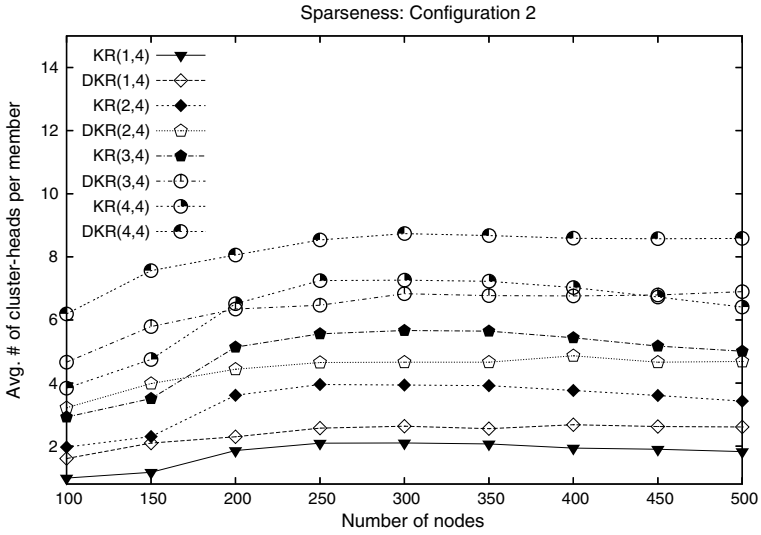


Fig. 17. Configuration 2, computing $(k,4)$-DS: *cluster-head sparseness*.

## 6. Case study

Multicast routing protocols can be classified as tree-based and mesh-based [50]. Tree-based can be further classified as single-source, shortest-path trees and shared, core-based trees. Core-based trees are more scalable compared to shortest-path trees, but usually present higher end-to-end delay and poor fault tolerance.

To improve the performance of core-based trees, multiple cores are deployed. Because more than one core is simultaneously active, the protocol tolerates core failure. With multiple cores there are two *one-to-all* designs [38]: *senders-to-all*, and *members-to-all*. In *senders-to-all*, senders transmit to all cores, and members join to just one core (usually the nearest one). In *members-to-all*, senders select one of the cores to send their data packets, and members need to join all cores. As for an *one-to-one* approach, each of the cores must join to at least one other [16]. In this case, senders send to just one of the cores, and receivers join to just one of the cores.

Senders-to-all has several advantages compared to members-to-all. Both approaches use one tree per core, but in members-to-all each tree connects all members, increasing the routing state in each router. In senders-to-all members decide which core to join, allowing members to choose the core that better satisfy their requirements (e.g., lower end-to-end delay).

While the *one-to-one* approach combines advantages of both *one-to-all* designs, it requires a reachability and maintenance protocol between the cores. Failure between any two virtually adjacent cores partitions the network.

Core placement has a direct impact on the performance of multi-core multicast routing protocols [38]. The selection of cores could be further extended to include a minimum number of cores within a maximum distance. In this context, the problem of finding the location of the cores is similar to computing a $(k,r)$-DS of the network. When selecting cores, redundancy is achieved by choosing a value for the parameter $k$ greater than one. At the same time, the distance parameter $r$ allows increasing local availability by reducing the distance to the cores.

### 6.1. Core hierarchical election for multicasting in ad hoc networks (CHEMA)

We present a novel multi-core multicast protocol named *core hierarchical election for multicasting in ad hoc networks* (*CHEMA*), which uses DKR for the election of cores, and is designed to work in the context of multiple-channel and multiple-interface. CHEMA's main features can be summarized as follows:

- Deploy multiple cores, with DKR as the core selection mechanism. This allows flexibility in terms of redundancy, and a bounded distance to the selected cores. *Core announcements* are used to disseminate core information throughout

the network. To reduce overhead, a single announcement aggregates information about all known cores.

- Use the senders-to-all approach. To reduce overhead, the packet header lists which neighbors should relay the multicast packet on a core basis. Instead of sending one packet toward each core, nodes relay the packet whenever they are listed at least once as a next-hop to any core. Before relaying the packet, the header is updated with entries for those cores for which the node is requested to forward the packet.
- Multi-channel and multi-interface: Each node has at least two interfaces. One is dedicated to receiving multicast transmissions from cores, and the other is used for any other transmission. Each core transmits in a channel different than any possible interfering core. That is, through core announcements nodes learn about all cores in the network, and the distance to each one. Using this information, cores select channels so that they do not interfere with other cores.
- Single *shot* approach: Once a multicast packet reaches the core, the packet is transmitted just once via the dedicated interface. In order to reach all receivers, the packet is transmitted with an increased power so that all nodes within $r$ hops from the core can receive it (i.e., any node that elected the core can receive the transmission). Because cores use different channels, and an interface is dedicated for receiving packets from the core, receivers should receive all transmissions sent by the core.

Multiple cores are selected via DKR. While cores have not yet been elected, multicast data packets are transmitted via *blind flooding*. After cores are elected, receivers join the nearest core by sending a join message to the core. Nodes aggregate all the fresh core announcements they receive, and broadcast them periodically every *core announcement interval* (which by default is set to be $3s$). Core announcements also include the number of members each core has. To let cores know about any associated members, an explicit *multicast join* message is sent from the receiver to the desired core whenever a node wants to join a multicast group. Note that this is not the same as the association provided by the join messages sent during the execution of DKR, which provides for connectivity from any node to at least $k$ cores within distance $r$.

Senders send multicast packets to all cores with members. Instead of sending one packet per core, the sender broadcasts just one packet with all the information regarding the cores that need to be reached. That is, the packet header includes an entry for every core and the corresponding next hop toward the core (recall that in DKR nodes keep information about which neighbors are used to reach each core). A node receiving the packet for which it is listed as a next hop to any core forwards the packet. Before relaying the packet, entries for which the node is listed as a next-hop are updated with the current information, and any other entries are excluded.

Because multicast packets are broadcast unreliably, a node may retransmit a packet up to $N$ times, unless it receives an implicit acknowledgment. That is, for every multicast packet transmitted the node relaying the packet keeps record of the packet and which neighbors should relay the packet. After a period of time equivalent to an *acknowledgment timeout*, the node checks if it has overheard any of the relayers transmissions. If the node fails to hear any of the relayer's transmission, the node retransmit the packet including only those nodes from which it has not yet heard from. Ideally, the length of an *acknowledgment timeout* should be set dynamically, because it depends on the level of contention, which is higher with a larger number of transmitting nodes.

Nodes are equipped with two radio interfaces. One is used for communication between cores and receivers, and the other is for general communication. More specifically, cores use a dedicated interface for transmitting multicast packets to their members, and the receivers use the same interface to receive packets from their cores. To allow for multiple cores to transmit simultaneously without interference, we assume that each core transmits on a different channel than any possibly interfering core. Therefore, the dedicated interface is set for transmitting and listening using a specific channel.

The problem of assigning non-interfering channels to cores is similar to the *graph coloring* problem in graph theory. Considering a core with transmission range of $r$, any core within distance $2r$ may interfere (i.e., even though the two cores may be out of range of each other, they may have members within range of both cores). In the context of MANETs, any distributed approximation to the graph coloring problem could be applied for assigning channels to the cores. Instead, we choose to limit

the total number of cores in the network to the maximum number of orthogonal channels available for the dedicated interface. Because nodes learn about all cores in the network (through the periodical core announcements), channels are assigned lexicographically.

To reduce delay, and to avoid retransmissions from nodes between the core and members located more than one hop away from the core, cores transmit multicast data packets with a larger power. The transmit power should be set so that the packet can be successfully received by any node up to $r$ hops from the core. Even though this approach increases energy consumption, it is expected to reduce the end-to-end delay and control overhead, because a single transmission from the core is supposed to reach all core members at the same time.

#### 6.1.1. Performance

We compare CHEMA against the *protocol for unified multicasting through announcements* (PUMA) [50]. PUMA has been shown to outperform two of the state of the art multicast routing protocols for MANETs (i.e., ODMRP [51] and MAODV [52]). PUMA presents the following characteristics: receiver-oriented; core based (one core per group); mesh-based, providing multiple routes from senders to receivers.

Only the core performs control packet flooding in PUMA. In CHEMA it is the same, but information about multiple cores are aggregated to reduce control overhead (PUMA also applies aggregation when flooding information about multiple groups).

We compared CHEMA against PUMA using the Qualnet$^{TM}$ [53] network simulator. Each simulation was run with four different random-number seeds. Timer values (i.e., *core announcements* in CHEMA, and *multicast announcements* in PUMA) were set to 3 s. Table 4 presents details about the simulation parameters.

Table 4
Simulation parameters

| | |
|---|---|
| Simulator | Qualnet$^{TM}$3.5 |
| Simulation time | 350 s |
| Terrain size | $1000 \times 1000$ m |
| Number of nodes | 50 |
| Node placement | Random |
| Mobility | Static |
| Radio range | 250 m |
| MAC protocol | 802.11 |
| Channel capacity | 2 Mbps |
| Data packet size | 512 Bytes |

In CHEMA, cores use a multiple of the regular radio range (i.e., for distance domination $r$, cores have a radio range of $r \cdot 250$ m) for the dedicated interface. The shorter radio range is used during the election process, and the larger radio range is used only by the cores when transmitting multicast data packets to their members. Because core members are at most $r$ hops distant, a larger radio range proportional to the distance parameter is enough for reaching all core members. DKR is executed every 16 s for core assignment. Because only static topologies are considered, the cores remain the same throughout the simulation.

Four performance metrics are evaluated:

- Packet delivery ratio: The ratio of the data packets delivered to the receivers to those data packets expected to be delivered (i.e., data packets sent times the number of receivers).
- Average end-to-end delay for data packets, including all possible delays caused by queuing at the interface, retransmission delays, and propagation and transfer times.
- Control overhead: The number of control packets transmitted per data packet delivered.
- Total overhead: The ratio of the total packets transmitted (i.e., control + data) to the data packets delivered.

For the simulation scenario, traffic load is varied across $\{1, 2, 5, 10, 25, 50\}$ packets/s. There are five senders, and 20 receivers for one multicast group. That is, the number of packets expected to be delivered varies from 20 packets/s to 1000 packets/s. Both senders and receivers are chosen randomly among the nodes in the network, and traffic load is equally distributed among all senders.

Even though DKR allows a myriad of scenarios for core selection, we consider just a few configurations for the purpose of simulation. For a 50-nodes network, at most eight cores are allowed (and there are eight orthogonal channels for the dedicated interface). Values 3 and 4 are tested for the *distance domination*, and at least one core is selected within the specified distance. These two configurations are presented in the graphs as CHEMA (1,3)-DS, and CHEMA (1,4)-DS, respectively. For the networks considered, three cores are elected in average in the first configuration, and fours cores in the second configuration.

Fig. 18 presents the results for packet delivery ratio. CHEMA delivers almost 100% of the data

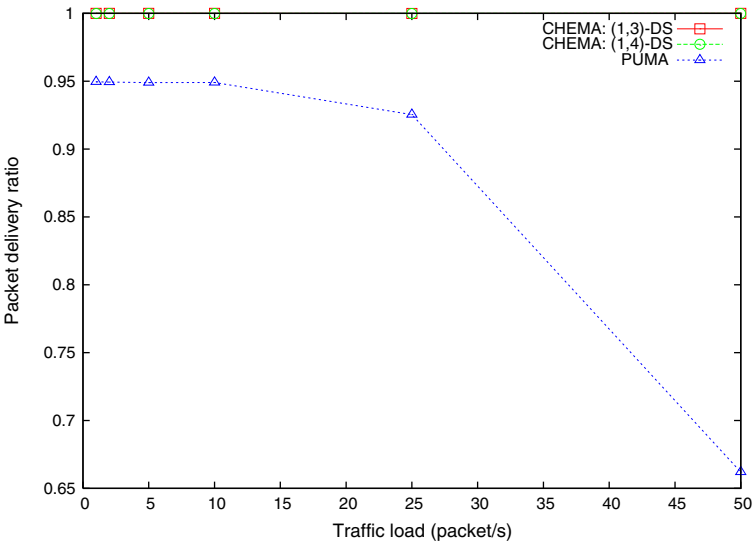


Fig. 18. Packet delivery ratio.

packets for all trafic loads considered. But PUMA cannot deliver more than 70% of packets for traffic load of 50 packets/s, due to increasing contention and collision of packets. On the other hand, mainly because CHEMA applies the *one shot* approach and the non-interfering channels for cores, once packets are transmitted by the core the packets are successfully received by the receivers.

For flows of up to 10 packets/s both protocols present similar results for the end-to-end delay (Fig. 19). While CHEMA has a small increase in terms of end-to-end delay for flows larger than 10 packets/s, PUMA experiences an exponential increase in average end-to-end delay. These results, together with the delivery ratio, indicate that CHEMA not only delivers more packets but does so incurring smaller end-to-end delays. This shows that it pays off sending packets to multiple cores and using a single transmission per packet from the cores to their members.

Even though CHEMA sends more control packets (mainly due to the election process) compared to

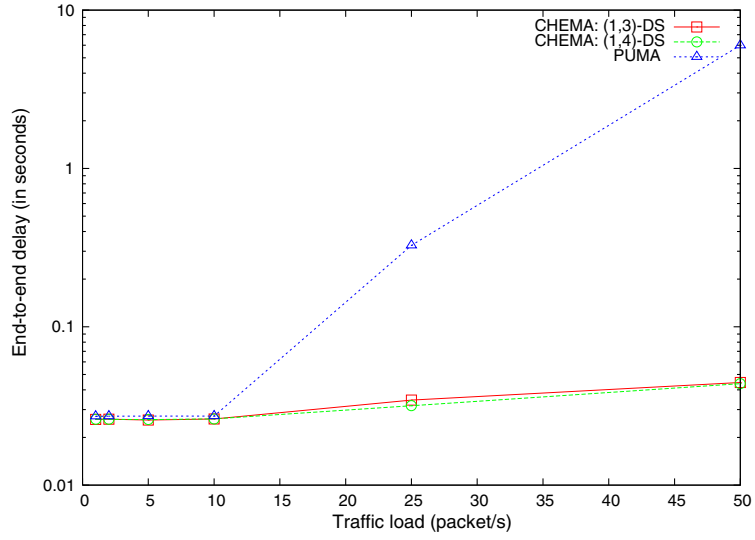


Fig. 19. End-to-end delay.

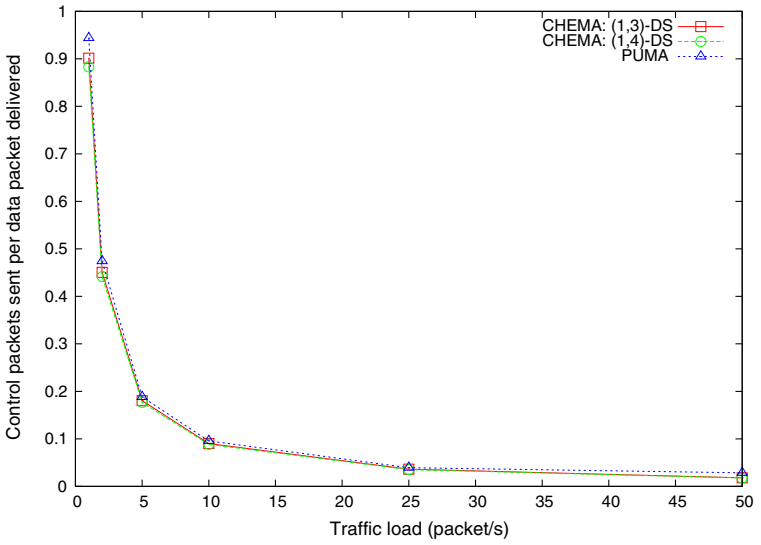


Fig. 20. Control overhead.

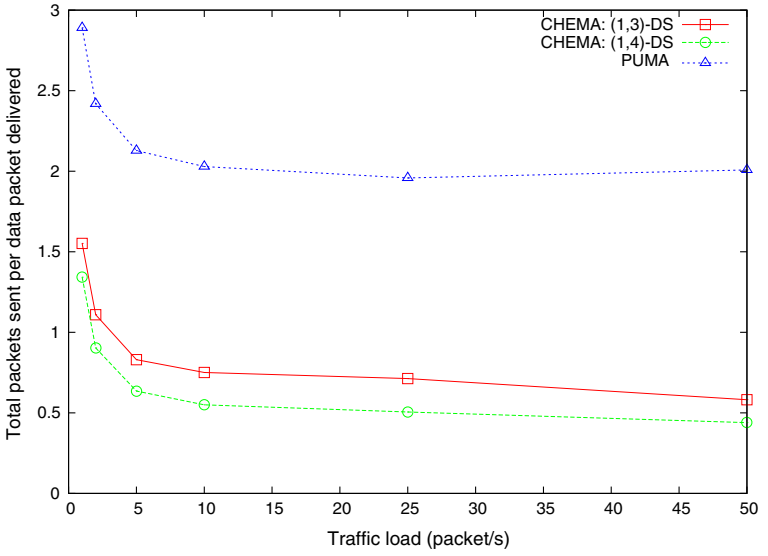


Fig. 21. Total overhead.

PUMA, both protocols present similar control overhead because CHEMA delivers more packets (as shown in Fig. 20). However, in terms of total overhead CHEMA incurs less than half total overhead compared to PUMA (Fig. 21). CHEMA requires fewer transmissions for every data packet delivered, specially because once the packets reach the targeted core it takes just one transmission per data packet to reach all core's receivers.

## 7. Conclusion

*Clustering* is the problem of building a hierarchy among nodes, and usually entails the computation of a *dominating set* (DS) of the network. The $(k,r)$-*dominating set* problem, $(k,r)$-DS, is defined as the problem of selecting a minimum cardinality vertex set $D$ (*dominating* nodes, also referred as *cluster-heads*) of a graph $G=(V,E)$,

such that every vertex $u$ not in $D$ is at a distance smaller than or equal to $r$ (*distance* domination parameter) from at least $k$ (*multiple* domination parameter) vertices in $D$. When applying $(k,r)$-DS for clustering, one can define the degree of redundancy (i.e., minimum number of *cluster-heads* per cluster) for clusters of variable radius (maximum distance from regular cluster members to their cluster-heads).

We have presented the first centralized and distributed solutions to the $(k,r)$-DS problem for arbitrary network topologies. The centralized solution, KR, is appropriate for wired networks, provides an approximation to the optimal solution, and is used as a lower bound when evaluating the performance of the distributed solution. The distributed solution, DKR, computes clusters with variable degree of redundancy, and variable radius. DKR selects *cluster-heads* through an election process. After the election, an approach is used for pruning redundant cluster-heads, and for connecting regular nodes to their cluster-heads.

We have conducted extensive simulations comparing KR against DKR. As expected, KR produces fewer cluster-heads than DKR. However, KR incurs too much *signaling overhead*, making it not appropriate for MANETs. There is a clear trade-off between efficiency (i.e., which approach reduces most the number of *cluster-heads*), and communication cost. While DKR usually produces more cluster-heads per cluster, it does not incur much control overhead.

As an example of an application of DKR, we proposed a novel multicast protocol named *core hierarchical election for multicasting in ad hoc networks* (*CHEMA*), which is designed to work in the context of multiple-channel and multiple-interface. CHEMA applies DKR for core election, with a dedicated interface using non-interfering channel for communication between cores and any multicast member. Because cores use a larger radio range for the dedicated interface, it requires just one transmission per data packet for any core member to receive the packet. CHEMA is compared against the *protocol for unified multicasting through announcements* (PUMA), which is one of the best performing multicast routing protocols for MANETs. CHEMA is shown to outperform PUMA in all aspects. CHEMA delivers more packets, incurs small end-to-end delays, and drastically reduces the total control overhead.

## Acknowledgements

We would like to thank Ravindra Vaishampayan for providing us with the code for PUMA. This work was supported in part by CNPq (Brazil), the Baskin Chair of Computer Engineering at UCSC, the National Science Foundation under grant CNS-0435522, the UCOP CLC under Grant SC-05-33, and by the US Army Research Office under grant No. W911NF-05-1-0246. Any opinions, findings, and conclusions are those of the authors and do not necessarily reflect the views of the funding agencies.

## Appendix A. Analysis of KR

In this section we show that KR computes a $(k,r)$-DS of any arbitrary network in polynomial time, yielding a dominating set of size at most $k \cdot \ln \Delta$ times the optimal.

### A.1. Correctness and time complexity

Let us consider an arbitrary graph $G = (V,E)$ of order $n = |V|$, represented using adjacency-lists. The initialization takes $O(n \cdot \log n)$ time, because each node in the graph needs to have its *Domin* and *total* parameters set, after which the node is inserted in the *HEAP*. Following the initialization, the selection of dominating nodes takes place. Dominating nodes are selected from the *HEAP*. The *while* loop is executed $O(n)$ times. The first *for* loop is executed $O(\Delta)$ times, and removing a node from the *HEAP* takes $O(\log n)$, what gives a $O(\Delta \cdot \log n)$ time for this internal loop. The second *for* loop is also executed $O(\Delta)$ times, and its internal *for* loop is executed $O(\Delta)$ times, what gives a $O(\Delta^2)$ time for the second *for* loop. Given that $\Delta \leqslant n$ (i.e., $\Delta$ increases as $r$ approaches the network diameter, and is at most $n$ when $r = \mathit{diameter}$), KR runs in $O(n^3)$ time.

**Theorem 3.** *KR correctly computes a $(k,r)$-dominating set of any connected graph $G = (V,N)$.*

**Proof.** Initially all nodes are in the *Heap*. While there are nodes in the *Heap*, the node at the top (i.e., with the largest *total* value) is selected as dominating. A node is also removed from the *Heap* when its *Domin* value is equal to zero (i.e., the node is $(k,r)$-dominated). Hence, when the *Heap* is empty, any node in the graph is either $(k,r)$-dominated or dominating. □

### A.2. Approximation ratio

**Theorem 4.** *KR, with parameters k and r, yields a dominating set of size at most* $\mathbf{k} \cdot \ln \Delta \cdot |\mathbf{OPT_{DS}}|$, *where* $\mathrm{OPT_{DS}}$ *is an optimal* $(k,r)$*-dominating set in the graph.*

**Proof.** Let $\mathrm{OPT_{DS}}$ be the set of vertices in an optimal $(k,r)$-dominating set. The set of vertices of $G$ dominated by vertex $i \in \mathrm{OPT_{DS}}$ is called $S_i$ (assuming that $i$ also belongs to $S_i$). If a vertex is dominated by more than $k$ dominating nodes, we arbitrarily put it in $k$ of such sets. The proof is based on amortized analysis. Each time a dominating node $y$ is selected, the operation has cost 1. This cost is equally distributed among all newly dominated vertices in $\mathfrak{N}_{r,y}$. We then prove that the total charge on the vertices belonging to a set $S_i$ (for any $i$) is at most $k \cdot \ln \Delta$. Because there are $|\mathrm{OPT_{DS}}|$ sets in the optimal solution, the theorem follows.

Assume that $N$ vertices are *newly dominated* when a node $s$ is selected as dominating. We charge each *newly dominated* vertex $\frac{1}{N}$, and a vertex can be *newly dominated* at most $k$ times. That is, while a vertex has been dominated less than $k$ times, it is considered as *newly dominated*. In case a node is dominated by more than $k$ dominating nodes the node is arbitrarily put in $k$ dominating sets.

We now prove the upper bound on the total charges to vertices belonging to a single set $S_i$. At each step, some vertices become *newly dominated*. The number of vertices not covered in $S_i$ is initially $u_0$, and finally drops to 0. In step $j$, the number of vertices dominated in $S_i$ is $u_j - u_{j+1}$. For simplicity, it is assumed that some vertices of $S_i$ are *newly dominated* at each step, decreasing the number of vertices not covered yet.

In the worst case, no two nodes in $S_i$ are *newly dominated* together. In this case, during step $j$ each node in $S_i$ can be charged at most $\frac{1}{u_j}$. Otherwise, node $i$ could be selected, because it would have dominated at least $u_j$ nodes (i.e., node $i$ would have the largest *total* value). Let $u_m = 0$ (i.e., at step $m$, all nodes in $S_i$ have already been covered). Adding up all the charges, assuming nodes are charged (i.e., *newly dominated*) at most $k$ times, we get:

$$k \cdot \sum_{j=0}^{m-1} \frac{u_j - u_{j+1}}{u_j} = k \cdot \sum_{j=1}^{m} \frac{1}{u_{j-1}} (u_{j-1} - u_j) \quad (A.1)$$

$$\leqslant k \cdot \sum_{j=1}^{m} H(u_{j-1}) - H(u_j) \quad (A.2)$$

$$= k \cdot (H(u_0) - H(u_m)) \quad (A.3)$$

$$= \mathbf{k} \cdot \ln \boldsymbol{\Delta}, \quad (A.4)$$

where $H(d) = \sum_{i=1}^{d} \frac{1}{i} = \ln d + \mathrm{O}(1), H(b) - H(a) \geqslant \frac{b-a}{b}$, $H(0) = 0$, and assuming the fact that $u_0 \leqslant \Delta$. □

## Appendix B. Analysis of DKR

To prove the correctness of DKR, we have to show that it is *safe* (i.e., the algorithm computes a $(k,r)$-DS of the network), and that it is *live* (i.e., it completes within a finite period of time).

**Lemma 5.** *Phase one of DKR has time complexity of* $\mathrm{O}(n\delta r)$, *where n is the number of nodes in the network,* $\delta$ *is the largest node degree, and r is the distance parameter.*

**Proof.** During each round, nodes send messages to all their one-hop neighbors. *Phase one* takes $r$ rounds. Assuming a network of $n$ nodes, and that nodes have at most $\delta$ links, *phase one* of DKR requires $\mathrm{O}(n\delta r)$ messages to complete. Therefore, the time complexity of *phase one* is $\mathrm{O}(n\delta r)$. □

**Lemma 6.** *After r rounds of successful transmission of message* **m**, *the message is propagated up to r hops away from the originating node.*

**Proof.** This can be proved by induction on the distance $d$ from the originating node. The base case is when $d = 0$, and corresponds to the originating node $n_0$. Now consider a node $v$ at distance $r$ from $n_0$. A neighbor $u$ of node $v$ at distance $r - 1$ received the message. Therefore, node $u$ sends the message to all neighbors, including $v$. Eventually node $v$ receives the message. □

**Theorem 7.** *Phase one of DKR correctly computes a* $(k,r)$*-DS of any arbitrary connected graph* $G = (V,E)$.

**Proof.** We assume that nodes know their one-hop neighbors. The system is either synchronous or asynchronous. In the latter case, rounds are simulated by tagging advertisements with the round number. By Lemma 6, after $r$ rounds a node ID is propagated at most $r$ hops away. Because nodes advertise their $K \leqslant k$ known smaller IDs, after $r$ rounds every node $n \in V$ learns the $K \leqslant k$ nodes, $D'_n$, with smaller IDs located within distance $r$. Let us assume that $S$ is the set composed of *proper-elected*, *self-elected*, and unsatisfiable nodes

(i.e., a node $i$ is deemed unsatisfiable if $|D'_i| < k$), and that $R$ is the set of satisfiable non-*proper/self-elected* nodes (i.e., $R = V - S$). It follows that the set $D = \{\bigcup_{n\in R} D'_n + S\}$ is a $(k,r)$-DS of $G$. □

**Theorem 8.** *Phase two of DKR correctly connects dominated nodes to at least k dominating nodes at most r hops away.*

**Proof.** At the beginning of *phase two*, any *dominated* node $i$ advertises its list of elected nodes, $D'_i$, by locally broadcasting the list via an LA message. Any node $n$ with status *dominated*, or *gateway*, upon receiving an LA message from neighbor $m$, changes its status to *dominating* if the advertisement lists node $n$, implying that node $n \in D'_m$. In this case, node $n$ sends an NA message which is then propagated to all nodes within distance $r$ from node $n$. For every node $k$ in the LA message such that $m \ni D'_n$, a notification is sent to node $k$ if node $n$ is on the path to node $k$ (i.e., for node $n$ the next-hop to node $k$ is known, and it is not $m$). If this is the case, at least one neighbor of node $m$ has a route to node $k$, because nodes are elected based on the advertisements sent by one-hop neighbors during *phase one*. Eventually, the LA message sent by node $m$ reaches all its one-hop neighbors, including the nodes with routes to the nodes elected by node $m$. A Notification for node $n$, when necessary, is issued (initiated or relayed) just once by any node $i$. A *dominated* node relaying a notification changes its status to *gateway*. Once a *Notification* reaches the destination, if the status of the destination is not *dominating*, it changes its status to *dominating*, and advertises itself sending an NA message. If any node $i$ relaying an NA messages currently has fewer than $k$ validated entries in $D'_i$, then an entry with the validated node is inserted in $D'_i$. After a period of time equal to *Wait Period* (starting from the beginning of *phase two*), every node $i$ in the network checks the number of *validated* entries in $D'_i$. If node $i$'s status is *pending dominating*, and it has $l < k$ validated entries in $D'_i$, then node $i$ changes its status to *dominating*, and it sends an NA message. Otherwise, if node $i$'s status is not *dominating*, it sends a *Join* message to all its dominating nodes (validated or not). Any node $n$ receiving a *Join* message, for destination $d$, does not need to relay the message in case node $n$ had already initiated, or relayed, a *Join* or *Notification* message to node $d$ before. After all the *Notification*, or *Join*, messages have reached their destinations, the paths from dominated nodes to their respective dominating nodes are formed by nodes that are either *dominating* or *gateway*. Because all nodes must check their status after a finite period of time (i.e., *Wait Period*), and any non-*dominating* node $i$ must send *Join* messages to all its *dominating* nodes, $(k,r)$-coverage is guaranteed. □

## References


[1] O. Younis, S. Fahmy, Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks, Transactions on Mobile Computing 3 (4) (2004) 366–379.

[2] B. Chen, K. Jamieson, H. Balakrishnan, R. Morris, Span: an energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks, in: MobiCom'01: Proceedings of the 7th Annual International Conference on Mobile Computing and Networking, ACM Press, 2001, pp. 85–96.

[3] S. Bandyopadhyay, E. Coyle, An energy efficient hierarchical clustering algorithm for wireless sensor networks, in: IEEE INFOCOM, 2003, pp. 1713–1723.

[4] L. Bao, J.J. Garcia-Luna-Aceves, Topology management in ad hoc networks, in: MobiHoc'03: Proceedings of the 4th ACM International Symposium on Mobile Ad hoc Networking and Computing, ACM Press, 2003, pp. 129–140.

[5] X.-Y. Li, Ad Hoc Networking, IEEE Press, 2003 (Chapter: Topology control in wireless ad hoc networks).

[6] W. Heinzelman, Application-specific protocol architectures for wireless networks, Ph.D. thesis, Massachusetts Institute of Technology, 2000.

[7] Y.P. Chen, A.L. Liestman, J. Liu, Ad Hoc and Sensor Networks, Nova Science Publisher, 2004 (Chapter: Clustering algorithms for ad hoc wireless networks).

[8] M.R. Garey, D.S. Johnson, Computers and Intractability, Freeman, San Francisco, 1978.

[9] T.W. Haynes, S.T. Hedetniemi, P.J. Slater (Eds.), Fundamentals of Domination in Graphs, Marcel Dekker, Inc., 1998.

[10] D. Joshi, S. Radhakrishnan, C. Narayanan, A fast algorithm for generalized network location problems, in: Proceedings of the ACM/SIGAPP Symposium on Applied Computing, 1993, pp. 701–708.

[11] L. Kleinrock, F. Kamoun, Optimal clustering structures for hierarchical topological design of large computer networks, Networks 10 (1980) 221–248.

[12] P.F. Tsuchiya, The landmark hierarchy: a new hierarchy for routing in very large networks, in: SIGCOMM'88: Symposium Proceedings on Communications Architectures and Protocols, ACM Press, 1988, pp. 35–42.

[13] N. Shacham, J. Westcott, Future directions in packet radio architectures and protocols, Proceedings of the IEEE 75 (1) (1987) 83–99.

[14] G. Pei, M. Gerla, X. Hong, C. Chiang, A wireless hierarchical routing protocol with group mobility, in: Proceedings of the IEEE Wireless Communications and Networking Conference, vol. 3, 1999, pp. 1538–1542.

[15] G. Pei, M. Gerla, X. Hong, Lanmar: landmark routing for large scale wireless ad hoc networks with group mobility, in: MobiHoc'00: Proceedings of the 1st ACM International

Symposium on Mobile Ad hoc Networking and Computing, IEEE Press, 2000, pp. 11–18.
[16] T. Ballardie, P. Francis, J. Crowcroft, Core based trees (cbt), in: SIGCOMM'93: Conference Proceedings on Communications Architectures, Protocols and Applications, ACM Press, 1993, pp. 85–95.
[17] I. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, A survey on sensor networks, Communications Magazine 4 (8) (2002) 102–114.
[18] G. Chen, F. Nocetti, J. Gonzalez, I. Stojmenovic, Connectivity based *k*-hop clustering in wireless networks, in: HICSS'02: Proceedings of the 35th Annual Hawaii International Conference on System Sciences, IEEE Computer Society, 2002, pp. 2450–2459.
[19] M. Gerla, J.K. Taek, G. Pei, On-demand routing in large ad hoc wireless networks with passive clustering, in: Proceedings of the IEEE Wireless Communications and Networking Conference, 2000, pp. 100–105.
[20] J. Wu, F. Dai, Broadcasting in ad hoc networks based on self-pruning, in: IEEE INFOCOM, 2003, pp. 2240–2250.
[21] J. Wu, H. Li, On calculating connected dominating set for efficient routing in ad hoc wireless networks, in: DIALM'99: Proceedings of the 3rd International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications, ACM Press, 1999, pp. 7–14.
[22] R. Sivakumar, P. Sinha, V. Bharghavan, Cedar: a core-extraction distributed ad hoc routing algorithm, IEEE Journal on Selected Areas in Communications 17 (8) (1999).
[23] F. Dai, J. Wu, Distributed dominant pruning in ad hoc networks, in: Proceedings of the IEEE International Conference on Communications, 2003, pp. 353–357.
[24] J. Sucec, I. Marsic, An efficient distributed network-wide broadcast algorithm for mobile ad hoc networks, Tech. Rep. CAIP-248, Rutgers University, September 2000.
[25] I. Stojmenovic, M. Seddigh, J. Zunic, Dominating sets and neighbor elimination-based broadcasting algorithms in wireless networks, IEEE Transactions on Parallel and Distributed Systems 13 (1) (2002) 14–25.
[26] M.A. Spohn, J.J. Garcia-Luna-Aceves, Enhanced dominant pruning applied to the route discovery process of on-demand routing protocols, in: Proceedings of the 12th IEEE International Conference on Computer Communications and Networks, 2003, pp. 497–502.
[27] M.A. Spohn, J.J. Garcia-Luna-Aceves, Enhancing the route discovery process of on-demand routing in networks with directional antennas, Lecture Notes in Computer Science, vol. 3042, Springer-Verlag GmbH, 2004, pp. 1000–1011.
[28] W. Lou, J. Wu, On reducing broadcast redundancy in ad hoc wireless networks, IEEE Transactions on Mobile Computing 1 (2) (2002) 111–122.
[29] P.-J. Wan, K.M. Alzoubi, O. Frieder, Distributed construction of connected dominating set in wireless ad hoc networks, in: IEEE INFOCOM, 2002, pp. 1597–1604.
[30] H. Lim, C. Kim, Flooding in wireless ad hoc networks, Computer Communications 2 (3–4) (2001) 353–363.
[31] B. Das, R. Sivakumar, V. Bharghavan, Routing in ad hoc networks using a spine, in: ICCCN, 1997, pp. 34–39.
[32] B. Das, V. Bharghavan, Routing in ad-hoc networks using minimum connected dominating sets, in: ICC (1), 1997, pp. 376–380.
[33] T. Clausen, P. Jacquet, A. Laouiti, P. Muhlethaler, A. Qayyum, L. Viennot, Optimized link state routing protocol for ad hoc networks, in: Proceedings of the IEEE International Multi Topic Conference, 2001, pp. 62–68.
[34] M.A. Spohn, J.J. Garcia-Luna-Aceves, Improving broadcast operations in ad hoc networks using two-hop connected dominating sets, in: Proceedings of the IEEE Global Telecommunications Conference Workshops, 2004, pp. 147–152.
[35] A.D. Amis, R. Prakash, T.H.P. Vuong, D. Huynh, Max–min d-cluster formation in wireless ad hoc networks, in: IEEE INFOCOM, 2000, pp. 32–41.
[36] B. Clark, C. Colbourn, D. Johnson, Unit disk graphs, Discrete Mathematics 86 (1990) 165–177.
[37] B. Liang, Z.J. Haas, Virtual backbone generation and maintenance in ad hoc network mobility management, in: IEEE INFOCOM, 2000, pp. 1293–1302.
[38] D. Zappala, A. Fabbri, V. Lo, An evaluation of shared multicast trees with multiple cores, Telecommunication Systems 19 (3–4) (2002) 461–479.
[39] E.M. Belding-Royer, Multi-level hierarchies for scalable ad hoc routing, Wireless Networks 9 (5) (2003) 461–478.
[40] R. Ramanathan, M. Steenstrup, Hierarchically organized, multihop mobile wireless networks for quality-of-service support, Mobile Networks and Applications 3 (1) (1998) 101–119.
[41] F. Dai, J. Wu, On constructing *k*-connected *k*-dominating set in wireless networks, in: Proceedings of the IEEE IPDPS, 2005.
[42] Z. Shi, P.K. Srimani, A new adaptive distributed routing protocol using d-hop dominating sets for mobile ad hoc networks, in: Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA), 2004.
[43] Y. Fernandess, D. Malkhi, K-clustering in wireless ad hoc networks, in: Proceedings of the ACM Workshop on Principles of Mobile Computing (POMC), 2002.
[44] P. Krishna, N.N. Vaidya, M. Chatterjee, D.K. Pradhan, A cluster-based approach for routing in dynamic networks, in: ACM SIGCOMM Computer Communication Review, 1997, pp. 49–64.
[45] D. Kim, S. Ha, Y. Choi, *k*-hop cluster-based dynamic source routing in wireless ad-hoc packet radio networks, in: Proceedings of the IEEE VTC, 1998, pp. 224–228.
[46] E. Gafni, Round-by-round fault detectors (extended abstract): unifying synchrony and asynchrony, in: PODC'98: Proceedings of the Seventeenth Annual ACM Symposium on Principles of Distributed Computing, ACM Press, 1998, pp. 143–152.
[47] M. Mosko, J.J. Garcia-Luna-Aceves, A self-correcting neighbor protocol for mobile ad hoc wireless networks, in: Proceedings of the IEEE International Conference on Computer Communications and Networks, 2002, pp. 556–560.
[48] M. Gerla, J. Tsai, Multicluster, mobile, multimedia radio network, ACM Baltzer Journal of Wireless Networks 1 (3) (1995) 255–265.
[49] J. Wu, F. Dai, A generic distributed broadcast scheme in ad hoc wireless networks, in: Proceedings of the IEEE International Conference on Distributed Computing Systems, 2003, pp. 460–467.
[50] R. Vaishampayan, J.J. Garcia-Luna-Aceves, Efficient and robust multicast routing in mobile ad hoc networks, in: Proceedings of the IEEE International Conference on Mobile Ad-hoc and Sensor Systems, 2004, pp. 304–313.

[51] S.-J. Lee, M. Gerla, C.-C. Chiang, On-demand multicast routing protocol, in: Proceedings of the IEEE Wireless Communications and Networking Conference, 1999, pp. 1298–1302.
[52] E.M. Royer, C.E. Perkins, Multicast operation of the ad-hoc on-demand distance vector routing protocol, in: MobiCom'99: Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking, ACM Press, 1999, pp. 207–218.
[53] Scalable Network Technologies, qualnet simulator, Available from: <http://www.scalable-networks.com>, 2003.

**Marco Aurélio Spohn** received the B.S. and M.Sc. degrees in Computer Science from Universidade Federal do Rio Grande do Sul (UFRGS), Brazil, in 1995 and 1997, respectively, and the Ph.D. degree in Computer Science from University of California Santa Cruz (UCSC), USA, in 2005. He is currently an Associate Professor at the Department of Computer Science from Universidade Federal de Campina Grande (UFCG) (Federal University of Campina Grande), Brazil. His current research interest is the design of algorithms and protocols for wireless ad hoc networks.

**J.J. Garcia-Luna-Aceves** received the B.S. degree in Electrical Engineering from Universidad Iberoamericana (''La Ibero''), Mexico City, Mexico in 1977. M.S. and Ph.D., Electrical Engineering, University of Hawaii at Manoa in 1980 and 1983, respectively. He holds the Baskin Chair of Computer Engineering. Prior to joining UCSC in 1993, he was a Center Director at SRI International in Menlo Park, California. He first joined SRI as an SRI International Fellow in 1982. He was a Visiting Professor at Sun Labs in Menlo Park, California in 1999, and was a Principal of Protocol Design for NOKIA from 1999 to 2003. He is a Principal Scientist at the Palo Alto Research Center (PARC). He has coauthored the book Multimedia Communications: Protocols and Applications (Prentice-Hall), and has published six patents and more than 270 papers on computer communication in journals and conferences. His current research interest is the analysis and design of algorithms and protocols for computer communication. At UCSC, he leads the Computer Communication Research Group (CCRG), which is home to many research projects that focus on wireless networks and inter-networking, and which are currently sponsored by the National Science Foundation (NSF), the UC Office of the President (UCOP), the Defense Advanced Research Projects Agency (DARPA), the US Army Research Office (ARO), the US Air Force Office of Scientific Research (AFOSR), and industry.