Extended Multipoint Relays to Determine Connected Dominating Sets in MANETs *

Jie Wu
Department of Computer Science and Engineering
Florida Atlantic University
Boca Raton, Florida 33431

Wei Lou
Department of Computing
Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong

Fei Dai
Department of Electrical and Computer Engineering
North Dakota State University
Fargo, ND 58105

*This work was supported in part by NSF grants ANI 0073736, CCR 0329741, CNS 0434533, CNS 0422762, and EIA 0130806. Wei Lou's work was supported in part by the HK PolyU ICRG A-PG53. The preliminary version appeared in IEEE SECON 2004. Contact address: jie@cse.fau.edu

Abstract

Multipoint relays (MPR) [18] provide a localized and optimized way of broadcasting messages in a mobile ad hoc network (MANET). Using partial 2-hop information, each node chooses a small set of forward neighbors to relay messages and this set covers the node's 2-hop neighbor set. These selected forward nodes form a connected dominating set (CDS) to ensure full coverage. Adjih, Jacquet, and Viennot [1] later proposed a novel extension of MPR to construct a small CDS, and it is source-independent. In this paper, we provide several extensions to generate a smaller CDS using complete 2-hop information to cover each node's 2-hop neighbor set. We extend the notion of coverage in the original MPR. We prove that the extended MPR has a constant local approximation ratio compared with a logarithmic local ratio in the original MPR. In addition, we show that the extended MPR has a constant global probabilistic approximation ratio, while no such ratio exists in the original MPR and its existing extensions. The effectiveness of our approach is confirmed through a simulation study.

Keywords: Approximation ratio, broadcasting, connected dominating set (CDS), heuristic solutions, mobile ad hoc networks (MANETs), multipoint relays (MPR).

1 Introduction

A mobile ad hoc network (MANET) is a wireless network that comprises of mobile computing devices for wireless communication, without the help of fixed infrastructures. Wireless interfaces pose a unique challenge in designing efficient broadcasting in MANETs.The simplest way to perform a broadcast is based on the following *blind flooding* rule: a node retransmits the message only once. The blind flooding may cause excessive redundancy and results in channel contention and message collision (also called the *broadcast storm problem* [15]).

Efficient broadcasting in MANETs can be formulated by identifying a small *connected dominating set* (CDS) in the network where only the nodes in the set relay the message (also called *CDS rule*). A *dominating set* (DS) is a subset of nodes in the network that assures every node is either in the subset or a neighbor of a node in the subset. A DS is called a CDS if the subgraph induced by the DS is connected. In Figure 1 (b), a sample network with 15 nodes is shown with double circled nodes forming a CDS. Many existing works on finding a small CDS are not suitable for MANETs, since they rely on either global information (such as a global network topology) or global infrastructure (such as a spanning tree). In a MANET, network topology changes frequently

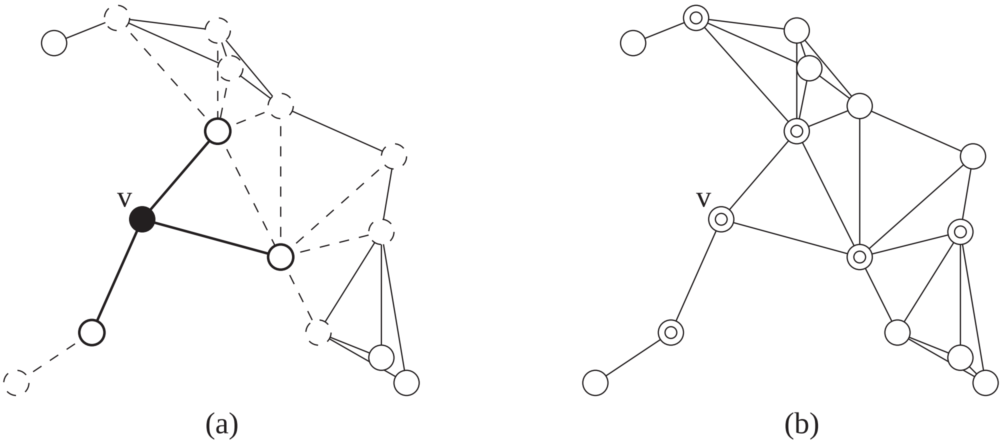


Figure 1: A sample network: (a) 2-hop information of node v and (b) double circled nodes forming a CDS.

and, hence, a global information/infrastructure approach may not be *combinatorially stable*. In a combinatorially stable system, the propagation of all topology updates is sufficiently fast to reflect the topology changes.

The k-hop localized approach is a solution to ensure the combinatorially stable property for a small k in MANETs. In this approach, each node determines its status and/or the status of neighbors (forward or non-forward) based on its k-hop information (such as local network topology within k hops). In general, a k-hop neighbor set of a given node consists of nodes that are at most k hops away from this node. If the neighborhood information is collected via periodical "Hello" message exchanges, it takes k rounds for each node to collect its k-hop neighbor set. However, k rounds can only collect partial topology information for the k-hop neighbor set (or simply, *partial k-hop information*). Specifically, links between k-hop neighbors are not included. Figure 1 (a) shows the partial 2-hop information of v after 2 rounds of "Hello" message exchanges with thick lines for links in the first "Hello" message exchanges and dashed lines for links in the second "Hello" message exchanges. To collect *complete k-hop information*, each node needs to either exchange positional information (obtained through GPS or non-GPS localization methods) or perform $k + 1$ rounds of "Hello" message exchanges. It is clearly impossible to collect up-to-date network topology information for a large k; therefore, k is usually a small integer such as 2 or 3 in MANETs.

Multipoint relays (MPR) [18] are a special 2-hop localized approach, in which each forward node determines the status of its neighbors based on its partial 2-hop information through node coverage. It should be stressed that in the MPR each node does not determine its forward status. Instead each forward node determines forward status for each of its neighbors. Specifically, each forward node selects a subset of 1-hop neighbors to cover its 2-hop neighbor set (and this node is also called *selector* for its neighbors). A link state routing protocol based on MPR, OLSR [6], is part of the standardization efforts along with DSR [10] and AODV [17]. In OLSR, MPR is used as a virtual backbone to disseminate link state information to the entire network.

The original MPR [18] is source-dependent (also called broadcast-dependent); that is, the forward node set is determined during a broadcast process and is dependent on the source of the broadcast and communication latency. Adjih, Jacquet, and Viennot [1] later proposed a novel source-independent (also called broadcast-independent) MPR. Specifically, the forward node set is determined before any broadcast process and is constructed based on the MPR following two simple rules. In [24], Wu enhanced the source-independent MPR through several modifications where the notion of free neighbor is introduced. However, neither the original MPR nor its extensions have local or global constant approximation ratio.

In this paper, we provide several extensions to generate a smaller forward node set using complete 2-hop information to cover each node's 2-hop neighbor set. Our focus will be on the source-independent MPR, which is widely used for applications that require a relatively stable CDS (e.g., to form a virtual backbone for efficient routing). In addition, we extend the notion of coverage in the original MPR. Specifically, each node v selects a set of node pairs (u, w) with u being a 1-hop neighbor and w a 2-hop neighbor to cover v's 2-hop neighbor set. Our results show that, using complete 2-hop information instead of partial 2-hop information, the extended MPR has a constant local approximation ratio compared with a logarithmic local ratio in the original MPR. We further prove that the extended MPR has a constant global probabilistic approximation ratio, while no such ratio exists in the original MPR. The effectiveness of our approach is confirmed through a simulation study.

The main contributions of this paper are as follows: (1) We propose an efficient extension of the MPR that forms a substantially smaller source-independent CDS than the original MPR and its extensions. (2) We show that our selection of forward pairs based on complete 2-hop information has a $O(1)$ local approximation ratio. (3) We prove that the (global) CDS derived from our extension has an expected approximation ratio of $O(1)$, and show that such an approximation ratio does not exist in the original MPR. (4) We conduct a simulation study to evaluate the average performance of our extension.

The remainder of the paper is organized as follows. Section 2 provides preliminaries on the general graph model, and reviews MPR with its existing extensions and the corresponding CDS selection algorithms. Section 3 proposes the extended MPR. Section 4 gives a local upper bound of the proposed algorithm. In Section 5, we prove a probabilistic global bound of the new algorithm and show that a similar upper bound does not exist in the original MPR. Section 6 provides some simulation results. The related work is discussed in Section 7. In Section 8, we draw our conclusion.

2 Preliminaries

This section reviews the basic graph model for MANETs, discusses MPR-based broadcasting and its existing extensions, and overviews some existing algorithms for CDS selection in MPR-based broadcasting.

2.1 Graph model

A MANET is represented by a graph $G = (V, E)$, where the node set V represents a set of wireless mobile nodes and the edge set E represents a set of bi-directional links between the neighboring nodes. We assume each node has a distinct ID. A node ID can be any unique attribute of a node, such as its network or MAC layer address. G is usually a *unit disk graph*, where two nodes are considered neighbors if and only if their geographic distance is no more than a given transmission range r (as shown in Figure 1); however, G can also be a non-unit disk graph (as in Figure 3). Let $N(v)$ denote node v's neighbor set (including v), and $N(V) = \bigcup_{v \in V} N(v)$ is the set of all nodes that are in V or have a neighbor in V. V *covers* U if $U \subseteq N(V)$.

The *(partial) k-hop information* of a node v is a subgraph $G_k(v) = (N_k(v), E_k(v))$ of the network G, where $N_k(v)$ is the k-hop neighbor set and $E_k(v)$ the partial k-hop link set of v. Specifically, $N_0(v) = \{v\}$ and $N_k(v) = \bigcup_{u \in N_{k-1}(v)} N(u)$ for $k \geq 1$. $E_k(v) = \{(u, w) | u \in N_{k-1}(v) \wedge w \in N_k(v)\}$ includes links among node in $N_k(v)$, but excludes those between two nodes that are exactly k hops from v. $G_k(v)$ is collected at v through k rounds of "Hello" message exchanges. As shown in Figure 3 (a), if v has 1-hop information, then it knows all its neighbors, but not the links between these neighbors. The *complete k-hop information* of a node v is a subgraph $G'_k(v) = (N_k(v), E'_k(v))$, where $E'_k(v) = \{(u, w) | u, w \in N_k(v)\}$ is the complete k-hop link set of

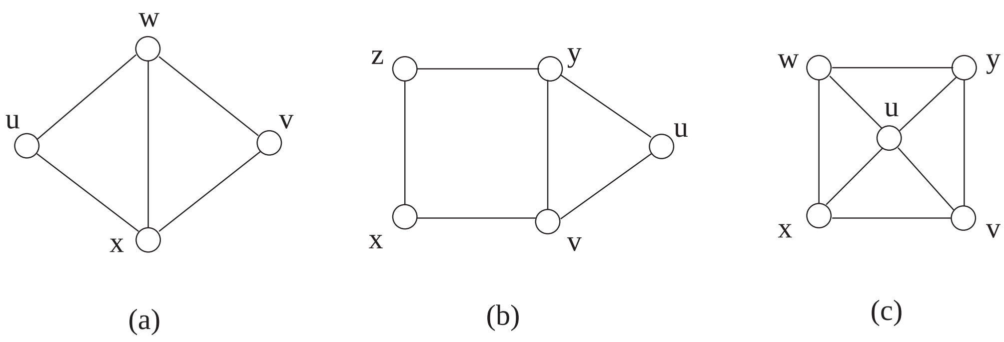


Figure 2: Three sample networks.

v. Such information is normally obtained from $(k+1)$ rounds of "Hello" message exchanges, or from k rounds of message exchanges if positional information is available.

2.2 MPR-based broadcasting

Multipoint relays (MPR). In multipoint relays (MPR) [18], each node v maintains 2-hop subgraph $G_2(v) = (N_2(v), E_2(v))$. Node v selects a small forward node set, $C(v)$, from its 1-hop neighbor set $N(v)$ to cover its 2-hop neighbor set $N_2(v)$; that is, $C(v) \cup \{v\}$ is a CDS for $G_2(v)$. $C(v)$ is also called the *coverage set* for v. When u is selected by v as a forward node, v is called the *selector* of u. Note that several selectors may exist for a particular node. A forward node may or may not actually retransmit the message; its actual status is determined by the following MPR rule [18], a special CDS rule.

- **MPR rule**: a node retransmits the message once if the first message received is from a selector.

The collection of nodes that have retransmitted the message plus the source node form a CDS called a forward node set.

Source-independent MPR. The original MPR is source-dependent. Adjih, Jacquet, and Viennot [1] proposed a novel extension of the MPR to construct a small CDS that is source-independent. The source-dependent approach depends on a particular broadcast; the resultant forward node set depends on many factors, such as the locations of neighbors, node priority, message propagation delay, and back-off delay. The source-independent approach does not depend on a particular broadcast; the resultant forward node set depends only on local topology and node priority. In addition, the forward node set is generic, so it can be used for any broadcast.

A node belongs to a CDS if

- **Rule 1**: the node has a smaller ID than all its neighbors.
- **Rule 2**: the node is a forward node selected by its smallest ID neighbor.

Note that node ID, $id(v)$, is not the only way to define the priority of node v. Other priority schemes exist, for example, $(|N(v)|, id(v))$, where node degree is the primary priority measure and node ID is used in case of a tie in node degree.

Enhanced MPR. Wu [24] observed two potential drawbacks in the source-independent MPR and proposed two extensions: (1) Rule 1 is "useless" in many occasions; that is, the node selected based on Rule 1 is not essential for a CDS. (2) The original MPR forward node selection (Algorithm 1) does not take advantage of Rule 2.

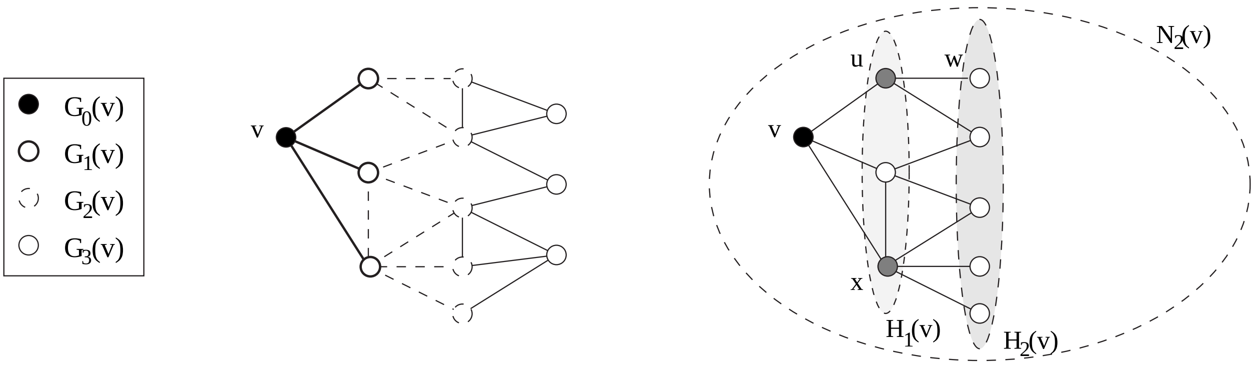
In Figure 2 (a), u and v are selected based on Rule 1; however, they are useless. In fact, node w alone is sufficient for a CDS. Similarly, u in Figure 2 (b) is useless. On the other hand, we might have to include some smallest ID nodes even if they are not selected by any of their neighbors as forward nodes. In Figure 2 (c), suppose node u is not selected by any of its neighbors, u has to be included (as it is selected by Rule 1), because any forward node selected by a node other than u will be ignored based on Rule 2.

Based on the first observation, Rule 1 was enhanced as follows:

- **Enhanced Rule 1**: the node has a smaller ID than all its neighbors and has two unconnected neighbors.

The Enhanced Rule 1 together with the original Rule 2 will generate a CDS under all cases except complete graphs.

In Figure 2 (b), we assume that v selects x as its forward node. Based on Rule 2, since v is the smallest ID neighbor of x, x cannot ignore v's choice. On the other hand, if v chooses y, since v is not the smallest ID neighbor of y, v's choice will be ignored by y. Therefore, forward node y



(a) (b)

Figure 3: (a) Partial k-hop information. (b) In MPR, each node v selects a few 1-hop neighbors (u and x in this example) to cover its 2-hop neighbor set. Links between nodes in $H_2(v)$ are not visible to v.

comes for "free" for v. That is, the inclusion of y does not increase the size of the global forward node set. Generally speaking, node u is a *free neighbor* of v if v does not have the smallest ID among u's neighbors.

2.3 CDS selection in MPR and its extensions

CDS selection in MPR. Let $H_k(v) = N_k(v) - N_{k-1}(v)$ denote the nodes k hops away from v. A simple greedy algorithm (Algorithm 1) for computing $C(v)$ (initially empty) at v is shown as Algorithm 1 [18]. This selection of $C(v)$ is illustrated by Figure 3 (b). Note that $N(v)$ is covered when v transmits. Therefore, $H_2(v)$ $(= N_2(v) - N(v))$ is used instead of $N_2(v)$. A neighbor u is *essential* if a node in $H_2(v)$ is solely covered by u. In Algorithm 1, essential neighbors are first included. Then neighbors with higher degrees (i.e., covering more uncovered 2-hop neighbors) are selected until $H_2(v)$ is covered. If there are two neighbors with the same degree, either one can be selected.

Algorithm 1 Greedy algorithm for selecting $C(v)$ [18]

1: Add $u \in H_1(v)$ to $C(v)$, if there is a node in $H_2(v)$ covered only by u. Any node in $H_2(v)$ that is not covered by $C(v)$ is called an uncovered node.
2: Add $u \in H_1(v)$ to $C(v)$, if u covers the largest number of uncovered nodes in $H_2(v)$.

Suppose the following coverage sets are selected based on the above greedy algorithm in Figure 2 (b): $C(u) = \{v, y\}$, $C(v) = \{x\}$, $C(x) = \{v\}$, $C(y) = \{z\}$, and $C(z) = \{y\}$. Collectively nodes v, x, y, and z form a CDS. As specified in the MPR rule, the actual set of forward nodes for a particular broadcast uses only a subset, and it depends on the location of the source and communication latency. For example, if v is the source and node x receives the first message from v, then x is a forward node. Also, if nodes w and y receive their first message from x and v, respectively, none of them will forward the message. Therefore, $\{v, x\}$ forms a CDS for this case. However, if node z is the source and node y receives the first message from z, then $\{y, z\}$ forms a CDS.

CDS selection in source-independent MPR. Applying Rule 1 and Rule 2 to Figure 2 (b) with the coverage set of each node selected as above, $\{v, x, y, z\}$ is changed to $\{u, v, x, y\}$ for a CDS. Compared with the set derived from the original MPR, node z is no longer in the final CDS since it is selected by y (which does not have the smallest ID among z's neighbors). Node u is included since it has a smaller ID than all its neighbors. The correctness of source-independent MPR is given in [1].

CDS selection in enhanced MPR. The greedy algorithm for selecting $C(v)$ (Algorithm 1) can be enhanced by giving free neighbors higher priorities. The modified greedy algorithm is shown in Algorithm 2. Simulation results in [24] show that this extension is effective when the network is sparse. The Enhanced Rule 1 is effective when the network is dense. Combining the Enhanced Rule 1 and the modified greedy algorithm (Algorithm 2) at each node v, the result is effective for both sparse and dense networks.

Algorithm 2 Modified greedy algorithm for selecting $C(v)$ [24]

1: Add all free neighbors to $C(v)$.
2: Follow steps 1. and 2. of Algorithm 1.

3 Proposed Approach

In this section, we first give a refined definition of coverage, where a node, using complete 2-hop information, can select both 1-hop and 2-hop neighbors to cover its 2-hop neighbor set. An enhanced Rule 2 for source-independent CDS construction is proposed based on the directed (1-hop) and indirected (2-hop) coverage. We then prove the correctness of the extended scheme.

A greedy algorithm is used to select such an extended coverage set. Finally, we show several examples.

3.1 Direct and indirect selector

The proposed approach is motivated by the case shown in Figure 2 (b). Suppose the current node is u. In the original MPR or its extensions, both y and v need to be selected to cover u's 2-hop neighbors z and x. However, z falls into the 2-hop neighbor set of v. That is, z can be covered by v via x when v calculates its forward node set. Motivated by this example, in our extended MPR, node v selects a pair of nodes to cover $H_2(v)$. Specifically, node v repeatedly adds a node pair (u, w) to $C(v)$ at each time, where $u \in H_1(v)$ and w is an uncovered node in $H_1(u) \cap H_2(v)$, until all nodes in $H_2(v)$ are covered. We now give an extended notion of coverage and selector:

Definition 1 (Coverage and Selector) *When node v adds two nodes $u \in H_1(v)$ and $w \in H_1(u) \cap H_2(v)$ to $C(v)$, neighbors of u are* directly covered *and neighbors of w are* indirectly covered*. In addition, v is a* direct selector *of u and an* indirect selector *of w.*

In the example of Figure 2 (b), when node u selects a pair (v, x) for $C(u)$, among 2-hop neighbors of u, x is directly covered by v and z is indirectly covered by v via x. In this case, u is a direct selector for v (to cover x) and an indirect selector for x (to cover z).

In the proposed approach, each node v still covers its 2-hop neighbor set, but uses complete 2-hop information. In fact, the only additional information used is about connections between any two 2-hop neighbors. We have then the following Enhanced Rule 2:

- **Enhanced Rule 2**: node u is a forward node if it is directly selected by a node in $H_1(u)$ that has the smallest ID in $H_1(u)$; node w is a forward node if it is indirectly selected by a node in $H_2(w)$ that has a smaller ID than all nodes in $H_1(w)$.

The following theorem shows that the Enhanced Rules 1 and 2 guarantee a CDS for a given connected graph.

Theorem 1 (Correctness of the Extended MPR) *If the given connected graph is not a complete graph, the set of forward nodes selected by the Enhanced Rule 1 and Enhanced Rule 2 forms a CDS.*

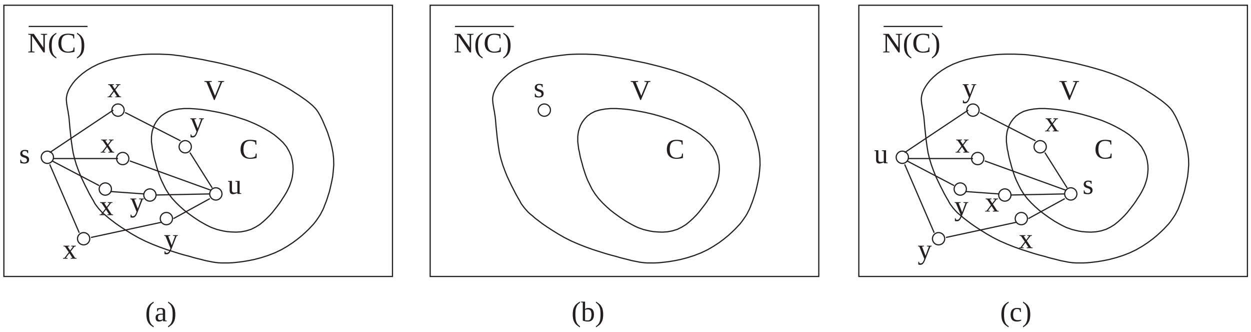


Figure 4: An illustration of the proof for Theorem 1: (a) Case 1, (b) Case 2, and (c) Case 3.

Proof: Assume that the graph is not a complete graph; we first show that there exists at least one node in the forward node set. Let c be the node with the smallest ID in the network. If all other nodes are neighbors of c, at least two neighbors are not directly connected. Based on the Enhanced Rule 1, c is selected. If there exists a node that is not a neighbor of c, c will designate a neighbor $c^{'}$ for relaying. Since c is the smallest ID node, $c^{'}$ is selected based on the Enhanced Rule 2. Let C be the connected component in the forward node set that contains the smallest ID node c and/or its designated neighbor $c^{'}$. We prove that C itself is a dominating set (DS).

We prove by contradiction. If C is not a DS, there must exist some nodes that are not in $N(C)$, i.e., $\overline{N(C)}$ is not empty. Let V be the set of nodes that have at least one neighbor in C and at least one neighbor in $\overline{N(C)}$. V cannot be empty, since the network is connected. Also, $V \cap C = \phi$. Consider the smallest ID node s in $N(V)$. s belongs to one of the following three cases (see Figure 4):

- Case 1: Assume s is in $\overline{N(C)}$ (which implies $s \notin V$). Since $s \in N(V)$ and $s \notin V$, there exists a neighbor v of s in V. (Note that in general when $s \in N(V)$, s may not have a neighbor in V.) Let u be a neighbor of v in C. Consider now the forward set for s. As u is a 2-hop neighbor of s, based on Definition 1, s has the following four choices to cover u:

  1. $s \rightarrow x(\in V) \rightarrow u$
  2. $s \rightarrow x(\in V) \rightarrow y(\in C \cap N(V)) \rightarrow u$
  3. $s \rightarrow x(\in \overline{N(C)} \cap N(V)) \rightarrow y(\in V) \rightarrow u$
  4. $s \rightarrow x(\in V) \rightarrow y(\in V) \rightarrow u$

  In the first case, s directly selects $x \in V$ to directly cover u; in the second case, s directly selects $x \in V$ to indirectly cover u (via y); in the third case, s indirectly selects $y \in V$ (via x) to directly cover u; in the forth case, s directly selects $v \in V$ to indirectly cover u (via y). In all these cases, s has the smallest ID among $N(V)$ which includes $N(x)$ or $N(y)$. With the Enhanced Rule 2, $x \in V$ or $y \in V$ will be selected in the forward node set C which contradicts $V \cap C = \phi$.

- Case 2: Assume s is in $N(C) \cap V$. Based on the Enhanced Rule 1, s is selected since its ID is smaller than that of all its neighbors. In addition, s has two unconnected neighbors, one in $\overline{N(C)}$ and one in C. This results in the contradiction that $V \cap C = \phi$.

- Case 3: Assume s is in $N(C) - V$. Let v be a neighbor of s in V, and let u be a neighbor of v in $\overline{N(C)}$. Consider now the forward set for s. Since u is a 2-hop neighbor of s, s has the following four choices to cover u:

  1. $s \rightarrow x(\in V) \rightarrow u$
  2. $s \rightarrow x(\in V) \rightarrow y(\in \overline{N(C)} \cap N(V)) \rightarrow u$
  3. $s \rightarrow x(\in C \cap N(V)) \rightarrow y(\in V) \rightarrow u$
  4. $s \rightarrow x(\in V) \rightarrow y(\in V) \rightarrow u$

  The rest of the proof is similar to Case 1.

In all cases, we reach a contradiction that $V \cap C = \phi$. Therefore, C has to be a DS. □

3.2 Extended greedy algorithm

The previous subsection provides general rules for the extended MPR. This subsection deals with efficient implementation of these rules (in terms of reducing CDS). The efficient implementation in this section does not change local and global asymptotic bounds discussed in the next two sections. With the Enhanced Rule 2, we extend the notion of a free neighbor to *1-hop free neighbor* and *2-hop free neighbor* as follows:

Definition 2 (Free Neighbor) *Node u is a 1-hop free neighbor of v if u is in $H_1(v)$ and v's ID is not the smallest ID in $H_1(u)$. Node w is a 2-hop free neighbor of v if w is in $H_2(v)$ and v's ID is larger than at least one node ID in $H_1(w)$.*

The greedy algorithm can then use these free neighbors for neighbor coverage without any cost. In the extended greedy algorithm (Algorithm 3), two nodes, u and w, as a pair are selected with each selection operation performed by the current node v, where u is a 1-hop neighbor of v and w is a 2-hop neighbor of v which is also a 1-hop neighbor of u. We introduce the concepts of "cost" and "yield" to measure the quality of each selection.

Definition 3 (Cost and Yield) *The "cost" of a selection operation is the number of selected nodes that are not free neighbors in the selection. The "yield" of a selection operation is the total number of uncovered nodes that are covered by the selection divided by the cost of the selection.*

Note that each node v knows its complete 1-hop and partial 2-hop free neighbors because v has complete 2-hop information, which includes the neighbor set of each of its 1-hop neighbors, and the partial neighbor set ($H_1(w) \cap N_2(v)$) for each 2-hop neighbor w. Also, as the cost of a selection can be zero, the corresponding value of the yield will be infinite.

Algorithm 3 Extended greedy algorithm for selecting $C(v)$

1: Add all pairs of 1-hop free neighbor u and uncovered 2-hop free neighbor w to $C(v)$ and remove all their covered nodes from $H_2(v)$.

2: Add a pair of nodes $u \in H_1(v)$ and uncovered $w \in H_1(u) \cap H_2(v)$ to $C(v)$ that gives the highest yield in $H_2(v)$.

The major modification in Algorithm 3, compared with Algorithm 2, is that a 2-hop neighbor w of v can be indirectly selected to cover other 2-hop neighbors. That is, a 1-hop neighbor u directly covers $H_1(u) \cap H_2(v)$ and u indirectly covers $H_1(w) \cap H_2(v)$ via w. Also, w always exists as long as $H_2(v)$ is not empty and is included even if it does not "contribute" additional coverage beyond what u does. Algorithm 3 takes the following considerations when selecting node pair (u, w) at v:

1. Both the 1-hop free neighbor u and the 2-hop free neighbor w can contribute additional coverage without any cost. Therefore, a pair of free neighbors should be included first.

2. Either the 1-hop free neighbor u or the 2-hop free neighbor w can decrease the total cost by half which leads to a higher yield.

3. Nodes u and w have equal cost and their contributions (in terms of coverage) are treated equally. Therefore, whichever covers a larger number of uncovered nodes will give a higher yield.

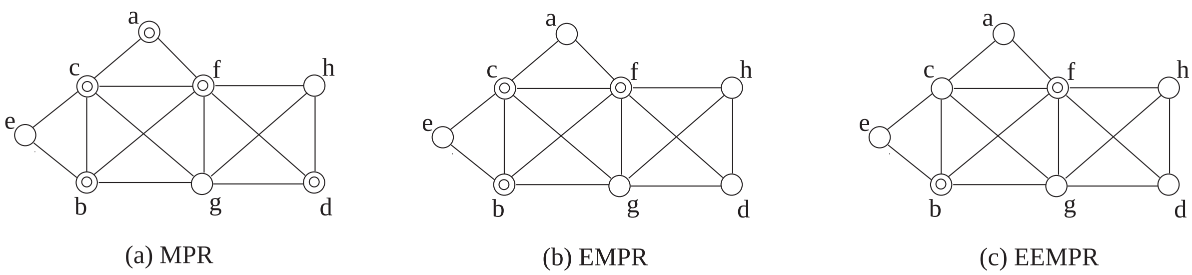
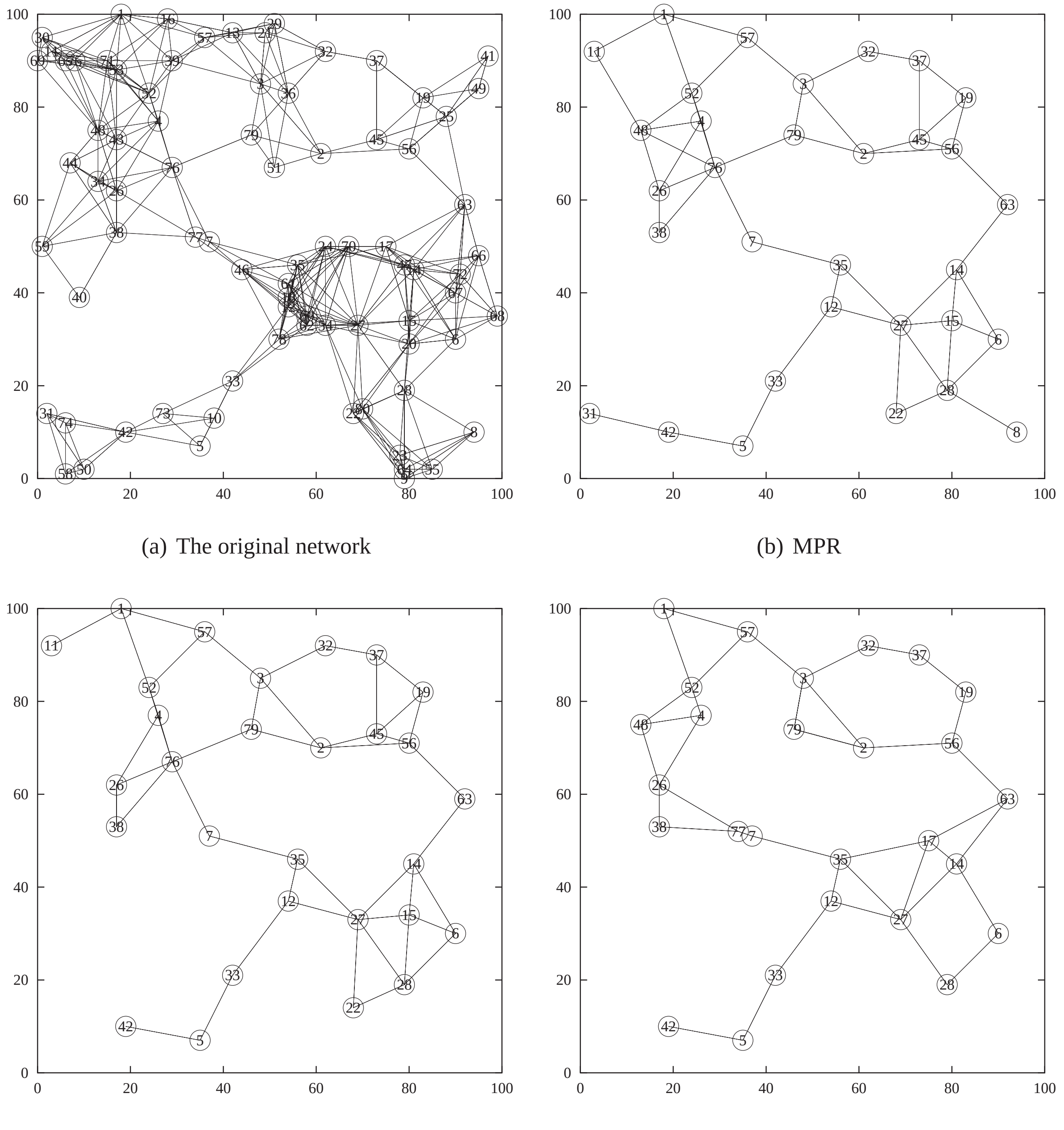


Figure 5: A sample network with 8 nodes. The double-circled nodes are selected forward nodes by (a) the MPR, (b) the EMPR, and (c) the EEMPR.

Figure 5 shows a sample network with 8 nodes. The double-circled nodes are the selected forward nodes by the source-independent MPR [1] (MPR) using Algorithm 1 with the regular Rules 1 and 2, the enhanced MPR [24] (EMPR) using Algorithm 2 with the Enhanced Rule 1 and regular Rule 2, and the proposed extended MPR (EEMPR) using Algorithm 3 with the Enhanced Rules 1 and 2. In Figure 5 (a), nodes a, b, and d are the nodes with the smallest ID within their corresponding 1-hop neighbors, so they are included in the CDS by Rule 1. Nodes c and f are selected as forward nodes by node a, which is the node with the smallest ID within c and f's 1-hop neighbors (Rule 2). Also, it is assumed that node b, the smallest ID neighbor of node g, selects $\{f\}$ to cover $H_2(b)$. Therefore, $\{a, b, c, d, f\}$ are in the CDS for the MPR. In Figure 5 (b), nodes a and d are removed from the CDS by the Enhanced Rule 1 because node a's 1-hop neighbors (c and f) are connected and d's 1-hop neighbors (f, g, and h) are pairwise connected. Therefore, $\{b, c, f\}$ are in the CDS for the EMPR. In Figure 5 (c), node c is removed from the CDS by the Enhanced Rule 2 because c's 1-hop neighbor with the smallest ID, a, selects f and b to indirectly cover e. Thus, only $\{b, f\}$ are in the CDS for the EEMPR.

Figure 6 (a) shows a relatively sparse network with 80 nodes. Figures 6 (b–d) show the results with MPR (Figure 6 (b)), EMPR (Figure 6 (c)), and EEMPR (Figure 6 (d)). In these figures, only nodes in the CDS and their induced subgraphs are shown. The size of the CDS's for MPR, EMPR, and EEMPR are 32, 29, and 27, respectively. Compared with previous work, EEMPR does not reduce significantly the CDS size in sparse networks. The major benefit of EEMPR is its probabilistic approximation ratio, which predicts a small average CDS size in dense networks, as will be discussed in the following sections.



(a) The original network

(b) MPR

(c) EMPR

(d) EEMPR

Figure 6: Applying different CDS algorithms in a sample network with 80 nodes.

4 Deterministic Local Approximation Ratio

This section shows local upper bounds of the original MPR and the proposed scheme. That is, the worst case performance of selecting a coverage set to cover the 2-hop neighbor set of a single node. The global performance will be discussed in the next section. Note that all results in this and the next section on the extended MPR (EEMPR) apply to both the general rules in Section 3.1 and its special cases (such as the greedy algorithm using free neighbors in Section 3.2). In both sections, G is assumed to be a unit disk graph.

4.1 Original MPR

In [18], Qayyum, Viennot, and Laouiti proved that the local upper bound of the ratio of the size of their proposed heuristic to that of the optimal multipoint relays is $O(\log n^{'})$, where $n^{'}$ is the maximum size of 2-hop neighbor set. Note that this ratio is with respect to multiple relays methods only (i.e., methods where 2-hop nodes are covered by selected 1-hop nodes). In fact, the approximation ratio is $O(n^{'})$ among all algorithms that cover 2-hop neighbor sets locally. Consider the example in Figure 7 (a) where all 1-hop neighbors of v are on the circle of C (with radius r from center v) and all 2-hop neighbors of v are on the circle of C' (with radius $2r$ from center v). r is the uniform transmission range of each node. Clearly, when u computes its forward nodes, each 2-hop neighbor of v, say w, on the circle of C' can only be covered by exactly one 1-hop neighbor of v, say u, on the circle of C whose position is exactly on the line connecting v and w (that is, there is a one-to-one relation between v and w). When the number of nodes on C' increases, the number of selected forward nodes on C also increases at the same rate. In fact, as indicated Figure 7 (a), a constant number of nodes (9 double-circled nodes) are sufficient to cover all 1-hop and 2-hop neighbors of v. Therefore, the approximation ratio of the original MPR is $O(n^{'})$.

4.2 Extended MPR

Next, we prove that for each single node v, the EEMPR can provide a constant size of the forward node set $C(v)$.

Theorem 2 (Local Upper Bound) *The EEMPR has a constant local approximation ratio.*

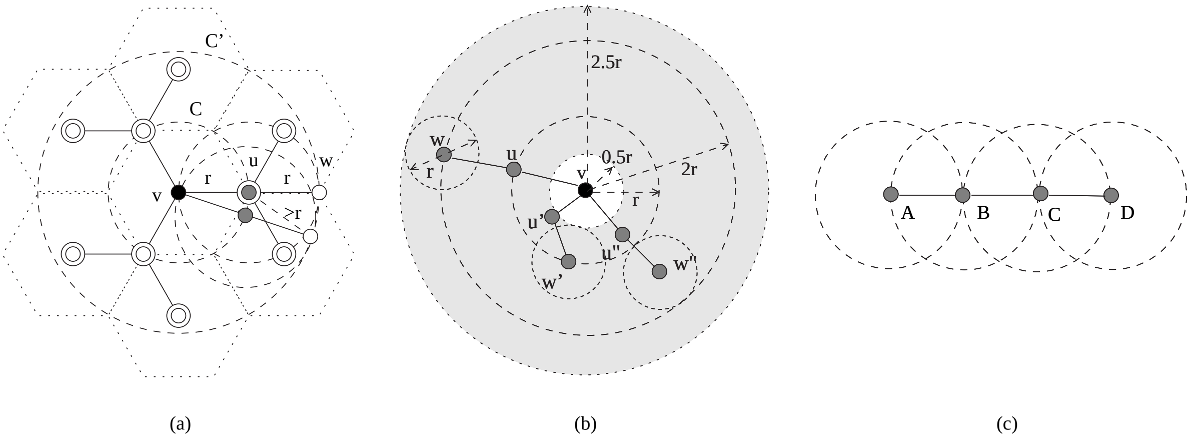


Figure 7: (a) An example of the worst case where node v's forward nodes are $O(n')$, (b) illustration of the EEMPR, and (c) the worst case where the CDS of the entire network is $O(n)$ for the EEMPR.

Proof: Suppose v is the node that selects a forward node set $C(v)$ to cover $H_2(v)$. Based on Definition 1, v selects a pair of nodes u and w, where u is in $H_1(v)$ and w is an uncovered node in $H_1(u) \cap H_2(v)$. The selected nodes are put into $C(v)$ and the nodes covered by $C(v)$ in $H_2(v)$ are removed. Node v continues to select pairs u' and w', u'' and w'', ..., and so on, until $H_2(v)$ becomes empty (see Figure 7 (b)). For each selection, the newly selected 2-hop forward node, say w', is not adjacent to any already selected 2-hop forward node (based on the property of unit disk graphs), say w, in $C(v)$. In other words, $\{w, w', w'', ...\}$ forms an independent set (IS). An IS is a set in which no two nodes are neighbors. This suggests that, within a disk whose diameter is r (or radius is $0.5r$), there exists at most one selected 2-hop forward node (of type w). In other word, such disks are non-overlapping. Notice that the possible location of v's 2-hop neighbor is only within the ring between r to $2r$. Thus, the disks with diameter r are confined within the ring between $0.5r$ to $2.5r$ (shadowed area in Figure 7 (b)). The maximum number of such disks is $\frac{\pi(2.5r)^2-\pi(0.5r)^2}{\pi(0.5r)^2} = 24$. Therefore, the total number of $\{w, w', w'', ...\}$ is no larger than 24, and the total number of nodes in $C(v)$, which is twice the size of $\{w, w', w'', ...\}$, is no larger than 48. Since the optimal number of forward nodes selected by each node to cover its 2-hop neighbor set is a constant, the proposed approach has a constant local approximation ratio. □

In [2], a disk with radius kr is proved to have an upper-bounded constant number of nodes l_k in an IS, where $l_k \leq (2k+1)^2$. The EEMPR provides a special case when $k = 2$. Although the EEMPR provides each node with a constant number of forward nodes, the upper bound of the CDS of the entire network is still $O(n)$ because the collection of the locally selected IS's does not correspond to a global IS. One worst case is shown in Figure 7 (c): all nodes are placed along line AD of length $3r$, and the node IDs monotonously increase along the line from the left end to the right end. Each node has its selector, which has the smallest ID among its 1-hop neighbor set. Based on the Enhanced Rule 2, a node will finally become a forward node if it is selected by its selector. Suppose n nodes are distinctively distributed along AD with $n/3$ nodes on each of the segments AB, BC, and CD. In addition, for each node u in BC (assuming one such u is at position B), there exist v in AB and w in CD, such that $Dis(u, v) = r$ and $Dis(u, w) = r$. All $O(n)$ nodes on the segment BC will be forward nodes. On the other hand, a CDS with only three nodes at positions A, B, and C is sufficient to cover the entire network. However, this situation corresponds to the worst case which rarely occurs. A study on average performance will be given in the next section.

5 Probabilistic Global Approximation Ratio

In this section, we first show that the original MPR and EMPR are probabilistically unbounded. That is, the average number of forward nodes in a finite region is infinite when the network is extremely dense. Then we prove that although the extended MPR (EEMPR) does not have a constant approximation ratio, it has a constant probabilistic approximation ratio. That shows the competitive average performance of the EEMPR.

5.1 Original MPR

Let CDS_{MPR} be the set of forward nodes selected by the original source-dependent MPR, source-independent MPR, or the EMPR based on partial 2-hop information. The following theorem shows that the MPR does not have a constant probabilistic upper bound in random unit disk graphs. We assume that all nodes are deployed in a 2-D plane. For each node v in a 2-D plane, $Disk(v, r)$ denotes a disk centered at v with a radius r. All nodes within $Disk(v, r)$ are 1-hop neighbors of v. All 2-hop neighbors of v are within $Disk(v, 2r)$, but not vice versa. As shown in Figure 8 (a), the deployment region (represented by the shadowed area) may not contain $Disk(v, 2r)$. Let $\omega \leq 2\pi$ be the arc of the disk within the deployment region. We assume the deployment region is rectangular and sufficiently large, such that $\omega \geq \pi/2$.

Theorem 3 (Probabilistically Unbounded) *In a network with n nodes randomly and uniformly*

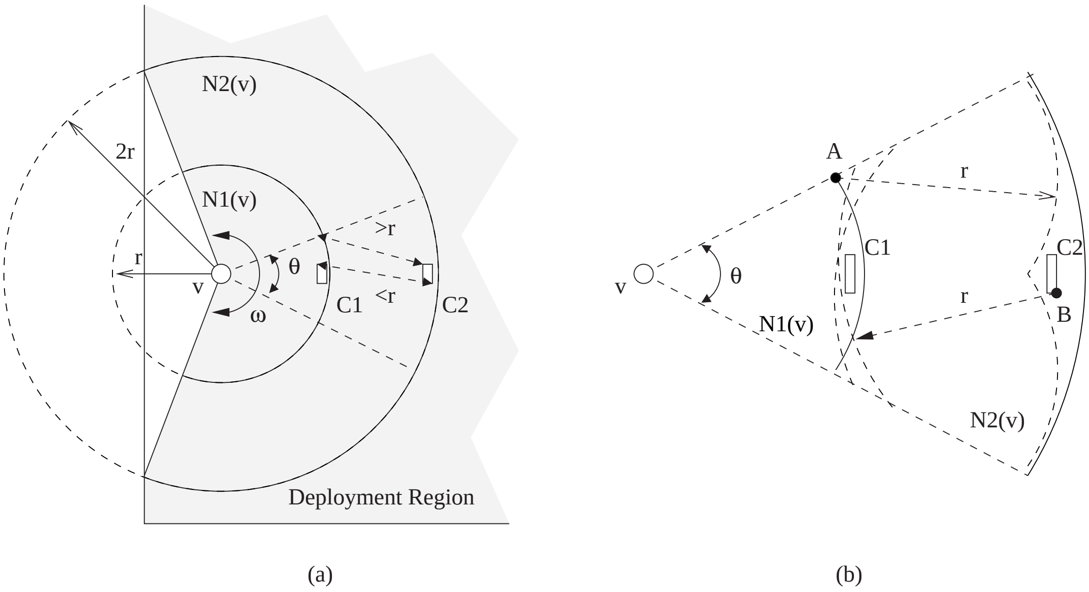


Figure 8: MPR is probabilistically unbounded.

deployed in a finite region that is sufficiently large, $\lim_{n\to\infty} E(|CDS_{MPR}|) = \infty$.

Proof: Let $C(v)$ be the set of forward nodes selected by a node v. We show that $\lim_{n\to\infty} E(|C(v)|) = \infty$. That is, for any integer $k_0 > 0$, there exist an n_0 satisfying $E(|C(v)|) > k_0$ for all $n > n_0$. As shown in Figure 8 (a), v's 2-hop neighborhood can be evenly divided into $4k_0$ sectors with angle $\theta = \omega/4k_0$. The following discussion shows that, for a sufficiently large n, the expected number of forward nodes is at least $1/4$ in each sector.

Within each sector, we construct two small regions C_1 and C_2 within $Disk(v, r)$ and $Disk(v, 2r)$, respectively. As shown in Figure 8 (b), C_2 is placed such that the distance between a node in C_2 and a node from a neighboring sector is always larger than r. C_1 is placed such that the distance between a node in C_1 and a node in C_2 is always less than r. The sizes of C_1 and C_2 can be very small but are both positive constants under a fixed k_0. If there exists a node u in C_1 and another node w in C_2, then w is a 1-hop neighbor of u and a 2-hop neighbor of v. In addition, w cannot be reached from v's 1-hop neighbors in any other sectors. Therefore, v must select at least one node from this sector of $Disk(v, r)$ to cover w.

Let S be the area of the deployment region and S' the minimum area of C_1 and C_2. The probability that v selects at least one forward node in a sector is

$$p = p_1 p_2 = (1 - (1 - \frac{S'}{S})^n)^2$$

where p_1 (p_2) is the probability that there exists at least one node in C_1 (C_2), and $p_1 = p_2 = 1 - (1 - \frac{S'}{S})^n$. The expected number of forward nodes selected by v is

$$E(|C(v)|) \geq 4k_0 p$$

Let $n_0 = \log_{1-\frac{S'}{S}}(\frac{1}{2})$. We have $p > \frac{1}{4}$ and $E(|C(v)|) > k_0$ for all $n > n_0$. If we let v be the source node in the original MPR, and a node with the minimal ID in the source-independent MPR and EMPR, then all nodes selected by v become forward nodes. That is, $C(v) \subseteq V_{MPR}$ and

$$\lim_{n\to\infty} E(|CDS_{MPR}|) \geq \lim_{n\to\infty} E(|C(v)|) = \infty$$

□

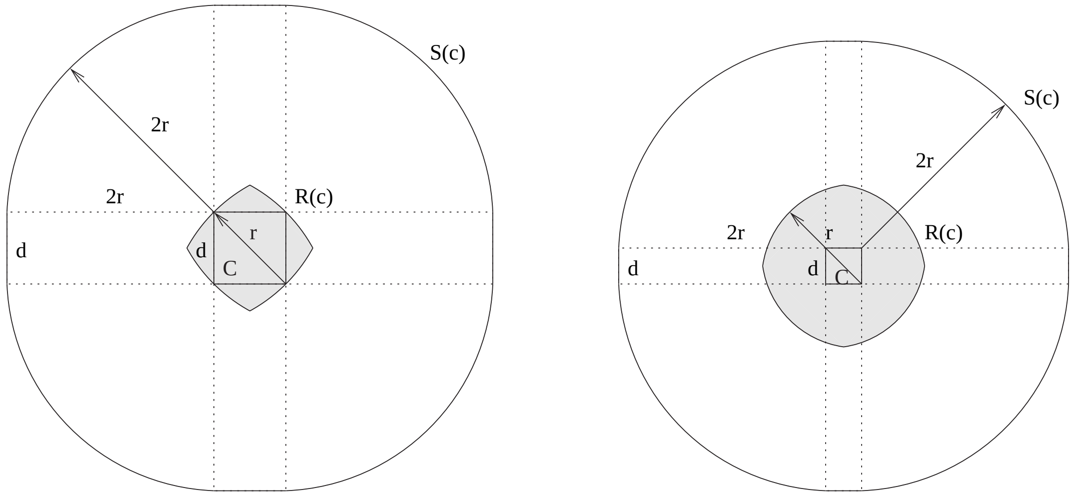
5.2 Extended MPR

Next, we give a probabilistic upper bound on the size of CDS derived from the extended MPR (EEMPR). Lemmas 1 and 2 show that the average number of forward nodes in each unit area is bounded by a constant. Theorem 4 further concludes that the average number of froward nodes in the entire network is constant times that of a minimum CDS.

Let $Disk(v,r) \cup Disk(u,r)$ denote the union (i.e., the combined area) of two disks, and $Disk(v,r) \cap Disk(u,r)$ the intersection (i.e., the common area) of two disks. Given a finite region C in the deployment region, a node v within region C is represented by $v \in C$. In the EEMPR, the number of forward nodes in C is determined by the decisions made by nodes in the following two regions.

Definition 4 (Selector and Releaser Regions) *Given a finite region C, its selector region is $S(C) = \bigcup_{v\in C} Disk(v, 2r)$ and its releaser region is $R(C) = \bigcap_{v\in C} Disk(v, r)$.*

Based on the Enhanced Rule 2, a node $v \in C$ is a forward node if (1) v is selected by a node $u \in S(C)$ as a forward node, and (2) v is not released by a node $w \in R(C)$, which does not select w as a froward node and has a lower ID than u. Figure 9 (a) shows a square region with side $d = \frac{\sqrt{2}}{2}r$ and its selector and releaser regions. Here we assume a very high node density such that there exists a node at every point of C. A lower node density will yield a smaller $S(C)$ and a larger $R(C)$. The area of $S(C)$ is $d^2 + 4d(2r) + \pi(2r)^2 = (0.5 + 4\sqrt{2} + 4\pi)r^2 \approx 18.72r^2$. The area of $R(C)$ is larger than that of C, which is $d^2 = 0.5r^2$. The area ratio between $S(C)$ and $R(C)$ is at



(a) (b)

Figure 9: Selector (white) and releaser (gray) regions of a square region C with (a) side $d = \frac{\sqrt{2}}{2}r$ and (b) $d < \frac{\sqrt{2}}{2}r$, where r is the transmission range.

most 37. This ratio is smaller when $d < \frac{\sqrt{2}}{2}r$, because $S(C)$ shrinks and $R(C)$ expands while C becomes smaller, as shown in Figure 9 (b).

Definition 5 (Node and Region Ranks) *Given a finite region C and a node $w \in C$, $rank(v) = |L_v|$, where $L_v = \{u \mid u \in S(C) \wedge id(u) \leq id(v)\}$ includes all nodes in $S(C)$ with an ID no higher than v. $rank(C) = \min_{v \in R(C)} rank(v)$ is the rank of a node with the minimal ID in $R(C)$.*

Lemma 1 *For any square region C with side $d \leq \frac{\sqrt{2}}{2}r$ and n nodes randomly and uniformly deployed in $S(C)$, $\lim_{n\to\infty} E(rank(C)) = O(1)$.*

Proof: First consider the case of $d = \frac{\sqrt{2}}{2}r$, where the probability that a node is placed in $R(C)$ is $p > \frac{1}{37}$. For any integer $i \geq 1$, event "$rank(C) = i$" is equivalent to "the first $i-1$ nodes with the lowest IDs are outside $R(C)$ and the i-th node is inside $R(C)$", and the corresponding probability is

$$Pr\{rank(C) = i\} = (1-p)^{i-1}p$$

When $n \to \infty$, the expected value of $rank(C)$ is

$$E(rank(C)) = \sum_{i=1}^{\infty} i \cdot Pr\{rank(C) = i\} = p\sum_{i=1}^{\infty} i(1-p)^{i-1} = \frac{1}{p} < 37$$

When $d < \frac{\sqrt{2}}{2}r$, p is larger and $E(rank(C))$ is smaller. Overall

$$\lim_{n \to \infty} E(rank(C)) < 37$$

□

Lemma 2 *The number of forward nodes selected by EEMPR in a finite region C is bounded by $O(1) \cdot rank(C)$.*

Proof: Let v be the node with the lowest ID in $R(C)$; that is, $rank(v) = rank(C)$. For any node $w \in C$, it becomes a forward node only if (1) it has the lowest ID in its neighborhood and, therefore, $rank(w) \leq rank(v)$ (Enhanced Rule 1), or (2) it is selected by a node $u \in S(C)$ and $rank(u) \leq rank(v)$ (Enhanced Rule 2). In case (1), the corresponding number of forward nodes is at most $rank(v)$. In case (2), there are at most $rank(v)$ such selectors in $S(C)$. From Theorem 2, each selector selects $O(1)$ forward nodes. Therefore, the total number of selected nodes in C is $O(1) \cdot rank(v)$. Overall, the total number of forward nodes in C is bounded by

$$rank(v) + O(1) \cdot rank(v) = O(1) \cdot rank(v) = O(1) \cdot rank(C)$$

□

Let CDS_{EEMPR} be the set of forward nodes selected in the EEMPR, and CDS_{OPT} be a minimum CDS selected by an optimal algorithm. The following theorem shows that the EEMPR has a constant probabilistic upper bound.

Theorem 4 (Probabilistic Upper Bound) *In a network with n nodes randomly and uniformly deployed in a finite region, $\lim_{n \to \infty} E(|CDS_{EEMPR}|) = O(1) \cdot E(|CDS_{OPT}|)$.*

Proof: Here we assume that the deployment region can be divided into many small but homogeneous squares with side no larger than $\frac{\sqrt{2}}{2}r$. From Lemmas 1 and 2, the expected number of forward nodes in each small square region is $O(1)$. Consider non-empty squares that contain at least one node. Suppose there are N such squares, then

$$\lim_{n \to \infty} E(|CDS_{EEMPR}|) = O(1) \cdot N \tag{1}$$

Now consider the minimum dominating set CDS_{OPT}. Each non-empty square must be (at least partially) covered by a node in CDS_{OPT}. On the other hand, each node in CDS_{OPT} can cover only $\Omega(1)$ non-empty squares. Therefore,

$$N = O(1) \cdot |CDS_{OPT}| \tag{2}$$

The theorem is proved by combining (1) and (2). □

In the above theorem, we assume global uniform node distribution for clarity. The probability density that a node is deployed at a given position (x, y) is assumed to be $f(x, y) = 1/S$ for all (x, y) within the deployment area, where S is the area of the deployment region. Nevertheless, Theorem 4 also holds in many networks with non-uniform node distributions. Note that the uniform distribution assumption is used only in Lemma 2 to establish a probabilistic upper bound of $rank(C)$. From the proof of Lemma 2, it is clear that $\lim_{n\to\infty} E(rank(C)) = O(1)$ as long as the probability p that a node in $S(C)$ is placed in $R(C)$ is a constant c. That is

$$\frac{\oint_{(x,y)\in R(C)} f(x,y)\mathrm{d}x\mathrm{d}y}{\oint_{(x,y)\in S(C)} f(x,y)\mathrm{d}x\mathrm{d}y} = c \tag{3}$$

Many non-uniform distributions exist that satisfy equation (3). For example, when the deployment region is an $\ell \times 1$ rectangle. The following node distribution

$$f(x,y) = \begin{cases} \dfrac{2^{\lfloor x \rfloor}}{2^{\ell}-1} & : \quad x \in [0,\ell] \wedge y \in [0,1] \\ 0 & : \quad x \notin [0,\ell] \vee y \notin [0,1] \end{cases}$$

satisfies equation (3) for any constants r and d in Lemma 2, while the node density changes dramatically from the left to the right of the deployment region.

6 Simulation

We compare the size of CDS for the following four algorithms: (1) MPR: the source-independent MPR algorithm proposed in [1], (2) EMPR: the enhanced MPR algorithm proposed in [24], (3) EEMPR: the extended MPR algorithm, and (4) REMPR: the random-selected extended MPR algorithm. The REMPR allows each node to randomly select a pair of 1-hop and 2-hop neighbors at each iteration of the forward node selection process. Both the EEMPR and REMPR apply Enhanced Rules 1 and 2 and have the same asymptotic performance as shown in Sections 4 and 5. The difference is that the EEMPR provides an efficient implementation of these rules using free neighbors.

All simulations were conducted using a custom simulator, which assumes an ideal network without node movement or channel collision. This simulation study focuses on efficiency (i.e., CDS size) instead of reliability (i.e., delivery ratio). However, such simplifications will not affect

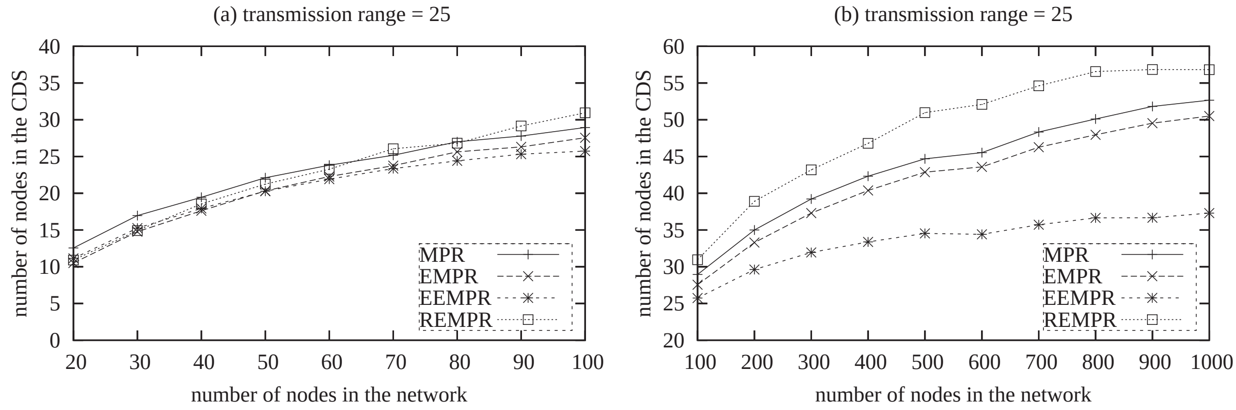


Figure 10: The number of nodes in the CDS when r is 25: a) n ranges from 20 to 100, and (b) n ranges from 100 to 1000.

our conclusion significantly. In the first scenario, a given number of nodes (ranging from 20 to 100 with a step of 10 and from 100 to 1,000 with a step of 100, respectively) are randomly distributed in a 100×100 2-D deployment region. Each node has a fixed uniform transmission range r (r is 25 and 50, respectively). If the generated network is not connected, it is discarded. The 90% confidence intervals of all simulation results are within $\pm$ 5%.

Figures 10 (a) and 10 (b) show the simulation results when the node's transmission range is 25. Figure 10 (a) shows the trend when the number of nodes in the network ranges from 20 to 100 (the corresponding graph is sparse), whereas Figure 10 (b) shows the trend when the number of nodes in the network is from 100 to 1000 (the corresponding graph is dense). We find that all three curves have a rising trend as the number of nodes in the network increases. The number of nodes in the CDS increases because, when more nodes join in the network, the network density increases and a node may select more 1-hop neighbors as forward nodes, which increases the size of the CDS. From the figure, we also notice that the rising trend is more sensitive to the number of nodes in the range from 20 to 100 (relatively sparse) than to that in the range from 100 to 1000 (relatively dense). The effect is more remarkable when the network is sparse because the greedy algorithm is a node coverage algorithm, that is, it selects 1-hop forward nodes to cover 2-hop neighbors. When the network is sparse, the collective coverage of the forward nodes may still leave some blank areas (i.e. areas with no nodes) within the 2-hop neighborhood. As more nodes join in, new nodes may appear in these blank areas thus resulting in the selection of more forward nodes. As the network density increases, the number of the blank areas reduces as does the number of newly selected

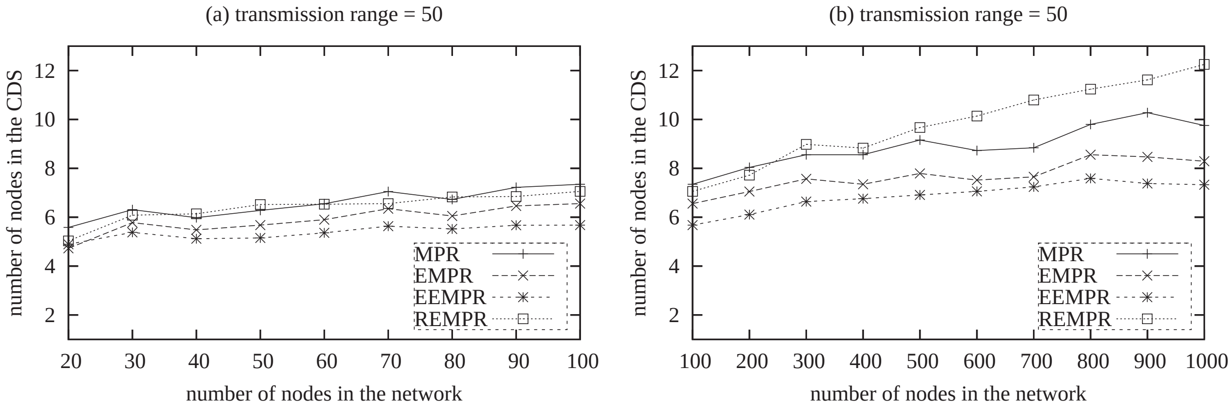


Figure 11: The number of nodes in the CDS when r is 50: a) n ranges from 20 to 100, and (b) n ranges from 100 to 1000.

forward nodes. Therefore, the rising trend slows down as the number of nodes increases.

Among these algorithms, the performance of the MPR and REMPR is the worst in all ranges. When the network is sparse (n is from 20 to 80), the curves of the EMPR and EEMPR are almost the same. But as the number of nodes increases, the gap between the EMPR and EEMPR becomes significant. When the number of nodes in the network is 1000, the number of nodes in the CDS determined by the EEMPR is only around 70% of that determined by the EMPR or MPR. The reason that the EEMPR shows great improvement in dense networks is that the selection of the forward nodes for one node has an upper bound that is irrelevant to the network density. Thus, the size of the CDS is less influenced by the network density. The size of CDS of the REMPR is larger than that of the MPR when the network size is over 60. The relatively low performance of the REMPR shows that, although both the REMPR and EEMPR have the same asymptotic upper bound, optimization techniques such as free neighbor can still make significant impact in their average performance.

Figures 11 (a) and 11 (b) show the results when the node's transmission range is 50 and number of nodes in the network is from 20 to 100 and from 100 to 1000, respectively. When the transmission range increases, the graph becomes denser if the number of nodes is fixed. In this case, the size of the CDS only increases slightly as the size of the network increases. This is because, when the transmission range is 50, the corresponding graph is sufficiently dense for the number of nodes to have little effect on network density. Among these algorithms, the EEMPR performs the best, followed by the EMPR. The MPR is the worst when the size of network is less than 200, and the

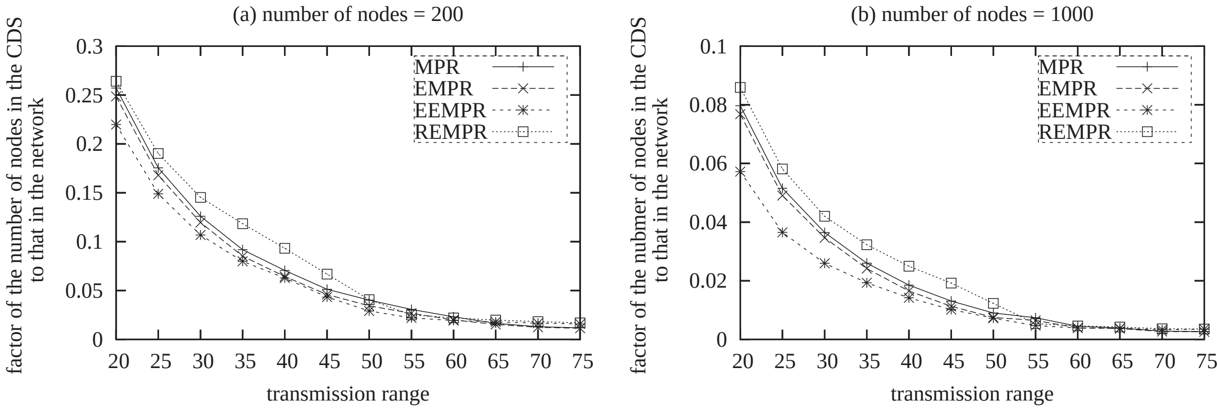


Figure 12: The factor of the number of nodes in the CDS to that in the network when r is from 20 to 75: (a) n is 200 , and (b) n is 1000.

REMPR has the largest CDS size when the network population is over 200.

Comparing Figures 10 (a) and 10 (b) with Figures 11 (a) and 11 (b), we find that increasing the node's transmission range can increase the coverage area of each node and, therefore, reduce the diameter of the network, which leads to a smaller size of the CDS.

In the second scenario, a fixed number of nodes (n = 200 and 1000, respectively) are randomly distributed in the same 2-D space. The network density is determined by the node's transmission range r. For each fixed number of nodes, we run different experiments where the value of r changes from 20 to 75. The results of sufficient number of experiments for each fixed network density are averaged to guarantee the same confidence interval.

Figures 12 (a) and 12 (b) show the percentage f of forward nodes versus the node's transmission range when the number of nodes is 200 and 1000, respectively. When the transmission range r increases, the factor decreases because the increase in r results in the decrease of the diameter of the network. Thus, less nodes are needed to cover the deployment region.

From the above results, we conclude that the proposed EEMPR always outperforms the MPR and EMPR regardless of the size of the network and the density of the network. The performance difference is more significant in relatively sparse networks. REMPR has poor performance, which confirms the effectiveness of free neighbor enhancement. All localized approaches demonstrate good scalability in dense networks, although the EEMPR is the only one that has a constant approximation ratio in expectation.

7 Related Work

The problem of finding a minimum CDS (MCDS) for a general network is proved to be NP-complete [5]. Even for a unit disk graph, such a problem is NP-complete [13]. Therefore, only heuristic algorithms can be applied. CDS construction algorithms can be classified into four groups: global [8, 9], quasi-global [23], quasi-local [27], and local [1, 4, 7, 11, 12, 14, 16, 18, 19, 20, 21, 22, 24, 25, 26].

Some earlier researchers proposed global greedy algorithms [9] that provide an approximation ratio of $O(\ln \Delta)$ for general networks, where Δ is the maximum node degree of the network. Quasi-global CDS algorithms [23] build shortest-path-tree-based CDS structures to provide $O(1)$ approximation ratio for unit disk graphs. In contrast, quasi-local CDS algorithms construct a CDS by first electing clusterheads [27] and then using selected gateway nodes to connect them. Quasi-local methods also have $O(1)$ approximation ratios in unit disk graphs.

Among local algorithms, a generic localized broadcast scheme [25] has been proposed to unify source-independent and source-dependent approaches. Source-independent approach forms a "static" CDS of the network that only depends on the network topology and node priority. Many algorithms belong to this group, such as source-independent MPR [1] and its extension EMPR [24], marking process [26] and its extensions [7], SPAN [4], CEDAR [20], and d-hop CDS [19]. In contrast, the source-dependent approach depends on the source of a specific broadcast operation. When a specific broadcast starts, after receiving a broadcast packet, the node determines both its own and/or some of its neighbors' forward/non-forward statuses under a local view of its neighbor set. The local view of its neighborhood can be updated by "Hello" messages and the broadcast packet. As the broadcast packet traverses the network, the forward nodes eventually form a "dynamic" CDS of the given network. Algorithms that belong to this group are source-dependent MPR [18], dominant pruning [11] and its extensions [12, 14], LENWB [22], SBA [16], and neighbor-elimination-based broadcasting [21].

A local broadcast algorithm is either *self-pruning*, *neighbor-designating*, or *hybrid* [25]. In self-pruning approaches [4, 7, 16, 19, 21, 22, 26], each node determines its forward/non-forward status based on local information, which is usually its 2- or 3-hop information. A node becomes a forward node when it has two 1-hop neighbors that cannot be connected via other nodes with higher priorities. In neighbor-designating approaches [1, 11, 12, 14, 18, 24], forward nodes are nominated by their 1-hop neighbors. Depending on the broadcast scheme, a nominated node may or may not becomes a forward node. Hybrid approaches [25] combine both self-pruning and neighbor-designating methods.

The three algorithms (MPR [1], EMPR [24], and EEMPR) discussed in this paper are neighbor-designating approaches. MPR has been used to generate both source-dependant and source-independent (and relatively large) CDS. An extension was proposed in [1] to reduce the size of a source-independent CDS, which was further extended in [24] using free neighbors. Other neighbor-designating approaches, such as the dominant pruning algorithm [11] and its extensions [12, 14], have been proposed to produce a smaller source-dependent CDS. All these algorithms use a similar greedy strategy in selecting 1-hop forward nodes, which has a logarithmic local approximation ratio. When node position information is available, a better strategy can be used to achieve a constant local approximation ratio [3]. The drawback of this method is the overhead for obtaining node positions and inaccuracy due to channel fading and shadowing. Note this approximation ratio applies to the case of selecting 1-hop neighbors to cover 2-hop neighbors, not to the case of selecting both 1-hop and 2-hop neighbors. As discussed early in this paper, an optimal solution to the former can be asymptotically larger than a solution to the latter.

8 Conclusions

We have proposed an extended source-independent MPR based on the recently proposed source-independent MPR. The enhancement is done by using complete 2-hop neighborhood information to cover each node's 2-hop neighbor set and by extending the notion of coverage in the original MPR. The effectiveness of the enhancement is confirmed through both probabilistic analysis and simulation study. In addition, we prove a constant probabilistic upper bound for the extended MPR and the non-existence of such a bound for the original MPR and its existing extensions.

In this paper, we did not consider energy-aware multiple relays selection. One straightforward extension is to use residue energy level as the selection criterion instead of using node ID. That is, the smallest ID node is replaced by the node with the highest residue energy level. In this case, a node with the highest residue energy in its 1-hop neighborhood has a better chance to become a forward node based on the enhanced rules. In this way, we can conduct an energy-aware broadcasting [21]. Another future work will be the study of different non-uniform node distributions that can still provide a constant probabilistic bound for the extended MPR.

References

[1] C. Adjih, P. Jacquet, and L. Viennot. Computing connected dominated sets with multipoint relays. *http://www.inria.fr/rrrt/rr-4597.html*, 2002.

[2] K. Alzoubi, X. Y. Li, Y. Wang, P. J. Wan, and O. Frieder. Geometric spanners for wireless ad hoc networks. *IEEE Trans. of Parallel and Distributed Systems*, 14(5):408–421, May 2003.

[3] G. Calinescu, I. Mandoiu, P. J. Wan, and A. Zelikovsky. Selecting forwarding neighbors in wireless ad hoc networks. *Proceedings of the 5th international workshop on Discrete algorithms and methods for mobile computing and communications*, pages 34–43, Oct. 2001.

[4] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris. Span: an energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks. *ACM Wireless Networks Journal*, 8(5):481–494, Sep. 2002.

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

[6] T. Clausen and P. Jacquet. Optimized link state routing protocol. IETF drafts (draft-ietf-manet-olsr-11.txt), July 2003.

[7] F. Dai and J. Wu. An extended localized algorithm for connected dominating set formation in ad hoc wireless networks. *IEEE Transaction on Parallel and Distributed Systems*, 15(10):908–920, Oct. 2004.

[8] B. Das, R. Sivakumar, and V. Bharghavan. Routing in ad-hoc networks using a spine. *Proc. of the 6th Int'l Conf. on Computer communications and Networks (ICCCN'97)*, pages 1–20, Sep. 1997.

[9] S. Guha and S. Khuller. Approximation algorithms for connected dominating sets. *Algorithmica*, 20(4):374–387, 1998.

[10] D. B. Johnson and D. A. Maltz. *Dynamic source routing in ad-hoc wireless networks*, in Mobile Computing, edited by Imielinski and Korth, pages 153–181, Kluwer Academic Publishers, Boston, MA, 1996.

[11] H. Lim and C. Kim. Flooding in wireless ad hoc networks. *Computer Communications Journal*, 24(3-4):353–363, 2001.

[12] W. Lou and J. Wu. On reducing broadcast redundancy in ad hoc wireless networks. *IEEE Trans. on Mobile Computing*, 1(2):111–123, Apr.-Jun. 2002.

[13] M. V. Marathe, H. Breu, H. B. Hunt III, S. S. Ravi, and D. J. Rosenkrantz. Simple heuristics for unit disk graphs. *Networks*, 25:59–68, 1995.

[14] M. Mosko and J. J. Garcia-Luna-Aceves. A self-correcting neighbor protocol for mobile ad-hoc wireless networks. *Proc. of IEEE IC3N'02*, pages 556 – 560, Oct. 2002.

[15] S. Ni, Y. Tseng, Y. Chen, and J. Sheu. The broadcast storm problem in a mobile ad hoc network. *Proc. of ACM/IEEE MOBICOM'99*, pages 151–162, Aug. 1999.

[16] W. Peng and X. C. Lu. On the reduction of broadcast redundancy in mobile ad hoc networks. *Proc. of First Annual Workshop on Mobile and Ad Hoc Networking and Computing (MOBIHOC'2000)*, pages 129–130, Aug. 2000. Boston, USA.

[17] C. Perkins and E. M. Royer. Ad-hoc on-demand distance vector routing. *Proc. of 2nd IEEE Workshop on Mobile Computing Systems and Applications (WMCSA)*, pages 90–100, Feb. 1999. New Orleans, LA.

[18] A. Qayyum, L. Viennot, and A. Laouiti. Multipoint relaying for flooding broadcast message in mobile wireless networks. *Proc. of 35th Hawaii Int'l Conf. on System Sciences (HICSS-35)*, pages 3898–3907, Jan. 2002.

[19] M. Q. Rieck, S. Pai, and S. Dhar. Distributed routing algorithms for wireless ad hoc networks using d-hop connected dominating sets. *Proc. of HPC-ASIA*, pages 443–450, Dec. 2002.

[20] P. Sinha, R. Sivakumar, and V. Bharghavan. Enhancing ad hoc routing with dynamic virtual infrastructures. *Proc. of IEEE INFOCOM'2001*, 3:1763–1772, Apr. 2001.

[21] I. Stojmenovic, S. Seddigh, and J. Zunic. Dominating sets and neighbor elimination based broadcasting algorithms in wireless networks. *IEEE Trans. on Parallel and Distributed Systems*, 13(1):14–25, Jan. 2002.

[22] J. Sucec and I. Marsic. An efficient distributed network-wide broadcast algorithm for mobile ad hoc networks. submitted for publication,2002.

[23] P. J. Wan, K. Alzoubi, and O. Frieder. Distributed construction of connected dominating set in wireless ad hoc networks. *Proc. of IEEE INFOCOM'2002*, 3:1597–1604, Jun. 2002.

[24] J. Wu. An enhanced approach to determine a small forward node set based on multipoint relay. *Proc. of IEEE VTC 2003 fall*, Sep. 2003.

[25] J. Wu and F. Dai. A generic distributed broadcast scheme in ad hoc wireless networks. *Proc. of ICDCS'2003*, pages 460–468, May 2003.

[26] J. Wu and H. Li. On calculating connected dominating sets for efficient routing in ad hoc wireless networks. *Proc. of ACM DIALM'99*, pages 7–14, Aug. 1999.

[27] J. Wu and W. Lou. Forward-node-set-based broadcast in clustered mobile ad hoc networks. *Wireless Networks and Mobile Computing, a Special Issue on Algorithmic, Geometric, Graph, Combinatorial, and Vector Aspects*, 3(2):155–173, 2003.