

# Theoretical Bound and Practical Analysis of Connected Dominating Set in Ad Hoc and Sensor Networks

Alireza Vahdatpour, Foad Dabiri, Maryam Moazeni, and Majid Sarrafzadeh

Computer Science Department  
University of California Los Angeles  
{alireza,dabiri,mmoazeni,majid}@cs.ucla.edu

**Abstract.** *Connected dominating set* is widely used in wireless ad-hoc and sensor networks as a routing and topology control backbone to improve the performance and increase the lifetime of the network. Most of the distributed algorithms for approximating connected dominating set are based on constructing *maximal independent set*. The performance of such algorithms highly depends on the relation between the size of the maximum independent set ( $mis(G)$ ) and the size of the minimum connected dominating set ( $cds(G)$ ) in the graph  $G$ . In this paper, after observing the properties of such subgraphs, we decrease the previous ratio of 3.453 to 3.0 by showing that  $mis(G) \leq 3 \cdot mcds(G) + 3$ . Additionally, we prove that this bound is tight and cannot be improved. Finally, we present practical analysis of constructing connected dominating set based on maximal independent set in wireless networks. It is shown that the theoretical bound for *unit disk graph* is still practically applicable for general wireless networks.

## 1 Introduction

Wireless ad-hoc and sensor networks are gaining more interest in a variety of applications. In addition to applications such as habitat and environment monitoring, these networks are useful in emergency operations such as search and rescue, crowd control and commando operations [15]. In such wireless networks, each node communicates with other nodes either directly or through other intermediate nodes. These networks can not have fixed or centralized infrastructure, since network nodes are not assumed to be fixed in the environment. Although a wireless ad-hoc network has no physical infrastructure, connected dominating set can be used as a virtual backbone in the network [19]. A connected dominating set (CDS) of a graph  $G = (V, E)$  is a subset  $S \subseteq V$  such that each node in  $V - S$  is adjacent to some nodes in  $S$  and the communication graph induced by  $S$  is connected. Centralized and distributed construction of dominating set in sensor networks has been studied widely before, since it has great impact on the efficiency of routing, power management, and topology control protocols.

The problem of finding a minimum CDS (MCDS) in a graph has been shown to be NP-hard [2]. Many efforts [1], [6], [7], [18], [19], [20] have been made to design

approximations and heuristics for the MCDS in a network. Among distributed approximation algorithms, which are the most suitable algorithms for ad-hoc and sensor networks, using the idea of constructing a maximal independent set and then converting it to a connected graph is a common trend. This approach has several advantages comparing to centralized and other distributed MCDS approximation algorithm. Sensor and ad-hoc networks have severe limitation in processing and communication capabilities, therefore, localized distributed algorithms are the most suitable solutions for issues arising from inefficient energy consumption in the network. Since every maximal independent set is a dominating set and can be constructed locally, most of the well-known algorithms usually construct a maximal independent set at the first step and then convert it to a CDS with a local algorithm. Consequently, the approximation performance ratio would be determined by two factors: 1) How large a maximal independent set can be compared to a minimum connected dominating set. 2) How many vertices are required to connect a maximal independent set. Simple algorithms are available that construct a connected graph from maximal independent set by the ratio of 2 [13], [19].

Several studies have been done to find the ratio of number of nodes in MIS to the size of MCDS in a graph. All the studies assume the graph to be unit disk graph. A *unit disk graph (UDG)* is the intersection graph of a family of unit circles in the Euclidean plane. In such graph, a vertex  $v$  is connected to vertex  $u$ , if and only if  $|uv| \leq 1$ , where  $|uv|$  represents the Euclidean distance between  $u$  and  $v$ . Even assuming the graph to be UDG, finding MCDS has been proven to be NP-Hard [2]. Using the geometric properties of unit disk graph, [13], [19], [10], [21], [14] have shown different constant ratios for the relation between number of nodes in MCDS and MIS subgraphs. The study in [21] showed that in every unit disk graph:  $mis(G) \leq 3.8cds(G) + 1.2$ . where  $mis(G)$  is the size of a maximum independent set and  $cds(G)$  is the size of a minimum connected dominating set in  $G$ . More recently, Funke et. al. [10] decreased the bound to  $3.453 \cdot cds(G) + 8.291$ . In this paper, we decrease the bound and show that  $mis(G) \leq 3 \cdot cds(G) + 3$ . Additionally, we analyze the ratio in practical networks, where the network does not necessarily obey the characteristics of a unit disk graph. Through running large set of simulation experiments, we observed that although theoretically the bound is not applicable to a general network where nodes have indeterministic wireless coverage. Furthermore, we analyze the effect of network density and connectivity on the construction of connected dominating set and its practical relation to maximal independent set in wireless networks.

The rest of this paper is organized as follows. Section 2, discusses the applications and usage of MCDS in sensor and ad-hoc networks. In section 3, related work and observations that lead to the main results of this paper are presented. Section 4 presents the main results which includes the proof of the new bound for the relation between the number of nodes in MCDS and MIS. Section 5 includes experimental results for practical analysis and observations. Finally, the last section contains some concluding remarks.

## 2 Applications of Connected Dominating Set in Sensor and Ad Hoc Networks

Several approaches have been proposed in the literature to maximize the network lifetime, as solutions to great challenges that energy constraints impose on sensor and ad-hoc networks. Using connected dominating set as a network backbone for routing and topology control is one of the most mature approaches that can lead to efficient energy consumption in the network. In this approach, messages can be routed from the source to one of its neighbors in the dominating set, along the CDS to the dominating set member closest to the destination node, and finally to the destination [12]. *Dominating set based routing* [7], [20], *spine based routing* [5] or *backbone based routing* [3] are the terms used in the literatures particularly addressing this routing method.

Furthermore, CDS has been widely used to increase the efficiency of multicast and broadcast routing. According to [12], [16] *broadcast storm problem* occurs when a large number of intermediate nodes unnecessarily forwards a message. Therefore, causing several nodes to hear the same message several times, which causes unwanted energy consumption in the network. In another approach CDS is applied to increase the number of nodes that can be in the sleep mode, while the connectivity of the network is maintained and messages can effectively be forwarded throughout the network [17], [20]. In [8] the nodes in dominating set coordinate with one another and use orthogonal spreading codes in their neighborhoods to improve spatial reuse with code division spread spectrum techniques.

Recently, Banik and Radhakrishnan [4] discussed that while using MCDS as backbone for broadcasting minimizes the the total number of transmissions, it does not guarantee minimum broadcast latency. Hence, they proposed a collision free broadcast scheduling algorithm among nodes in MCDS to minimize the broadcast latency.

Different approaches are applied to construct connected dominating set over a network. Centralized algorithms are not suitable for sensor and ad-hoc networks since they require one node to perform most of the computation and communication, which leads to unbalanced energy consumption in the network. In contrast, a common distributed approach that has been addressed by [13] for the first time, is to use maximal independent set to form the connected dominating set. This approach is popular and efficient among distributed CDS algorithms since distributed construction of maximal independent set in large networks needs few number of local communications among the nodes. Currently, the best known approximation algorithm for connected dominating set uses the above approach [14]. [21], [19], [13] investigated the relation of minimum connected dominating set and maximal independent set. The common assumption in most of the work proposing algorithms for constructing MCDS over the network is that a wireless network can be modeled as a Unit Disk Graph (UDG). Exploiting the geometric properties of unit disk graphs, introducing effective approximation algorithms for MCDS is possible. [13] has addressed few simple heuristics for unit disk graphs that leads to acceptable approximation bounds.

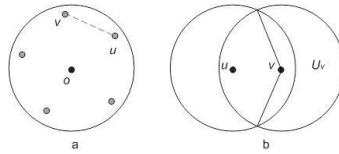
### 3 Motivation, Observations and Related Work

Table 1 summarizes performance comparison of well known distributed CDS construction algorithms [12]. The best performance of distributed algorithms is achieved by [14] where it first finds a maximal independent set and then converts it to a connected dominating set. Finding maximal independent set and converting it to a dominating set is relatively an easy task in a graph, therefore, most of distributed algorithms try to efficiently convert the dominating set constructed by maximal independent set to a connected tree which has the minimum cardinality from the optimal MCDS. In most of these conversion algorithms the size of the connected tree is at most twice of the input dominating set. Hence, the final approximation ratio of the connected dominating set is highly depended on the approximation ratio of the size of maximal independent set to the size of the MCDS.

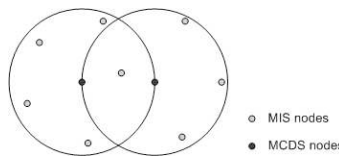
In [13] authors assumed Unit Disk Graphs (UDG) and showed that in this type of geometric graphs, the ratio of number of nodes in MIS to the number of nodes in MCDS is at most 5. In a UDG there is an edge between two vertices if the distance between them is less than one unit. Since wireless sensor/ad-hoc networks are usually considered to be consisted of homogeneous nodes with equal wireless ranges, using UDG for the theoretical model is reasonable. Environment effects and slight geometric conditions are ignored in UDG, however. Modeling a wireless network as a UDG, each vertex of MCDS cannot dominate more than five independent nodes, since no more than five independent nodes can lie in a unit disk area (Figure 1.a). This immediately results in that the maximum ratio of nodes in every MIS is not more than five times of the size of the MCDS. [19] decreased the ratio to 4, by showing that the intersection between the dominating areas (communication coverage) of each two adjacent nodes does not let every vertex in MCDS to dominate five independent nodes. Figure 1.b shows part of a minimum connected dominating set.  $v$  and  $u$  are two adjacent vertices in MCDS. It can be seen that  $U_v$  lie in a sector of at most  $240^\circ$  within the coverage range of vertex  $v$ . It can be easily shown that at most 4 independent nodes can lie in this area.

**Table 1.** Performance comparison of previous CDS construction distributed algorithms [12] (Here,  $n$  and  $m$  are the number of vertices and edges in the graph.  $\Delta$  and  $H$  denote the maximum degree of the vertices in the graph and the harmonic function, respectively.)

	Graph Model	Approximation Ratio	Time Complexity	Message Complexity
[7]-I	general	$2H(\Delta) + 1$	$O((n +  C )\Delta)$	$O(n C  + m + n\log(n))$
[7]-II	general	$2H(\Delta)$	$O( C (\Delta +  C ))$	$O(n C )$
[19]	UDG	$8opt + 1$	$O(n)$	$O(n\log(n))$
[1]	UDG	$192opt + 48$	$O(n)$	$O(n)$
[6]	UDG	$8opt$	$O(n)$	$O(n\log(n))$
[20]	general	$O(n)$	$O(\Delta^3)$	$\theta(m)$
[14]	UDG	$6.8opt$	-	-



**Fig. 1.** a) A neighboring area with five independent nodes b)  $U_v$  lies in a sector of at most  $240^\circ$



**Fig. 2.** An example where 8 nodes lie in the neighboring area of two vertices

[21] has extended the observation and proved that the number of nodes in the MIS is at most  $3.8cds(G) + 1.2$ . Using this upper-bound, the approximation ratio of the algorithms [1], [6], [19] that use MIS to find an approximate connected dominating set on a network would be 7.6. The key observation in [21] is that the neighboring area around two adjacent nodes in MCDS cannot have more than eight independent nodes. Figure 2 depicts an example where eight nodes lie in the neighboring area of two adjacent nodes in MCDS.

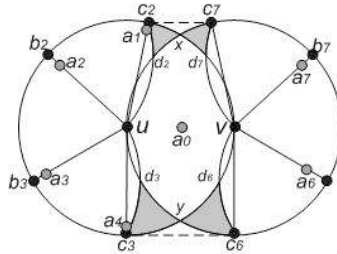
Recently, [10] et al. presented a better bound using Voronoi diagram of the vertices in the MCDS. Up to now,  $3.453 \cdot cds(G) + 8.291$  achieved by this paper is the best bound. In the next section, we will extend UDG properties discussed before and show that the ratio of number of nodes in MIS to the number of nodes in MCDS is  $3 \cdot cds(G) + 3$ . The idea behind the proof in the next section is to apply the previous observation to a number of nodes connected together. It will be shown that the number of independent nodes in one vertex's neighboring area depends on the position and the number of independent nodes in the adjacent MCDS vertices' neighboring areas and by generalizing the relation, the minimum upper bound can be driven.

### 4 Main Results

In this section, we will present our contribution which is a tight bound on the approximation ratio between *MIS* and *MCDS*.

**Definition 1.** For any node  $u \in G(V, E)$ , the neighboring set of  $u$ ,  $N(u)$  is defined to be the set of nodes adjacent to  $u$  in  $G$ , in other words:

$$N(u) = \{v | e_{uv} \in E\}$$



**Fig. 3.** Four small dark areas

**Lemma 1.** For any  $u \in G$ :

$$|N(u) \cap mis(G)| \leq 5 \tag{1}$$

where  $mis(G)$  is a maximal independent set of graph  $G$ .

*Proof.* If  $u$  and  $v$  are independent ( $|uv| > 1$ ), it is depicted in figure 1.a, that the angle made by connecting vertices  $v$  and  $u$  to vertex  $o$  is greater than  $60^\circ$ . Therefore, at most five mutually independent nodes can exist in a unit disk.  $\square$

**Lemma 2.** Each two adjacent nodes in MCDS cannot have more than 8 independent nodes in their neighboring area:

$$\forall u, v \in mcDs(G), |uv| \leq 1; |(N(u) \cup N(v)) \cap mis(G)| \leq 8 \tag{2}$$

*Proof.* The complete proof can be found in [21]. However, since the next lemma is a generalization of this lemma, we will briefly overview the proof scheme here. Figure 3 depicts two adjacent vertices  $u$  and  $v$ , and their neighboring area. Without loss of generality, we assume that  $|uv| = 1$ , therefore,  $N(u) \cup N(v)$  is maximum. First, if two independent nodes lie in the  $N(u) \cap N(v)$ , according to lemma 1,  $N(u) - N(v)$  and  $N(v) - N(u)$  can include at most three independent nodes, resulting in the total of at most 8 independent nodes in the  $N(u) \cup N(v)$ .

In the second case, assume that only node  $a_0 \in mis(G)$  lies in  $N(u) \cap N(v)$ . In order to have more than 8 independent nodes in  $N(u) \cup N(v)$ , four nodes must lie in both  $N(v) - N(u)$  and  $N(u) - N(v)$ . Lets assume that four independent nodes  $(a_1, a_2, a_3, a_4)$  lie in  $N(u) - N(v)$ . Considering that  $a_6$  and  $a_7$  are lying in  $N(v) - N(u)$  and the fact that the sector between each two independent nodes in a circle is at least  $60^\circ$  (see lemma 1),  $a_5$  and  $a_8$  must lie in the dark areas  $yd_6c_6$  and  $xd_7c_7$ .

According to the proof available in [21], the maximum euclidian distance between any point of  $c_3d_3y$  and any point of  $d_6c_6y$  is less than unit. Similarly, points in  $c_2d_2x$  and  $c_7d_7x$  has the same property accordingly. Therefore the dark areas in  $N(v)$  cannot include independent nodes and hence four nodes cannot lie in  $N(v) - N(u)$ . This completes the proof.  $\square$

**Definition 2. Semi-exclusive neighboring set:** let  $U$  be any set of independent nodes in  $G = (V, E)$ , and  $T$  be any spanning tree on  $mcDs(G)$ . Assume

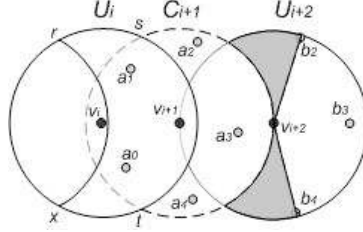


Fig. 4. Three adjacent vertices in MCDS and their neighboring areas

$v_1, v_2, \dots, v_T$  is an arbitrary preorder traversal of  $T$ . For any  $i, 2 \leq i \leq |T|$ , consider  $U_i = N_i \cap U - \bigcup_{j=1}^{i-1} N(j)$  be the subset of nodes in  $U$  that are adjacent to  $v_i$  but not to any of  $v_1, v_2, \dots, v_{i-1}$ . We will call  $U_i$  the semi-exclusive neighboring set of node  $v_i$ .

**Lemma 3.** For all  $i, j \in mcds(G), |U_i| = |U_j| = 4$  there exists a node  $k \in mcds(G)$  on the path between  $i$  and  $j$ , where  $|U_k| < 3$ .

*Proof.* Lets  $v_0, v_1, v_2, \dots, v_m$  be an arbitrary path in the  $mcds(G)$  and  $N_1, N_2, \dots, N_m$  and  $U_1, U_2, \dots, U_m$  be the neighboring and semi-exclusive sets associated with vertices  $v_1, v_2, \dots, v_m$ . Suppose that  $U_i$  and  $U_j$  contain four independent nodes. We will show that there exists  $v_k, i < k < j$  such that  $U_k$  contains less than three independent nodes.

Figure 4 shows vertex  $v_i$  and its three immediate neighbor vertices in  $mcds(G)$ ,  $v_{i-1}, v_{i+1}$  and  $v_{i+2}$ . As seen in Figure 4, since  $\widehat{st} \geq 120^\circ$  and  $\widehat{rx} \leq 240^\circ$ ,  $N_{i+1} \cap U_i$  will include either three or two independent nodes. In the first case, according to lemma 1  $|U_{i+1}| \leq 2$  and this satisfies the lemma. In the second case, assume  $a_0, a_1 \in N_{i+1} \cap U_i$ . According to lemma 2,  $U_{i+1}$  can contain at most three independent nodes  $a_2, a_3$  and  $a_4$  (otherwise, by putting one independent node in  $\widehat{rx}$ ,  $|N_{i-1} \cup N_i|$  would be greater than six). It is obvious that  $\widehat{a_4 v_{i+1} a_2} > 180^\circ$  (note that  $\widehat{a_4 v_{i+1} a_2}$  is the angle obtained by **counterclockwise** rotation of  $v_{i+1} a_2$  toward  $v_{i+1} a_4$ ). Assuming  $|U_{i+1}| = 3$ , either one or two of the independent nodes in it will also be in  $N_{i+2}$ . In the case that  $|N_{i+2} \cap U_{i+1}| = 2$ ,  $N_{i+2}$  has the same property as of  $N_{i+1}$  and we can continue the reasoning with the relation between  $N_{i+3}$  and  $U_{i+2}$ .

If only one independent node lies in the intersection area (as it is assumed in figure 4,) since **clockwise** angle between  $a_2 v_{i+1}$  and  $a_4 v_{i+1}$  is less than  $180^\circ$ , according to the proof available in [21], independent nodes in  $U_{i+2}$  have to be inside the  $\widehat{b_2 b_4}$  sector area. Since  $\widehat{b_1 v_{i+1} b_2} \leq 180^\circ$ , at most three independent nodes can lie in it. Therefore, when  $|U_{i+2}| = 3$ , the same angular property of these nodes enforces the same condition on the nodes in  $U_{i+3}$ .

The above reasoning can be extended to all  $U_h, h > i$ , while  $\forall s, i \leq s < h, U_s$  contains exactly three independent nodes. This immediately results in the fact that for all  $v_j, j > i, |U_j|$  is at most three unless the semi-exclusive neighboring set of a vertex on the path from  $v_i$  to  $v_j$  contains less than three independent nodes.  $\square$

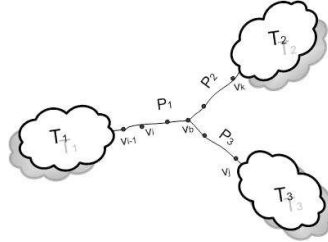


Fig. 5. Fork subgraph

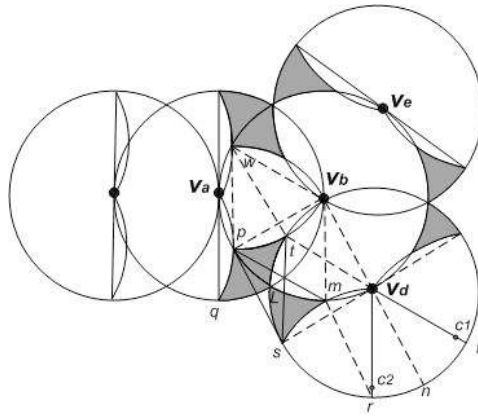


Fig. 6. Five vertices in the branch location and their neighboring areas

**Lemma 4.** Assume a preorder traversal of the vertices in  $mcds(G)$ . The fork subgraph and dominated independent nodes with the following properties doesn't exist:(figure 5, the root for traversal is assumed to be in  $T_1$ )

- $v_b$  is the vertex where paths from  $v_i$  to  $v_j$  and  $v_k$  get separated.
- $|U_i| = |U_j| = |U_k| = 4, |U_b| = 2$
- For every vertex  $v_h$  on the path from  $v_b$  to  $v_i, v_j$  or  $v_k, |U_h| = 3$

*Proof.* Figure 6 depicts  $v_b$  and its immediate neighbors in  $mcds(G)$ . Without loss of generality, assume  $\widehat{v_d v_b v_e} = \widehat{v_d v_b v_a} = \widehat{v_a v_b v_e} = 120^\circ$  (we have shown in the appendix why this assumption does not cause loss of generality.) Three cases are possible for two independent nodes in  $U_b$ :

First, assume  $|U_b \cap N_e| = 2$  and  $|U_b \cap N_d| = 0$ . As an immediate consequence, the vertices in the path followed by vertex  $v_e$  has the specification of the vertices in lemma 3 and for every  $v_h$  on the path followed from  $v_b$ , the semi-exclusive neighboring set  $U_h$  contain at most three independent nodes while for all  $v_s$  on the path between  $v_e$  and  $v_h, |U_s| = 3$ . This completes the proof of the lemma, assuming the first case.

In the second case, assume  $|U_b \cap N_e| = 2$  and  $|U_b \cap N_d| = 1$ . Since the path followed by vertex  $v_e$  has the same condition of the previous case, the lemma is also true in this case.

In the third case, assume  $|U_b \cap N_e| = 1$  and  $|U_b \cap N_d| = 1$ . We will show that the location of dark areas in figure 6 is such that either  $v_d$  or  $v_e$  has the properties of the nodes in lemma 3 (the nodes in semi-exclusive set are limited to a  $180^\circ$  sector), and consequently, the path continued from it cannot lead to a  $U_k$  with four dominating independent nodes unless there is a vertex  $v_h$  on the path between  $v_b$  and  $v_k$ , where  $U_h$  contains less than three independent nodes.

As figure 6 depicts, since the total neighboring area is maximum,  $\widehat{v_e v_b v_d} = 120^\circ$ . Consequently, because of the following geometric properties, the distance between the points in dark areas  $pql$  and  $msl$  is less than or equal to a unit. Hence, only one of them can include an independent node. Since  $rm = mv_b = v_b v_d = v_d r = 1$ , therefore  $rm \parallel mv_b \parallel v_b v_d \parallel v_d r$ . This follows that  $rm \parallel v_d n$  and since  $\widehat{nv_d r} = 30^\circ$ , therefore,  $\widehat{mr v_d} = 30^\circ$ . Hence, we can conclude that  $rm$  is dividing  $\widehat{sm v_d}$  to two equal sections and  $\widehat{sm} = \widehat{mv_d}$ . Similarly  $\widehat{lm} = \widehat{v_d m} = \widehat{pl} = 30^\circ$  and hence  $\widehat{pm} = 60^\circ$ . This results in  $pm = v_b p = v_b m = 1$ .

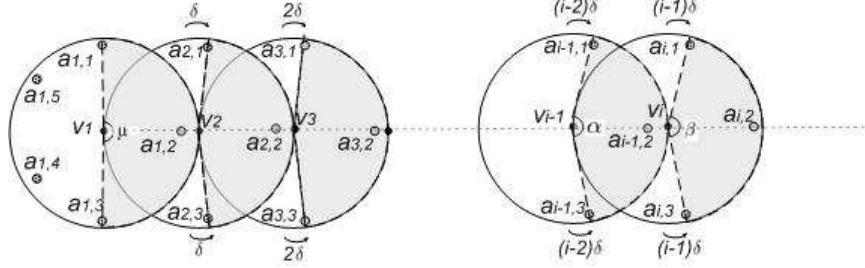
To show that all point in two dark area are closer than a unit to each other, it remains to prove that  $ps \leq 1$ . According to the assumption,  $\widehat{pv_b v_d} = \widehat{sv_d v_b} = 90^\circ$  and since  $pv_b = v_b v_d = sv_d = 1$ , consequently,  $pv_b v_d s$  is a square and  $ps = 1$ , which completes the claim.

**Note that**  $c_1$  and  $c_2$  are two independent nodes in  $U_d$  and the  $c_1 v_d c_2 = 60^\circ$ .  $v_d r$  and  $v_d t$  are achieved by connecting  $v_d$  to  $c_1$  and  $c_2$  respectively. Sector  $smv_d r$  and similar sectors are obtained by drawing the *unit arc* [21] of independent nodes in  $U_d$ .

By the same reasoning, two other pairs of dark areas have the same property, leading to the fact that either  $U_e$  or  $U_d$  is limited to a  $180^\circ$  sector. This follows the property of lemma 3 and as it was shown, the path continuing from such vertex cannot lead to a  $U_k$  containing four independent nodes while all the  $U_h$  between  $w$  and  $k$  has three independent nodes in their semi-exclusive neighboring areas and this completes the proof for lemma 4. □

**Theorem 1.** *For any unit disk graph  $G$ , the size of a maximal independent set is at most  $3c_{ds}(G) + 3$  where  $c_{ds}(G)$  is the size of a minimum connected dominating set.*

*Proof.* Let  $T$  be a subgraph induced by a minimum connected dominating set in the given unit disk graph. Let  $|T|$  be the number of vertices in  $G$ . We will show by induction on  $|T|$  that there exists at most  $3|T| + 3$  independent nodes in the neighboring area of  $T$ . According to lemmas 1 and 2, the claim is true for  $|T| = 1$  and 2. Now, we assume that  $|T| \geq 3$ . Consider the semi-exclusive neighboring area of a leaf vertex  $v_l$  in an arbitrary tree that spans all the vertices of  $T$ . By removing  $v_l$  from the MCDS tree and removing all the independent nodes in its semi-exclusive area from graph  $G$ , by induction hypothesis, the neighboring area of  $T' = T - v_l$  contains at most  $3c_{ds}(T') + 3$  independent vertices. Next, we consider  $v_l$  and its semi-exclusive neighboring area  $U_l$ . If  $U_l$  includes less



**Fig. 7.** Graph  $G$ , with  $mis(G) = 3 \cdot cds(G) + 2$ , Note that  $\mu = 180^\circ - \epsilon$ ,  $\epsilon \rightarrow 0$

than three independent vertices, then the proof is complete. In the case that  $U_l$  contains four independent nodes, consider a preorder traversal of  $T$ .

According to lemma 3 on the path between  $v_l$  and the closest predecessor vertex  $v_j$  of  $T$  that  $U_j$  contains four independent nodes, there exists at least a vertex  $v_k$ , where  $U_k$  includes less than three independent nodes. Assume  $T'' = T - \{v_k, v_{k+1}, \dots, v_l\}$ , according to induction hypothesis, the neighboring area of  $T''$  contains at most  $3cds(T'') + 3$ . Now we consider the subtree  $P$  rooted by  $v_j$  which contains  $v_m$ ,  $m > j$ . The number of independent nodes lie in the exclusive neighboring area of nodes in  $P$  is less than  $3|P|$ . This is because node  $e$  is the only vertex in  $P$  that its semi-exclusive neighboring set includes four nodes (more than three nodes), which cannot make counter example since  $U_j$  includes only two independent nodes ( $v_j$  in  $P$ ).

Furthermore, according to lemma 4 another branch cannot exist in the subtree which would lead to a vertex  $v_h$  containing more than three nodes in its semi-exclusive neighboring area unless there is a vertex in the branch and on the path to  $v_h$  with less than three independent nodes in its semi-exclusive neighboring area. Consequently, the neighboring area of  $T = T'' + P$  consists less than  $3|T| + 3$  independent nodes and this completes the proof of the theorem.  $\square$

**Corollary 1.** For approximation algorithms in [19], [6] for the minimum connected dominating set, the performance ratio can be reduced from 7.8 to 6.

#### 4.1 Tightness

**Theorem 2.** The upper bound  $3 \cdot cds(G) + C$  for the size of maximal independent set in unit disk graph  $G$ , where  $cds(G)$  is the size of the minimum connected dominating set and  $C$  is a constant, is tight.

*Proof.* To prove the theorem, we construct an extendable graph topology for vertices in a MCDS and MIS subgraphs such that for each  $cds(G) = n, n \in N$ , we have:

$$mis(G) = 3 \cdot cds(G) + C$$

Consider figure 7, where black vertices represent the nodes in the minimum connected dominating set and gray nodes are the set of independent vertices in

a maximal independent set. Most of the edges that are connecting gray vertices are not showed for clarity of the figure, however, since the graph is a UDG, if  $|uv| \leq 1$  there is an edge between  $u$  and  $v$ . Vertices in MCDS subgraph are all on a straight line and  $\forall i, |v_i v_{i+1}| = 1$ . As it is depicted in figure 7, the placement of the independent vertices in the MIS subgraph is such that  $N(v_1)$  includes five independent nodes. For each vertex  $v_i, i > 1, U_i$  includes three independent nodes, which results in the factor of 3 for the relation between the size MIS and the size of MCDS. Lets consider the relation of the semi-exclusive areas  $U_{i-1}, U_i,$  and  $U_{i+1}$ . As shown in figure 7, assume the sector containing three independent nodes in  $U_{i-1}$  is  $\alpha$  degree wide (The sector obtained by clockwise rotation of  $v_{i-1}a_{i-1,1}$  to  $v_{i-1}a_{i-1,3}$ ). According to lemma 3, the sector in  $U_i$  that can contain independent nodes is  $\beta$  degrees wide, where  $\beta$  is smaller than  $\alpha$  ( $\beta = \alpha - 2\delta$ ). In addition, according to lemma 1,  $\beta$  should be larger than  $120^\circ$  to be capable of containing three independent nodes. To construct a graph that meets the above relation with  $cds(G) = n$ , we define:

$$\delta = \frac{30}{n-1}$$

$$a_{i,1}\widehat{v_i}a_{i,2} = a_{i,2}\widehat{v_i}a_{i,3} = 90 - (i-1)\delta - \epsilon$$

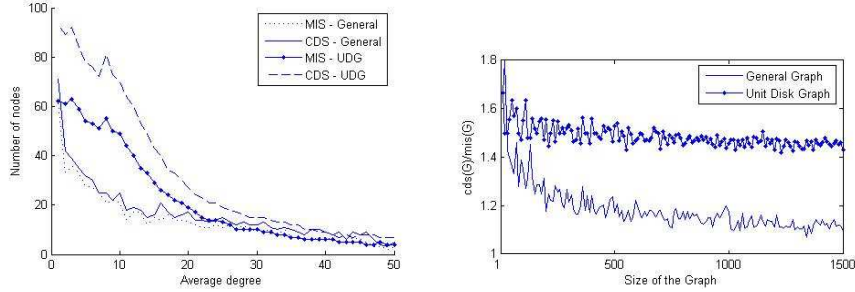
Therefore,

$$|a_{i,j}a_{i-1,k}| > 1, \begin{cases} 2 \leq i \leq n \\ 1 \leq j \leq 3 \\ 1 \leq k \leq 3 \end{cases}$$

which are the immediate results of lemmas 3 and 2. Hence, we place two nodes in each semi-exclusive area on the edges of the mentioned sector and one on the line connecting MCDS vertices. Consequently, the number of nodes in MIS subgraph is  $3(n-1) + C$ , where  $C$  is the number of independent nodes in  $N(v_1)$ . Up to five independent nodes can lie in this unit disk, and consequently  $mis(G) = 3 \cdot cds(G) + 2$  which completes the proof.  $\square$

## 5 Experimental Analysis

In this section, a practical analysis of the relation between connected dominating set and maximal independent set in wireless networks is presented. Since finding MCDS is NP-Hard, finding the size of optimal CDS in reasonable amount of time is not possible. Here, we will use the linear approximation algorithm introduced by Guha et. al. [11], which finds a CDS in an arbitrary graph with maximum cardinality of  $2(1 + H(\Delta)) \cdot |OPT_{mcds}|$ , where  $\Delta$  and  $H$  denote the maximum degree and the harmonic function respectively. The algorithm for constructing MIS is such that in each iteration a random node is selected to be a member of MIS, which results in the elimination of its neighbors from the list of candidate MIS members. This process is continued until all nodes in the graph are either selected or eliminated from being a member of MIS.



**Fig. 8.**  $cds(G)$  and  $mis(G)$  vs. the average degree in UDG (Left) and  $\frac{cds(G)}{mis(G)}$  vs. size of the graph (Right)

To run the experiments, 400 nodes were distributed randomly in a space of size  $100 \times 100$ . We controlled the random graph generation phase by setting the average degree of the nodes in the graphs. It can be seen in figure 8 that as the average degree of the graph is increasing, the connectivity of the nodes increases resulting in significant decrease in the size of both  $mis(G)$  and  $cds(G)$ . However, both sizes maintain the same slope as the connectivity increases.

Figure 8 also compares  $\frac{cds(G)}{mis(G)}$  resulted from a general graph and one resulted from a unit disk graph. While maintaining the size of the space fixed, the size of the connected graph is increased from 10 to 1500 (resulting in a denser graph). Here, a *general graph* is a graph that is structured by vertices with arbitrary coverage radii. We observe that in the general graph the size of MIS is still comparable to the size of CDS. While the ratio is approximately constant in the UDG, the size of MIS and CDS subgraphs of the general graph are getting closer. Hence, the probability that the ratio does not meet the theoretical bound of UDG is higher in the denser graph. However, since in both cases  $\frac{cds(G)}{mis(G)} \geq 1$ , it can practically be assumed that the theoretical bound for unit disk graph is still practically applicable for wireless networks.

## 6 Conclusion

Applying protocols and techniques based on minimum connected dominating set plays an important role in improving the performance and energy consumption of wireless sensor and ad-hoc networks. A common approach to distributively approximate MCDS is to construct maximal independent set and then convert it to a connected tree. In this paper, we decreased the known bound between the size of maximal independent set and the size of minimum connected dominating set. We proved that  $mis(G) \leq 3 \cdot cds(G) + 3$ , where  $mis(G)$  and  $cds(G)$  denote the number of nodes in MIS and MCDS subgraphs respectively. Also, by proving that this bound is tight, we show that it cannot be improved anymore. The new relation results in the new approximation ratio of 6 for distributed algorithm proposed by [6], [19]. In addition, we analyzed the behavior of general wireless networks in terms of size

of MIS and MCDS subgraphs and concluded that the theoretical bound for unit disk graph is still practically applicable for wireless networks.

## References

1. Alzoubi, K.M., Wan, P.-J., Frieder, O.: Message-optimal connected dominating sets in mobile ad hoc networks. In: *MobiHoc 2002: Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing*, pp. 157–164. ACM, New York (2002)
2. Clark, C.C.B., Johnson, D.: Unit disk graphs. *Discrete Mathematics* 86, 165–177 (1990)
3. Das, R.S.B., Bharghavan, V.: Routing in ad-hoc networks using a virtual backbone. In: *ICCCN 1997*, pp. 1–20 (1997)
4. Banik, S.M., Radhakrishnan, S.: Minimizing broadcast latency in ad hoc wireless networks. In: *ACM-SE 45: Proceedings of the 45th annual southeast regional conference*, pp. 533–534. ACM, New York (2007)
5. Bevan Das, R.S., Bharghavan, V.: Routing in ad-hoc networks using a spine. In: *International Conference on Computers and Communication Networks*, pp. 376–380 (1997)
6. Cardei, M., Cheng, X., Cheng, X., Du, D.-Z.: Connected domination in multihop ad hoc wireless networks. In: *Proceedings of Sixth International Conference in Computer Science and Informatics, CSI*
7. Das, B., Bharghavan, V.: Routing in ad-hoc networks using minimum connected dominating sets. In: *ICC*, pp. 376–380 (1997)
8. De Gaudenzi, T.G.F.L.M., Garde, R.: Ds-CDMA techniques for mobile and personal satellite communications: An overview. In: *IEEE Second Symposium on Communications and Vehicular Technology* (1994)
9. Deb, B., Nath, B.: On the node-scheduling approach to topology control in ad hoc networks. In: *MobiHoc 2005: Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing*, pp. 14–26. ACM, New York (2005)
10. Funke, S., Kesselman, A., Meyer, U., Segal, M.: A simple improved distributed algorithm for minimum CDS in unit disk graphs. *ACM Transaction on Sensor Networks* 2(3), 444–453 (2006)
11. Guha, S., Khuller, S.: Approximation algorithms for connected dominating sets. *Algorithmica* 20(4), 374–387 (1998)
12. Blum, A.T.J., Ding, M., Cheng, X.: Connected dominating set in sensor networks and manets. In: Du, D.-Z., Pardalos, P. (eds.) *Handbook of Combinatorial Optimization*, pp. 329–369. Kluwer Academic Publishers, Dordrecht (2004)
13. Marathe, M.V., Brey, H., Hunt III, H.B., Ravi, S.S., Rosenkrantz, D.J.: Simple heuristics for unit disk graphs. *Networks* 25, 59–68 (1995)
14. Min, M., Du, H., Jia, X., Huang, C.X., Huang, S.C.-H., Wu, W.: Improving construction for connected dominating set with steiner tree in wireless sensor networks. *J. of Global Optimization* 35(1), 111–119 (2006)
15. Mnif, K., Rong, B., Kadoch, M.: Virtual backbone based on mcds for topology control in wireless ad hoc networks. In: *PE-WASUN 2005: Proceedings of the 2nd ACM international workshop on Performance evaluation of wireless ad hoc, sensor, and ubiquitous networks*, pp. 230–233. ACM, New York (2005)
16. Ni, S.-Y., Tseng, Y.-C., Chen, Y.-S., Sheu, J.-P.: The broadcast storm problem in a mobile ad hoc network. In: *MobiCom 1999: Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking*, pp. 151–162. ACM, New York (1999)

17. Shaikh, J.A., Solano, J., Stojmenovic, I., Wu, J.: New metrics for dominating set based energy efficient activity scheduling in ad hoc networks. In: LCN 2003: Proceedings of the 28th Annual IEEE International Conference on Local Computer Networks, Washington, DC, USA, p. 726. IEEE Computer Society Press, Los Alamitos (2003)
18. Stojmenovic, I., Seddigh, M., Zunic, J.: Dominating sets and neighbor elimination-based broadcasting algorithms in wireless networks. *IEEE Transactions on Parallel and Distributed Systems* 13(1), 14–25 (2002)
19. Wan, P., Alzoubi, K., Frieder, O.: Distributed construction of connected dominating set in wireless ad hoc networks (2002)
20. Wu, J., Li, H.: On calculating connected dominating set for efficient routing in ad hoc wireless networks. In: DIALM 1999: Proceedings of the 3rd international workshop on Discrete algorithms and methods for mobile computing and communications, pp. 7–14. ACM, New York (1999)
21. Wu, W., Du, H., Jia, X., Li, Y., Huang, S.C.-H.: Minimum connected dominating sets and maximal independent sets in unit disk graphs. *Theoretical Computer Science* 352(1), 1–7 (2006)

### Appendix: Proof of Lemma 4 for a Branch Location with Unknown Angles

In this appendix we provide a proof to show why lemma 4 is still correct in general cases. To generalize the case, we consider rotating  $v_e$  and  $v_d$  in figure 6 for  $d_1$  and  $d_2$  degrees. Therefore,  $\widehat{v_a v_b v_e} = 120 + d_1^\circ$  and  $\widehat{v_a v_b v_d} = 120 + d_2^\circ$  (See figure 9).  $d_1$  and  $d_2$  are positive if the rotation is clockwise and are negative if the rotation is counterclockwise.

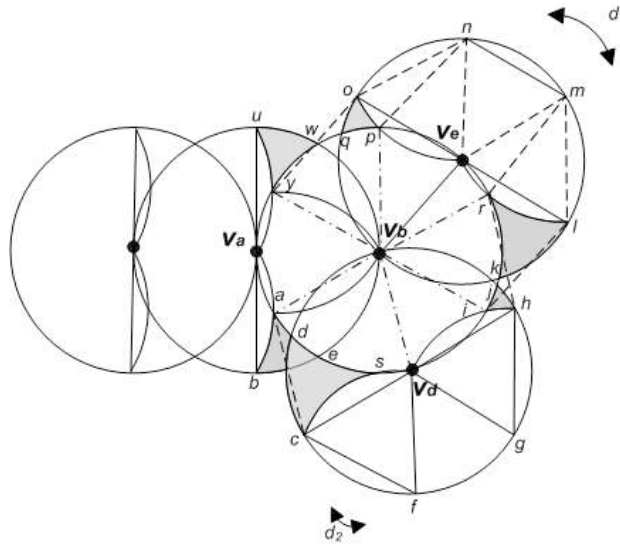


Fig. 9. The general case for the branch

Note that in figure 9, diameters  $ol$  and  $ch$  are such that  $|ac| = |yo| = 1$ . Also,  $f, g, m$  and  $n$  are location of independent nodes lying in exclusive neighboring sets of  $v_e$  and  $v_d$ . Arcs  $\widehat{cv_d}$ ,  $\widehat{v_dh}$ ,  $\widehat{lv_e}$  and  $\widehat{v_eo}$  are from the perimeters of unit arcs containing  $v_d$  and one of  $f, g, m$  or  $n$ , respectively. We first prove that the distance between all the points in  $rlk$  and  $hij$  is less than a unit by showing that  $|hr| = |lj| = 1$ . First, note that  $|ac| = |cv_d| = |v_dv_b| = |v_ba| = 1$  and hence,  $acv_dv_b$  is a parallelogram. Consequently,  $\widehat{av_bv_d} = \widehat{v_bv_dh}$ . Now by showing that  $v_br$  and  $v_ba$  are on a straight line, we show that  $ra$  is a diameter and as a result,  $v_dh \parallel v_br$  (since  $v_dh \parallel v_ba$ ). Now, assume parallelograms  $v_bv_emr$  and  $v_bv_enp$ . Since  $\widehat{nm} = 60^\circ$  and  $\widehat{mv_ev_b} + \widehat{v_bv_en} = 300^\circ$  ( $onv_e$  and  $lmv_e$  are two unit arcs and  $\widehat{ol} = 180^\circ$ , therefore  $\widehat{mn} = 60^\circ$ ), hence  $\widehat{rv_bp} = \widehat{rv_bv_e} + \widehat{pv_bv_e} = 360^\circ - (\widehat{mv_ev_b} + \widehat{v_bv_en}) = 60^\circ$ . Considering that  $\widehat{rv_e}$  and  $\widehat{v_e\hat{p}}$  are both  $30^\circ$  before rotating  $v_e$ ,  $\widehat{rv_e} = 30 - d_1^\circ$  and  $\widehat{pv_e} = 30 + d_1^\circ$ . In addition, according to the fact that  $\widehat{wq} = d_1^\circ$  (since it is the result of rotating  $v_d$ ), we can conclude that  $\widehat{ar} = 180^\circ$ . Therefore,  $v_dh$  is parallel to  $v_br$ , and since  $|v_dh| = |v_br| = |v_dv_b| = 1$ ,  $v_bv_drh$  is parallelogram, and hence  $|rh| = 1$ .

With similar reasoning (considering  $ilv_ev_b$  and showing that it is a parallelogram), we can see that  $|il| = 1$  and hence distance between all the points in regions  $krl$  and  $ijh$  are less than one. To show the same condition for other pairs of dark areas, similar set of proof will apply, resulting in the fact that rotating  $v_e$  and  $v_d$  around  $v_b$  doesn't have any effect on the generality of lemma 4.