Optimal Algorithm for Connected Dominating Sets

Nannan Zhao and Muqing Wu

Beijing University of Posts and Telecommunications,
Xitucheng str. 10, 100876 Beijing, China
znn812@gmail.com,
wumuqing@bupt.edu.cn

Abstract. There is no infrastructure in most energy constrained networks, such as WSN. Connected Dominating Set (CDS) has been proposed as virtual backbone. The CDS pays the way for routing, data aggregation and activity scheduling. In order to reduce the backbone size and prolong the lifetime of networks, it is desirable to construct a Minimum CDS (MCDS). Unfortunately, it is a NP-hard problem with a distribute manner. In this paper, a distributive algorithm for MESH CDS is introduced. Theoretical analysis and simulation results are also presented to verify the efficiency of our algorithm.

Keywords: WSN, Connected Dominating Sets, MESH CDS, Localized distributive algorithm.

1 Introduction

Wireless Sensor Networks (WSN) has attached more and more attention recently [1] [2] [3]. It revolutionizes information gathering and processing in both urban environments and inhospitable terrain. The nodes in WSN communicate with each other through multi-hop without physical infrastructure. Therefore a Connected Dominating Sets has been adopted in order to construct a virtual backbone.

The topology related problem in wireless sensor network usually has been modeled in Unit Disk Graph (UDG) [4], in which each node has the same transmission rage and there is an edge between two nodes if and only if their distance is within the transmission range. A Dominating Set (DS) is a subset of nodes so that each node is the UDG graph is either in it or adjacent to at least one node in it. If the DS is connected, it is a Connected Dominating Set (CDS).

The CDS plays a major role in routing, broadcasting, coverage and activity scheduling. To reduce the communication overhead, to simplify the network management, and to make more nodes stay in radio-sleep state which prolong the network life time, it is desirable to find a Minimum Connected Dominating Set (MCDS). Unfortunately, computing a MCDS for a given UDG has been proved to be NP-hard. Thus, only distributed approximation algorithms in polynomial time are practical for wireless sensor networks.

There are many distributed CDS algorithms which use a Maximal Independent Set (MIS) to build a Dominating Set (DS) [5] [6] [7]. An Independent Set (IS) is a subset of nodes which there is no edges between any two nodes in the subset. An MIS of a

P. Sénac, M. Ott, and A. Seneviratne (Eds.): ICWCA 2011, LNICST 72, pp. 340–350, 2012.

UDG is an IS that any other node is adjacent to some node in this IS. Later, the DS will be connected according to different criterion [8] [9].

In this paper, we proposed an algorithm to construct an Optimal CDS. It is a MESH CDS is size-and-energy optimized which has less number of nodes in the backbone we built, since the MIS node of CDS in qualified by degree. The energy cost of construction is lower than the previous CDS algorithms, because we do not use ACK messages in the CDS construction step. We choose MESH structure instead of tree structure, which make the virtual backbone have better robustness against nodes failure.

The remainder of the paper is organized as follows:

In section 2, we provide related research works on CDS and MCDS. Our MESH CDS algorithm is proposed in section 3, which also contains the performance analysis. The simulation results using OMNeT++ will be shown in section 4. At last, in section5, we conclude the paper.

2 Related Works

In WSN, most nodes have no need for global network information. Since MCDS problem is NP-hard in distributive manner, there are lots of approximation algorithms have been proposed. These approaches can be classified into two types. One type is to form a DS at first, for example, using MIS to construct a DS [8] [9]. At the second step, they try to find some connectors according to certain optimal principle, so that the DS in transformed to a CDS. Wan et al. proposed a distributed algorithm based on quasi-global information (Spanning Tree) with approximation ratio of 8, v_i time complexity and $O(n \log n)$ message complexity. Another type is to find a CDS at the initial stage and then prune some redundant nodes or links to archive final CDS. Wu et al. proposed in [10], a distributed algorithm with message complexity, $\theta(m)$ time complexity and $O(\Delta^3)$ the approximation ratio at most $O(n)$, where Δ and m are the maximum degree in graph and number of edges respectively.

As mentioned in [11], each of these two types has their own advantages and drawbacks. CDS built by the first type of algorithms have smaller size which means there would be fewer nodes whose radio state must be active. It seems that the energy efficiency of network is surely better than the CDS of the other type, however, we could not make this conclusion simply. The reason is that the latter type of algorithms has much lower energy cost due to its lower complexity. They do not have to send so many messages to achieve the CDS. What was worse, the network lifetime will be shorter using first type of CDS, because when the nodes in the CDS is fewer, the robustness will be weaker, so there will be much more times of CDS reorganization considering the nodes failure. This problem has been studied in [12]. The Second type of algorithms usually create a CDS which is obviously larger compared with the CDS produced by the first type of algorithm. In this situation, there must be a great number of nodes which could not close their communication module to save energy. As a result, the energy of nodes will not be used efficiently, since there will be a long time when most nodes in the CDS have no messages to send and receive but these nodes still wait to do that, spending lots of energy on communication module.

At last the time complexity and message complexity is also important. None of solutions mentioned above has $O(1)$ time complexity and $O(n)$ message complexity.

3 Our Proposal

3.1 Assumption

First, we assume that all the nodes in WSN are deployed in the 2-dimentional plane. Second, all the nodes have the same transmission range. So the network topology is modeled as a UDG. Each node v_i has a unique id ID_i. The number of links which is connected to v_i is represented as D_i. The maximum timeout for each node is noted $T_{\max}$. For each node v_i, the timeout is T_i set by:

$$T_i = \frac{T_{\max}}{D_i}.$$

So the node which has larger D_i has faster timeout.

We use colors to indicate the nodes' states. Each node has one of the four colors which represent its state: White, Black, Grey, Green and Blue. White color stands for the initial state which indicates that the node has not decided its role in MIS and joined the CDS neither. Black (Grey) color means the node has joined (or not) the MIS, but has not joined the CDS. The Blue node is the CDS node. At last, Green color represent a temporary state which tells that the node has already started the MIS/CDS processing but still not decided its final role .

3.2 Initial Process

Initially, each node is in White which means its role is not decided yet and all nodes' degree D_i is initialized as 0. At first, each node broadcasts its ID_i. When the node received the broadcast, it adds the ID_i to its neighbor table and D_i is increased by 1.

Algorithm 1: Initial process

1. Each node v_i set its color to white, set D_i to 0.
2. Each node broadcast message including its own id ID_i.
3. When received the id broadcast message, it adds the received ID_i to its neighbor table and increase the D_i by 1.

3.3 MIS Construction

There are many MIS calculation proposals, such as [5] [6] [10] [13]. In fact, almost all the MIS protocols have the common algorithm:

- Initially, each node is in White which means its role is not decided yet;
- Each node broadcasts its weight (ID, degree or level in a tree);
- Each node collects neighbors' information and sorts the neighbors according to the weight;
- If a node has the highest weight among its neighbors, the node declares it is IN the MIS and set its color to Black;
- When a node receives the declaration of a Black node, it declares it is OUT the MIS and set its color to Grey;
- If the node with lower weight did not receive any Black declarations, it should wait the decisions of all the higher weight nodes until it becomes the highest weight node among the nodes whose roles are still not decided;
- The neighbor table will be kept updating during this period according to the declaration messages;
- When all the nodes are either Black or Grey, the algorithm is terminated.

The scheme quite simple and effective to make a MIS, however, this kind of algorithms will bring unnecessary time delay. The reason is that some node which is not the largest-ID node within its neighbors must wait for its neighbors' messages to decide its own color. There is no constrains for this time delay, what was worse, the neighbors messages may not be received successfully every time in fact. As a result, there may be some nodes waiting perpetually. In another word, these nodes will not be colored and stay white. Finally, it causes the failure of MIS construction because these white nodes are not covered by any nodes in the independent set.

So our algorithms give some modifications.

- We choose ID to be the measurement of weight;
- When a White node receives the declaration of a Black node, it set its color to Grey and broadcast a Grey declaration to inform its neighbors;
- When a White node receives a declaration of a Grey node, it set its color to Green, decrease D_i by 1 and set timeout to $T_i = \frac{T_{\max}}{D_i}$. This means the color of node need to be decided through competition;
- When a Green node receives a declaration of a Black node, it means the competition fails. Then the node set its color to Grey and broadcast a Grey declaration to inform its neighbors;
- When a Green node's timeout expires, it means the node win the competition, it set its color to Black and broadcast a Black declaration to inform its neighbors;
- When a Black node receives a declaration of a Black node which has larger node ID, it means there is a neighbor which has higher priority to be MIS node than itself. Then the node set its color to Grey and broadcast a Grey declaration to inform its neighbors;
- Finally, all the Black nodes form the MIS.

Algorithm 2: MIS construction

1. Each node lookup its own neighbor table. If a node v_i has the largest ID among its neighbors, it set its own color to Black and broadcast a message which indicates its own color and ID. We note this Black Message for short. Similarly, there will be Grey Message later.
2. When a White node received a Black Message, it set its own color to Grey and broadcast a Grey Message.
3. When a White node received a Grey Message, it set its color to Green, decrease D_i by 1 and set a timer to $T_i = \frac{T_{max}}{D_i}$.
4. When a Green node received a Black Message, the node set its color to Grey and broadcast a Grey Message.
5. When a Green or Grey node received a Grey Message, decrease D_i by 1.
6. When a Green node's timer expires, the Green node set its color to Black and broadcast a Black Message.
7. When a Black node receives a declaration of a Black node which has larger node ID, the node set its color to Grey and broadcast a Grey declaration to inform its neighbors.

Theorem 1: The set of black nodes computed the MIS construction algorithm is a MIS.

Proof: First, all the black nodes forms a Independent Set (IS). The reason is that any two black nodes are not neighbors. If they are, one of them must be colored grey according to the modification next to last.

Second, we say there is not any node which is independent form the black nodes set. For the contradiction, we assume there is one node v_i which neither belongs to black nodes set nor covered by any black nodes. So node v_i will never receive any declarations of Black node, and then it must be colored into Black. So node v_i is a member of black nodes set and contradiction is made.

So we proved our black nodes set is a MIS.#

Theorem 2: If the size of MIS is no less than 2, for each MIS node calculated, it always has a non-MIS neighbor that connects it to at least another MIS node.

Proof: In order to make contradiction, we assume that there is a MIS node u which has not any neighbors that connect u to another MIS node.

We note all the neighbors of u forms a set N_u^1. All the neighbors of nodes in N_u^1 forms a set N_u^2. Because there is not any node which connects u and another MIS node, N_u^2 must be constitute of non-MIS node. Referring to the Algorithm2, there will not be another MIS node. This make contradiction to the precondition that the size of MIS is no less than 2.#

Apparently, after the MIS construction, the D_i of Grey nodes is the number of Black nodes it connects with.

3.4 CDS Formation

The second step is making a CDS. Since an MIS is also a DS, a CDS can be constructed by connecting the nodes in an MIS with some nodes not in the MIS which we call connectors.

In the second step, a tree is usually formed in most of the backbone formation proposals. The tree is formed from one selected node and then the formation process spreads over the network until all Black nodes are in the tree backbone. Tree formation is a sequential process. In a WSN of large scale, the failure and adding of nodes may take place all the time. As a result, the tree backbone has to be reorganized frequently. So compared with tree backbone, we consider that MESH backbone which has additional links is more applicable to WSN. Many algorithms use ack messages to make the backbone constructed successfully and collect the parents-and-children information, however, we think it is not necessary. Topology control only make decisions that which node/link exists in the network. The parents-and-children information is used for routing which is the successive problem of topology control. So we only need to form a CDS as well as determine which nodes belong to CDS. Whatever their parents and children should be, it is not important. Our proposal only gives out the nodes in CDS without parents-and-children information.

The CDS formation is started from the initiator which is assigned at first. The initiator set its color to Blue and then broadcast a Blue message. When a Grey node receives the Blue message, it changes its own color to Green and then enters a waiting period which has Ti seconds. During this period, Green node does not handle any Blue messages. Once the green node timeout, it set its color to Blue and broadcasts a CLR message and a Blue message. When a green node receives a CLR message, it timeout immediately and set its color back to Grey. When a BLACK node receives the Blue message, it set its color to Blue and then broadcast a Blue message. When a node turns into Blue, it discards any messages it receives.

Algorithm 3: CDS construction

1. One Black node which was assigned to be initiator set its own color to Blue and broadcast a Blue Message.
2. When a Black node received a Blue Message, it set its own color to Blue and broadcast a Blue Message.
3. When a Grey node received a Blue Message, it set its own color to Green and set a timer to $T_i = \frac{T_{\max}}{D_i}$. Then the Green node broadcast a Blue message.
4. Each Green node does not handle any Blue messages.
5. Once a Green node timeout, it broadcasts a CLR message
6. When a Green node receives a CLR message, its timer is cancelled immediately and set its color back to Grey.
7. When a Blue node received a message, the message will be ignored.

Theorem 3: The set of Blue nodes computed by the algorithm 3 is a CDS the network.

Proof: First we assume there is only one MIS node u which is not blue. According to theorem 2, there is at least one non-MIS node which connect to another MIS node v. Then v must be blue and has sent a Blue message. So the connector of u and v must have send a Blue message, which indicates that the node u must be blue. There is contradiction. If there is other non-Blue MIS node we can make the contradiction similarly one by one. Now we can say that all the MIS node must be Blue.

Second we assume there is one Blue MIS node u which is not connected to another part of MIS node. Then node u has not received any Blue message, so it would not be a Blue node. There is contradiction.

Finally, we can say that all the Blue node forms a CDS of the network in conclusion.

4 Simulation Resluts

In this section, we verify our algorithm by test its performance with the simulation software OMNeT++ in different network size. We have made the comparison between our proposal , K. M. Alzoubi et. al.'s algorithm in [5] and a MESH CDS algorithm in [14]. We will show the result in 2 aspects, that is CDS size and energy cost.

The number of nodes is from 10 to 500. Node density is $1/10000\, node/m^2$. Nodes are distributed randomly. Mobility model is set to be static. CSMA MAC is used. The energy model parameters is shown in TABLE 1.

Table 1. Energy model parameters

Radio state	Energy cost (mW)
Tx	78
Rx	78
IDEL	0
SLEEP	0

The CDS of algorithm in [5] is shown if Fig. 1 while the result of our proposal is shown in Fig. 2. All of them have the same network size, 50 nodes. The black node is the MIS nodes as well as the cyan nodes is the connectors. The Gray node stands for the non-CDS nodes. The dash line in Fig.1 show us the parent-child links of nodes'. We can see that the number of MIS nodes in Fig. 2 is than that in Fig. 1. Furthermore, the MIS nodes emerge at the area where nodes are distributed more densely. In another word, one MIS node computed by our algorithms covers more nodes. The reason is that when the nodes with larger degree have less time to wait in algorithm 2, which means they have larger probability to be selected as MIS node.

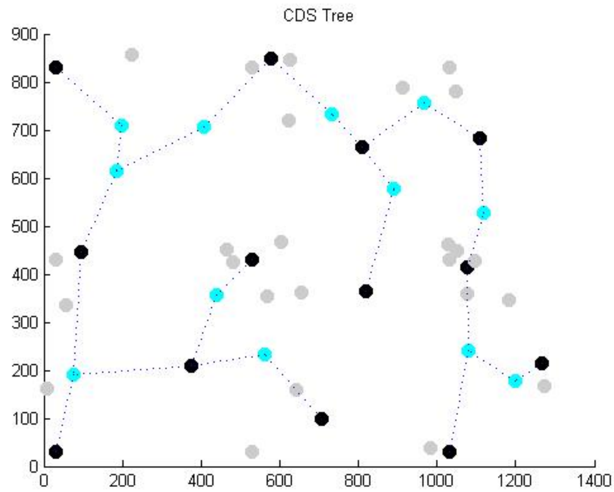

Fig. 1. The backbone of CDS Tree

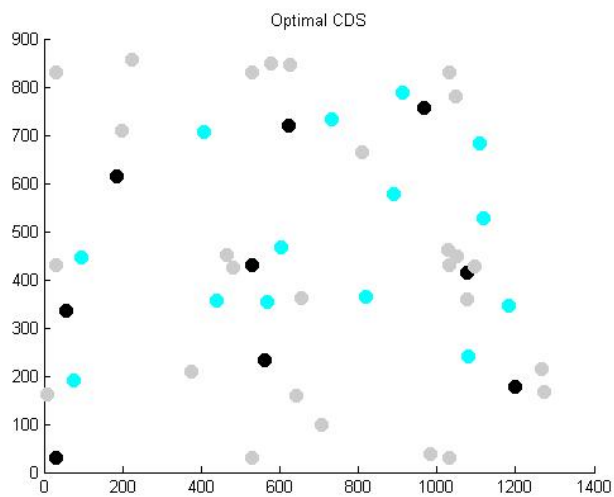

Fig. 2. The size of Optimal CDS

The size of MIS in different network scale is shown in Fig. 3. The number of nodes in MIS computed by the Optimal CDS is less than those computed by the other two algorithms. As we explained before, MIS nodes in our algorithm emerges at where the nodes are deployed more densely. While the network's scale increasing, the difference is growing larger. The fewer MIS nodes lead to the fewer connectors, so we can infer that the size of CDS constructed by our algorithm will be smaller than the MESH CDS in [14]. The results of CDS size is shown in Fig. 4. The CDS Tree in [5] has the least number of CDS, because the tree structure has less number of links.

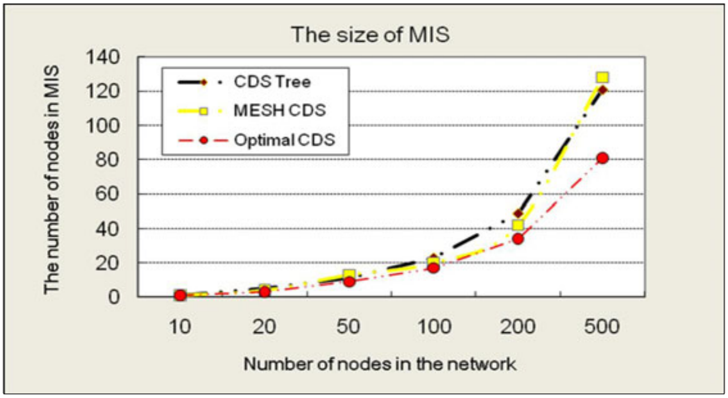

Fig. 3. The size of MIS

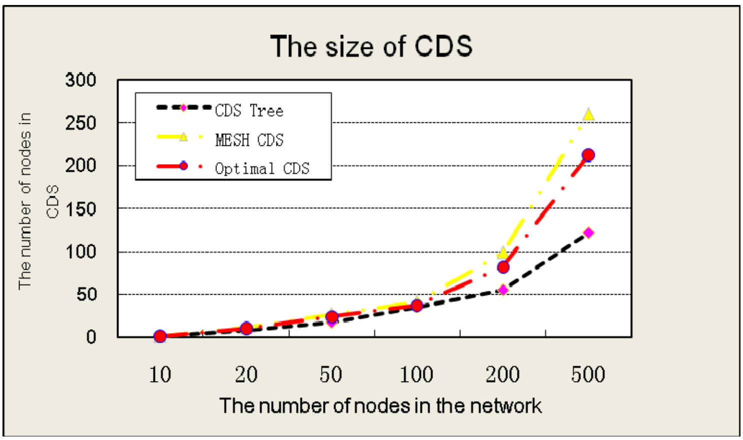

Fig. 4. The size of CDS

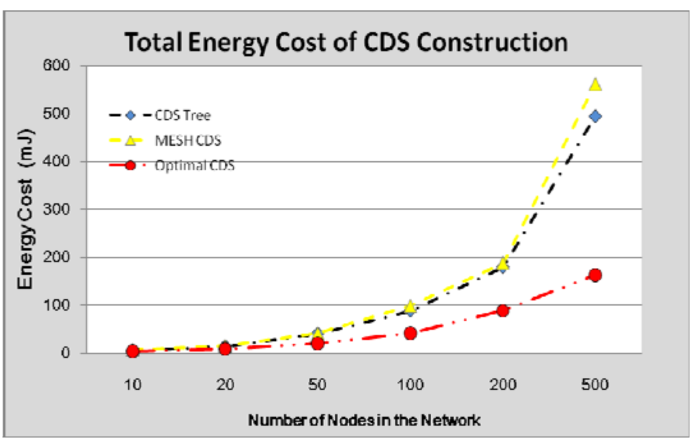

Fig. 5. The energy cost of CDS construction

The energy cost of CDS construction is shown in Fig. 5. It is clear to see that Optimal CDS saves a lot of energy than the other 2 algorithm. The reason is that there is no ACK message in our proposal, which reduces a great amount of messages in the CDS forming step. Though Optimal CDS has more connectors than those of CDS Tree, its energy cost is the least among the three.

5 Conclusion

In a random deployed network Connected Dominating Set plays an important role such as routing, broadcasting, coverage and activity scheduling in WSN. Our Optimal CDS is a CDS uses a time compete mechanism with the nodes' degree to make each MIS node covers other non-MIS nodes as many as possible. In order to enhance the energy efficiency of each node, we also removed the ACK messages and parent-child information which is unnecessary for topology control. The simulation results show us that the Optimal CDS has greatly reduced the number of MIS nodes and the energy cost of CDS construction.

References

1. Pottie, G.J.: Wireless sensor networks. In: Information Theory Workshop, June 22-26, pp. 139–140 (1998)
2. Sohrabi, K., Gao, J., Ailawadhi, V., Pottie, G.J.: Protocols for self-organization of a wireless sensor network. IEEE Personal Communications 7(5), 16–27 (2000)
3. Ahmed, A.A., Shi, H., Shang, Y.: A survey on network protocols for wireless sensor networks. In: Proceedings. International Conference on Information Technology: Research and Education, ITRE2003, August 11-13, pp. 301–305 (2003)
4. Clark, B.N., Colbourn, C.J., Johnson, D.S.: Unit disk graphs. Discrete Mathematics 86, 165–177 (1990)
5. Alzoubi, K.M., Wan, P.-J., Frieder, O.: Distributed Heuristics for Connected Dominating Sets in Wireless Ad Hoc Networks. Journal of Communications And Networks 4(1) (2002)
6. Alzoubi, K.M., Wan, P.-J., Frieder, O.: Weakly-connected dominating sets and sparse spanners in wireless ad hoc networks. In: Proceedings. 23rd International Conference on Distributed Computing Systems (2003)
7. Raei, H., Sarram, M., Adibniya, F., Tashtarian, F.: Optimal distributed algorithm for minimum connected dominating sets in Wireless Sensor Networks. In: 5th IEEE International Conference on Mobile Ad Hoc and Sensor Systems, MASS 2008 (2008)
8. Wan, P.-J., Alzoubi, K.M., Frieder, O.: Distributed Construction of Connected Dominating Set in Wireless Ad Hoc Networks. Mobile Networks and Applications 9, 141–149 (2004)
9. Wang, F., Thai, M.T., Du, D.-Z.: On the construction of 2-connected virtual backbone in wireless networks. Wireless Communications (March 2009)
10. 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 (1999)

11. Basagni, S., Mastrogiovanni, M., Petrioli, C.: A performance comparison of protocols for clustering and backbone formation in large scale ad hoc networks. In: IEEE International Conference Mobile Ad-hoc and Sensor Systems (2004)
12. Zhao, N., Zhang, Y.: The Relation between Organized Structure and Maintaining Cost in WSN. In: 5th International Conference on Wireless Communications, Networking and Mobile Computing, WiCom 2009, September 24-26, pp. 1–4 (2009)
13. Basagni, S., Petrioli, C., Petroccia, R.: Efficiently reconfigureureurable backbones for wireless sensor networks. Computer Communications 31(4), 668–698 (2008)
14. Zhao, N., Wu, M., Zhang, Y.: A Localized MESH-CDS Structure and Maintenance for Wireless Sensor Networks. In: 2010 6th International Conference on Wireless and Mobile Communications (ICWMC), September 20-25, pp. 262–267 (2010)