Technical Report

Department of Computer Science
and Engineering
University of Minnesota
4-192 EECS Building
200 Union Street SE
Minneapolis, MN 55455-0159 USA

TR 04-048

On Greedy Construction of Connected Dominating Sets in Wireless Networks

Yingshu Li, My T. Thai, Feng Wang, Chih-wei Yi, Pengjun Wan, and Ding-zhu Du

December 16, 2004

On Greedy Construction of Connected Dominating Sets in Wireless Networks

Yingshu Li * My Thai * Feng Wang * Chih-Wei Yi †
Peng-Jun Wan † Ding-Zhu Du *

Abstract

Since no fixed infrastructure and no centralized management present in wireless networks, a Connected Dominating Set (CDS) of the graph representing the network is widely used as the *virtual backbone* and plays an important role in the network. Constructing a minimum CDS is NP-hard. Many CDS construction algorithms have been designed. In this paper, we propose a new greedy algorithm, called S-MIS, with the help of Steiner tree that can construct a CDS within a factor of $4.8 + \ln 5$ from the optimal solution. We also introduce the distributed version of this algorithm. The theoretical proof shows that our algorithm is better than the current best performance ratio which is 6.8. A simulation is conducted to compare S-MIS with its variation which is rS-MIS. The simulation shows that the sizes of the CDSs generated by S-MIS and rS-MIS are almost the same.

*Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455, USA. Email:{yili,mythai,fwang,dzd}@cs.umn.edu. Support in part by National Science Foundation under grant CCR-0242520.

†Department of Computer Science, Illinois Institute of Technology, Chicago, IL 60616. Email: {wan,yich}@cs.iit.edu.

1 Introduction

Wireless networks are bringing more and more benefits to us. Wireless ad hoc networks are used in many fields such as battlefield, disaster recovery, conferences and concerts. Wireless sensor networks can also be employed to provide services in military fields, environmental detection, and agriculture. There are no fixed or pre-defined infrastructure in wireless networks and hosts in a wireless network communicate via a shared medium either through a single hop or multihops. Usually, there is no central management in wireless networks either. Therefore, each host also needs to serve as a router so that it can forward the received messages according to some routing protocols.

Broadcast and multicast are two popular communication methods in wireless networks. Broadcast is to send messages from one host to all the other hosts in the network. Multicast is to send messages from one host to a group of hosts in the network. Due to the different transmission medias and methods of wired networks and wireless networks, the broadcast and multicast protocols in wired networks are not suitable for wireless networks. Currently, virtual backbones are usually used to support broadcast and multicast in wireless networks and a Connected Dominating Set (CDS) is the best candidate to work as a virtual backbone stimulated by the characteristics of wireless networks.

In this study, we use $G = (V, E)$ to represent a wireless network where V is the set of hosts in the network and E represents all the links in the network. We assume that all the hosts are deployed in a 2-D plane and their maximum transmission range are the same. Thus the resultant topology of the network is modelled as an undirected *Unit Disk Graph* (UDG) [7]. In the context of graph theory, we call a host as a node. A Dominating Set (DS) of a graph G is a subset $S \subset V$ such that for each node in G, it either belongs to S or has at least one neighbor in S. A CDS is a DS which induces a connected subgraph. The nodes in the CDS are called the *dominators*, otherwise,

dominatees. It is desirable to build a Minimum-sized Connected Dominating Set (MCDS) in consideration of reducing more traffic and maintenance. However, the construction of an MCDS in a UDG is proved to be NP-hard in [7]. Figure 1 gives an example UDG containing a CDS which is also a MCDS.

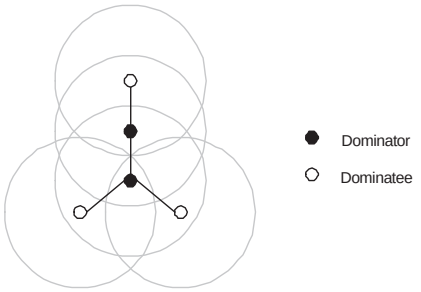

Figure 1: A UDG with a CDS

With the help of the CDS, routing including broadcast and multicast is easier and can adapt quickly to topology changes of a network. Only the nodes in the CDS need to maintain the routing information. Furthermore, if there is no topology changes in the subgraph induced by the CDS, there is no need to update the routing information, which reduces both storage and message complexities. If a dominatee wants to deliver a message to another dominatee, it first sends the message to its dominator. Then the search space for the route is reduced to the CDS. After the message is relayed to the destination's dominator, this dominator will deliver the message to the final destination.

To construct a CDS, we utilize an Maximal Independent Set (MIS) which is also a subset of all the nodes in the network. The nodes in an MIS are pairwise nonadjacent and no more nodes can be added to remain the non-adjacency property of this set. Thus each node which is not in the MIS is adjacent to at least one node in the MIS. Thus an MIS is a DS. If we connect the nodes in an MIS through some nodes not in the MIS (we call them Steiner nodes), a CDS is then constructed. We use performance ratio (PR) to evaluate a CDS construction algorithm. PR is defined as the ratio of

the size of the constructed CDS over the size of MCDS. In this paper, we propose a new greedy algorithm with PR of $4.8 + ln5$, which is better than the current best one.

The remainder of this paper is organized as follows. Section 2 briefly describes the related research works in literature. A new greedy algorithm for constructing a CDS and the analysis of this algorithm are illustrated in Section 3. The simulation results are presented in Section 4. Section 5 describes the distributed version of this algorithm. Finally, Section 6 ends this paper with a conclusion and some discussions.

2 Related Work

The idea of using a CDS as a virtual backbone for routing was proposed in [11]. Then many efforts have been made to design approximations algorithms for CDS construction. In most of the CDS construction algorithms, a coloring mechanism is used where initially all the nodes are white, a dominator is colored black and a dominatee is colored grey.

Guha and Khuller [14] first proposed two 2-phase centralized greedy algorithms to construct CDSs in general graphs. The number of the white neighbors of each node or a pair of nodes (a dominatee with one of its white neighbor) is the greedy function. The one with the largest such number will become dominator(s) at each step. In the first algorithm, the CDS is built up at one node, then the searching space for the next dominator(s) is restricted to the current dominatees and the CDS expands until there is no white nodes. In the second algorithm, all the possible dominators are determined in the first phase, then they are connected through some intermediate nodes in the second phase. Das *et al.* [8, 9, 23] gave the implementations of the algorithms in [14]. Ruan *et al.* [21] then designed a 1-phase greedy algorithm with PR of $2 + ln\Delta$ where Δ is the maximum degree in the graph.

Wu and Li [27] proposed a distributed algorithm where each node knows the connectivity information within the 2-hop neighborhood, but they did not specify the

PR. If a node has two unconnected neighbors, it becomes a dominator. The generated CDS is easy to maintain. But the size of the CDS is large. In [25], the authors gave out the PR of Wu and Li's algorithm which is $O(n)$.

In the recent years, it is popular to construct a CDS by first constructing an MIS, then by connecting the nodes in the MIS, a CDS is generated.

Alzoubi *et. al.* [1, 2, 25] made a great improvement by proposing two 2-phase distributed algorithms. A spanning tree is constructed first and then each node in the tree is labelled as either a dominator or a dominatee. The algorithms are employed in a UDG to obtain a constant PR which is 8.

Cardei *et. al.* [4] presented a 2-phase distributed algorithm. This algorithm requires a leader to be selected at the beginning of the first phase. The leader first becomes a dominator making its neighbors dominatees. They introduce a new *active* state for white nodes. A white node becomes activate only after one of its neighbors becomes a dominatee. All the active nodes will compete to become a dominator based on the pair (the number of white neighbors, ID). The improvement over Alzoubi *et. al.*'s algorithms is that the root do not need to wait for the COMPLETE messages from the furthest nodes. The root initiates the connecting phase just after it receives NUMOFBLACKNEIGHBORS from all of its neighbors. The PR of this algorithm is also 8.

Alzoubi *et. al.* [3] noticed the difficulty of the maintenance of the CDS constructed by their previous algorithms and designed a localized distributed 2-phase algorithm which is good at maintenance in general graphs. An MIS is generated in a distributed fashion without building a tree or selecting a leader. Once a node knows that it has the smallest ID within its 1-hop neighborhood, this node becomes a dominator. After there are no white nodes, the dominators are responsible for identifying a path to connect all the dominators. In this algorithm, no network connectivity information is utilized and the PR is 192. In [15], the authors gave another localized distributed algorithm with PR of 172.

Among all the approximation algorithms for CDS construction in UDGs, the best known PR is 6.8 [18]. In this paper, we will present an algorithm with PR of $(4.8 + \ln 5)$, which can also be implemented as a distributed algorithm.

Our main idea is to employ a Steiner tree in the second step to connect the nodes in the MIS. In a graph, a *Steiner tree* for a given subset of nodes, called *terminals*, is a tree interconnecting all the terminals. Every node other than the terminals in the Steiner tree is called a *Steiner node*. Clearly, a small number of Steiner nodes is expected in order to obtain a small CDS. Therefore, we will study the following Steiner tree problem in UDGs.

Steiner Tree with Minimum Number of Steiner Nodes (ST-MSN): Given a UDG G and a subset P of nodes, compute a Steiner tree for P with the minimum number of Steiner nodes.

The ST-MSN problem in UDGs has not been studied very much, unlike its geometric version in the Euclidean plane, which has been studied extensively [16, 5, 26]. However, some results cannot be extended to UDGs. For example, two points with distance 2 can be connected with a Steiner point in the Euclidean plane. But, two nodes with distance 2 may not be able to be connected by a Steiner node since such a node may not exist. Fortunately, a 3-approximation algorithm for ST-MSN can be extended from the the Euclidean plane to UDGs with a quite different proof, which becomes a fundamental part in our approximation algorithm.

3 The S-MIS Algorithm

In this section, a new greedy algorithm, which is called S-MIS, is introduced. S-MIS consists of two steps. At the first step, we construct a MIS. An important property of an independent set is that [17]

Lemma 1 *In a unit disk graph, every node is adjacent to at most five independent nodes.*

It is well known that every MIS is also a dominating set. The following lemma is a recent result in [28] about the relation between the size of the MIS and the MCDS in a UDG.

Lemma 2 *In any unit disk graph, the size of every maximal independent set is upper-bounded by* $3.8opt + 1.2$ *where opt is the size of the minimum connected dominating set in this unit disk graph.*

Especially, Wan [24] and Cheng [6] constructed a MIS having the following property.

Lemma 3 *Any pair of complementary subsets of the MIS are separate by exactly two hops.*

We assume throughout this paper that the MIS satisfies Lemma 3.

At the second step, we employ a greedy approximation for the ST-MSN to interconnect the nodes in the MIS. We will show that this greedy approximation has PR of $1 + \ln 5$. Note that the size of the optimal solution for the ST-MSN cannot exceed the size of the MCDS since the latter can also interconnect the MIS. Therefore, we spend at most $(1 + \ln 5)opt$ Steiner nodes in the second step. By Lemma 2, the resulting CDS would have size bounded by $(4.8 + \ln 5)opt + 1.2$.

Theorem 1 *The S-MIS algorithm produces a CDS with size bounded by* $(4.8+\ln 5)opt+1.2$ *where opt is the size of the MCDS.*

We can use the method in [24] or Cheng [6] to construct a MIS at the first step. At the second step, a greedy algorithm A is employed, which is described as following.

Algorithm A: Input a MIS and mark all the nodes in this MIS black. Mark the other nodes in the UDG grey. In the following, we will change some grey nodes to blue according to some certain rules. A *black-blue component* is a connected component of the subgraph induced only by black and blue nodes and by *ignoring connections between blue nodes*.

for $i = 5, 4, 3, 2$ **do**

while there exists a grey node adjacent to at least i

black nodes in different black-blue components

do

change its color from grey to blue;

end-while;

return all blue nodes.

We know that Theorem 1 follows immediately from the following Theorem.

Theorem 2 *Let T^* be an optimal tree for the ST-MSN problem on an input MIS. Then the number of the output blue nodes is at most $(1 + \ln 5)C(T^*)$, where $C(T^*)$ is the number of the Steiner nodes in T^*.*

Proof. If the input MIS contains only one node, then $C(T^*) = 0$. The theorem is trivial. Thus, we may assume that the input MIS contains at least two nodes and hence $C(T^*) \geq 1$. Let n be the number of the black nodes. Let $x_1, ..., x_k$ be the blue nodes in the order of appearance in the Algorithm A. Let a_i be the number of the black-blue components after $x_1, ..., x_i$ turns from grey to blue. Note that every black-blue component contains a black node which is adjacent to a Steiner node of T^*. Therefore, there exists a Steiner node x_i which is adjacent to at least $a_i/C(T^*)$ black nodes in different black-blue components, so does x_{i+1}. Hence,

$$a_{i+1} \leq a_i - a_i/C(T^*).$$

Note that $a_k = 1 \leq C(T^*)$ and $a_0 = n > C(T^*)$. There exists h, $1 \leq h \leq k$ such that $a_h \geq C(T^*)$ and $a_{h+1} < C(T^*)$. Now, we have

$$a_h \leq a_{h-1}(1 - \frac{1}{C(T^*)}) \leq ...$$

$$\leq a_0(1 - \frac{1}{C(T^*)})^h$$

$$\leq a_0 e^{-\frac{h}{C(T^*)}}$$

Here, we note that $1 + x \leq e^x$ for $x > -1$. Thus,

$$\frac{h}{C(T^*)} \leq \ln \frac{a_0}{a_h} \leq \ln \frac{n}{C(T^*)} \leq \ln 5.$$

Therefore,

$$k \leq h + 1 + a_{h+1} - 1 \leq (1 + \ln 5)C(T^*).$$

□

From Theorem 1 and Theorem 2, we conclude that the PR of the S-MIS algorithm is $(4.8 + \ln 5)opt + 1.2$.

4 Simulation Results

In the definition of a black-blue component, we ignore the connections between any two blue nodes. This is important to the proof of Theorem 2. Actually, ignoring those connections makes the following fact true: the connecting ability of each grey node cannot increase after some grey nodes change to blue. In other words, the number of the connected black-blue components is a submodular function on the grey nodes. Loss of this submodular property would make the theoretical analysis harder and the result would be worse (see [10] for the detailed discussion on this matter). Therefore, in order to see how this ignorance affects the performance of the S-MIS algorithm, we conduct a simulation to compare the results of the S-MIS algorithm and the revised S-MIS (rS-MIS) algorithm considering the connections between the blue nodes.

In this simulation, the number of the blue nodes is the measurement to evaluate the sizes of the CDSs generated by the S-MIS algorithm and the rS-MIS algorithm since

the number of the black nodes are the same for these two algorithms. Totally N hosts are randomly generated in a fixed 1000*1000 2-D square. The transmission range of each node is R. Only the connected networks are considered in this simulation. The algorithms are run 100 times for each group of N and R and the results averaged.

For $R \in [200, 800]$, we change N from 20 to 100. It is shown that the sizes of the CDSs generated by S-MIS and rS-MIS are quite similar to each other. Only occasionally, rS-MIS can generate a slightly smaller CDS. Let $P = \frac{\#\ blue\ nodes}{\#\ black\ nodes}$. Figure 2 illustrates the relation between P_{S-MIS} and P_{rS-MIS} when R is set to 400. In the figure, the two lines almost overlap with each other and have different values at only some points such as $N = 37$ and $N = 72$. Therefore, even if we do not consider the blue-blue connections, the result would not be affected greatly.

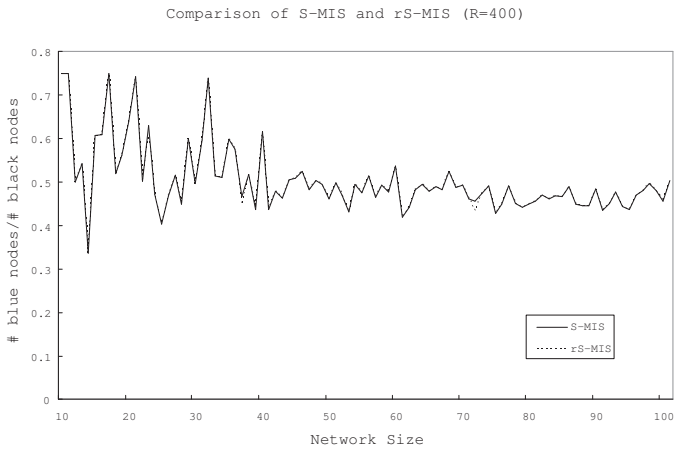

Figure 2: Performance comparison of S-MIS and rS-MIS

For different R, the improvement of rS-MIS over S-MIS, which is equal to $\frac{P_{S-MIS}-P_{rS-MIS}}{P_{S-MIS}} * 100\%$, is averaged for $N \in [10, 100]$. Figure 3 illustrates the result. Again, this figure shows that the improvement of rS-MIS over S-MIS is quite small.

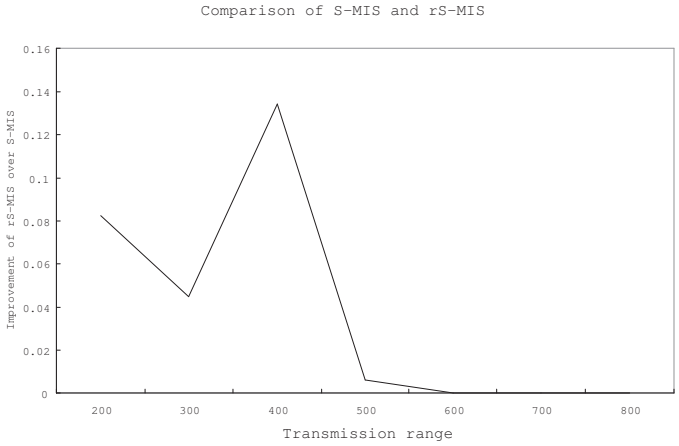

Figure 3: Performance comparison of S-MIS and rS-MIS

5 Distributed Implementation

In wireless networks, there is no centralized management and nodes may have mobility, therefore, distributed algorithms are expected. In this section, we briefly describe the distributed version of the S-MIS algorithm. There already exist several distributed algorithms for constructing a MIS in literature [24, 6]. Thus, we only introduce a distributed implementation of the greedy Algorithm A.

Each black or blue node carries a z-value which is an identification for the black component it belongs to, that is, all nodes with the same z-value form a black-blue component. Initially, the z-value of each black node equals its ID.

Grey nodes are ranked based on two values. The first one is the y-value which is the number of the adjacent black nodes in different black-blue components. The second one is its ID. The node with a larger y-value is ranked higher. If two grey nodes have the same y-value, then the one with a smaller ID is ranked higher.

A grey node is *adjacent* to a black-blue component if it is adjacent to a black node in the black-blue component. A grey node u is a *competitor* of another grey node v if u and v are adjacent to the same black-blue component. A grey node u is going to

change its color to blue if and only if u is ranked higher than every competitor of u.

Every grey node keeps two lists, a black list and a competitor list. The black list contains all the adjacent black nodes with their z-values, which enables the grey node to compute its y-value.

The competitor list contains all its competitors and their black lists so that each grey node can also compute the y-value of every competitor of it, which enables the grey node to make a decision on whether it should change its color nor not.

When a grey node u changes its color to blue, all its adjacent black-blue components are merged into one black-blue component and hence their z-values should be updated to the same one, say the smallest one among them. Meanwhile, all the competitors of u become the competitors of every competitor of u. Therefore, the competitor list of each competitor of u should also be updated. So, after u changes its color, u will send an UPDATE(u) message to all its neighbors. The message contains u' ID and its two lists.

When a black node v receives an UPDATE(u) message, it will update z, send out a COMPLETE(u) message and pass the UPDATE(u) message to its neighbors other than the nodes which already sent to the v UPDATE(u) or COMPLETE(u) message.

When a grey node receives an UPDATE(u) message, it updates both of its black and competitor lists and sends out a COMPLETE(u) message to its neighbors.

6 Conclusion

In this paper, we study the problem of constructing a CDS in wireless networks with the help of a Steiner tree. We propose a new greedy algorithm which is S-MIS with performance ratio of $(4.8 + \ln 5)opt + 1.2$. We also introduce the distributed version of this algorithm. In the S-MIS algorithm, we ignore the blue-blue connections when choosing the connectors for the black nodes. The simulation result shows that this ignorance does not affect the result much and the theoretical proof of the bound of

the size of the generated CDS still holds.

It is our interest to further investigate the maintenance of the CDS when nodes have mobility and the routing protocols based on the generated CDS. The work can also be extended to develop CDS construction algorithms when hosts in a network have different transmission ranges, not just be limited in UDGs.

References

[1] K.M. Alzoubi, P.-J. Wan and O. Frieder, New Distributed Algorithm for Connected Dominating Set in Wireleess Ad Hoc Networks, Proceedings of the 35th Hawaii International Conference on System Sciences, Big Island, Hawaii, 2002.

[2] K.M. Alzoubi, P.-J. Wan and O. Frieder, Distributed Heuristics for Connected Dominating Sets in Wireless Ad Hoc Networks, Journal of Communications and Networks, Vol.4, No.1, March 2002.

[3] K.M. Alzoubi, P.-J. Wan and O. Frieder, Message-Optimal Connected Dominating Sets in Mobile Ad Hoc Networks, *MOBIHOC*, EPFL Lausanne, Switzerland, 2002.

[4] M. Cardei, X. Cheng and D.-Z. Du, Connected Domination in Ad Hoc Wireless Networks, Proc. 6th International Conference on Computer Science and Informatics, 2002.

[5] D. Chen, D.-Z. Du, X. Hu, G.-H. Lin, L. Wang, and G. Xue: Approximations for Steiner trees with minimum number of Steiner points, *Journal of Global Optimization* Vol. 18, pp. 17-33, 2000.

[6] M. Cadei, M.X. Cheng, X. Cheng, and D.-Z. Du, Connected domination in ad hoc wireless networks, *Proc. the Sixth International Conference on Computer Science and Informatics (CS&I'2002)*, 2002.

[7] B. N. Clark, C. J. Colbourn and D. S. Johnson, Unit Disk Graphs, Discrete Mathematics, 86, Page 165-177, 1990.

[8] Bevan Das, Raghupathy Sivakumar and Vaduvur Bharghavan, Routing in Ad Hoc Networks Using a Spine, International Conference on Computers and Communications Networks, Las Vegas, NV. September 1997.

[9] Bevan Das and Vaduvur Bharghavan, Routing in Ad-Hoc Networks Using Minimum Connected Dominating Sets, International Conference on Communications, Montreal, Canada, June 1997.

[10] D.-Z. Du and K.-I Ko, Design and Analysis of Approximation Algorithms, Lecture Notes.

[11] A. Ephremides, J. Wieselthier, and D. Baker, A design concept for reliable mobile radio networks with frequency hopping signaling, Proceedings of the IEEE, 75, Page 56-73, 1987.

[12] Z. Gaspar and T. Tarnai, Upper bound of density for packing of equal circles in special domains in the plane, Periodica Polytechnica Ser. CIV. Eng. VOL. 44, NO. 1, pp. 13-32(2000).

[13] M. Gerla and J. T. Tsai, Multicluster, mobile, multimedia, radio network, ACM Journal on Wireless Networks, 1, pp. 255-265,1995.

[14] S. Guha and S. Khuller, Approximation algorithms for connected dominating sets, Algorithmica, 20(4), Page 374-387, April 1998.

[15] Yingshu Li, Shiwei Zhu, My T. Thai and Ding-Zhu Du, Localized Construction of Connected Dominating Set in Wireless Networks, NSF International Workshop on Theoretical Aspects of Wireless Ad Hoc, Sensor and Peer-to-Peer Networks (TAWN04), Chicago, June 2004.

[16] G.-H. Lin and G.L. Xue, Steiner tree problem with minimum number of Steiner points and bounded edge-length, *Information Processing Letters*, Vol. 69, pp. 53-57, 1999.

[17] M. V. Marathe *et al.*, Simple heuristics for unit disk graphs, *Networks*, vol. 25, pp. 59-68, 1995.

[18] Manki Min, Christina Xiao Huang, Scott C.-H. Huang, Weili Wu, Hongwei Du, Xiaohua Jia, Improving construction for connected dominating set with Steiner tree in wireless sensor networks, to appear in *Journal of Global Optimization*.

[19] J. Molnar, On the Packing of Unit Circles in a Convex Domain, Ann. Univ. Sci. Budapest. Eotvos, Sect. Math. Vol. 22/23, pp. 113-123, 1979/80.

[20] B. Ramamurthy, J. Iness and B. Mukherjee, Minimizing the number of optical amplifiers needed to support a multi-wavelength optical LAN/MAN, *Proc. IEEE INFOCOM'97*, pp. 261–268, April 1997.

[21] Lu Ruan, Hongwei Du, Xiaohua Jia, Weili Wu, Yingshu Li and Ker-I Ko, A Greedy Approximation for Minimum Connected Dominating Sets, to appear in Theoretical Computer Science, 2005.

[22] A. Salhieh, J. Weinmann, M. Kochha, and L. Schwiebert, Power Efficient topologies for wireless sensor networks, *ICPP'2001*, pp. 156-163.

[23] R. Sivakumar, B. Das, and V. Bharghavan, An Improved Spine-based Infrastructure for Routing in Ad Hoc Networks, IEEE Symposium on Computers and Communications, Athens, Greece, June 1998.

[24] P.-J. Wan, K.M. Alzoubi, O. Frieder, Distributed construction of connected dominating set in wireless ad hoc networks, *Proceedings of Infocom 2002*.

[25] P.-J. Wan, K.M. Alzoubi and O. Frieder, Distributed Construction of Connected Dominating Sets in Wireless Ad Hoc Networks, IEEE INFOCOM, 2002.

[26] Ding-Zhu Du, Lusheng Wang, Baogang Xu, The Euclidean Bottleneck Steiner Tree and Steiner Tree with Minimum Number of Steiner Points, *COCOON' 2001*, pp. 509-518, 2001.

[27] J. Wu and H. Li, On Calculating Connected Dominating Set for Efficient Routing in Ad Hoc Wireless Networks, Proc. of the Third International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications, Page 7-14, August 1999.

[28] Weili Wu, Hongwei Du, Xiaohua Jia, Yingshu Li, Chih-Hao Huang, Ding-Zhu Du, minimum connected dominating sets and maximal independent sets in unit disk graphs, technical report 04-047, Department of Computer Science and Engineering, University of Minnesota, 2004.