An energy-aware distributed algorithm for virtual backbone in wireless sensor network with different transmission range

Mohammad Amin Ebrahimzadeh Fakhar¹*, Seyed Reza Kamel Tabbakh²

¹ Department of Software Engineering, Khorasan Razavi, Neyshabur, Science and Research Branch-Islamic Azad University, Neyshabur, Iran, Amin_efakhar@yahoo.com

² Department of Software Engineering, Mashhad Branch - Islamic Azad University, Mashhad, Iran, DrKamel@mshdiau.ac.ir

Abstract

Since there is no fixed infrastructure or centralized management in Wireless Sensor Networks (WSNs), a Connected Dominating Set (CDS) has been proposed as a virtual backbone is efficient. A virtual backbone plays a major role in routing, broadcasting, coverage and activity scheduling. Wireless sensor networks to form a CDS usually by UDG (Unit Disk Graph) models that are used in this model, all nodes have the same message, but this article UDG model instead of a version that is closer to reality called DGB (Disk Graph with Bidirectional links) is used in which nodes can adopt different transmission intervals. In many applications, to reduce overhead, increase network lifetime, and so on, to find the MCDS (minimum connected dominating set) is desirable, but the point is that MCDS UDG models and DGB, the problem is NP-hard. In addition to the analysis of algorithms, the new algorithm will provide and the efficiency of the algorithm, especially in terms of energy consumption, through theoretical analysis and simulation algorithms are available to be checked out.

Keywords: wireless sensor networks, distributed, energy-aware, virtual backbone, different transmission range.
1 Introduction

Wireless sensor networks are a group of special networks of hundreds or even thousands of low costing, low power and multipurpose sensors be made that can sense their surroundings and to collect and process the relevant information and send this information via radio waves for nodes that are in transmission range[1].

Usually one or more nodes in a wireless sensor network are important to sink the intelligence gathered by the sensors, either directly or through intermediate nodes send to this node.

Wireless sensor networks usually by UDG (Unit Disk Graph) are model [2]. In this model, all nodes have the same transmission range and two other neighboring nodes within range if they are transmission range to each other. Since in practice the assumption that all nodes have the same transmission range due to the weakening of transmitter-receiver and many other factors is impossible, and in many applications require different transmission range are, in this paper, a model of the DGB (Disk Graph with Bidirectional links) we use [3] where nodes can adopt a different transmission range.

For the most efficient operations in a multi-hop wireless sensor network including communications and environmental monitoring, the number of sensors to form a CDS (connected dominating set) as the virtual backbone of a viable and desirable [4]. A subset of nodes, DS (dominating set) called if any network node, or a subset of the above, or in the neighborhood is one of the subsidiaries. A DS, connective (CDS) if the sub graph obtained from the DS, is connected. CDS is a virtual backbone on the network graph.

In many applications, to reduce traffic congestion, reduce overhead and increase network lifetime, and so on, to find MCDS is desirable, but the point is that a problem is NP-Hard is MCDS UDG model [2] and this DGB model is included as well [5]. So to solve this NP-Hard problem different solutions, including Centralized [6], based on cluster (Cluster-based) [7, 8], Greedy [9] and Heuristic [10] is offer.

Some of the methods for the calculation of a CDS, an MIS will first calculate the network graph. An independent set (IS) a non-directed graph G (V, E); a subset of V is between any of its nodes, not the edge. An independent set is maximal if the other node can not be added to this collection. This means that by adding new nodes, edge nodes between sub-subsidiaries is established and independent properties are violated. Therefore, a MIS is a DS of a non-directional graph. Black nodes in Figure 1 show an MIS.

![MIS](image)

**Fig 1: maximal independent set (MIS)**

Approximation ratio is ratio of the optimum an algorithm that shows. This ratio is defined as the relation 1:

\[ AR = \frac{D_{\text{Alg}}}{D_{\text{Opt}}} \] (1)
In this relation, $D_{\text{Alg}}$ is the presented CDS size (number of nodes selected as CDS in solution provider in the worst case) and $D_{\text{Opt}}$ is optimal CDS size (number of optimal (lowest) CDS nodes). In various applications, the constant approximation ratio, it is desirable that the algorithm presented by this paper has this feature.

The algorithm presented in this paper has the following characteristics:

* Distributed: this feature for wireless sensor network is very important because it requires no central control and each sensor can implement this algorithm.

* Localized: This means that the sensor doesn’t need to know the full network topology or general information and only it uses information of its range.

* Proposed algorithm has approximation ratio is better than existing algorithms. Approximation ratio, the CDS size (number of nodes) presented to optimal CDS size that proposed algorithm over existing algorithms provide the smaller CDS.

* Time complexity and message complexity are better than existing algorithms in the field.

* As an important factor in terms of energy consumption, compared to existing algorithms, is optimized. (Energy Efficient)

2 Previous works

In general, the methods of construction of a CDS divide into two categories, centralized and distributed which the distributed methods according to the nature of wireless networks is taken into consideration.

Because wireless sensor networks without centralized management and mass usually developed in a large space that is much more effective and efficient methods of distribution focused approach, because in this section we will investigate these methods.

Das and colleagues [11] an algorithm based on the greedy algorithm [12] provided that this algorithm Methods is: in the first phase a dominating set (DS) is calculated and then add nodes to the set to obtain CDS. In the first phase of each node by node is selected so that each node degree with the neighbors those are 2 jumps out of it, and if degree was higher comparing to be elected as a member of CDS is noteworthy that the degree in this algorithm in fact, is effective degree, that number of non-DS neighboring nodes. When a DS achieved the first phase is finished. In the second phase components obtained by a spanning tree are connected to each other with the aim of selecting the lowest number of nodes. This algorithm has $O(cn + |c|\Delta)$ time complexity and $O(n|c| + m + n\log(n))$ message complexity, which $|c|$ refers to the number of nodes selected as CDS, $\Delta$ the maximum degree of the graph and $m$ is the number of edges in the graph.

Wan et al. [13] offer an algorithm based on the MIS to find CDS in UDG that was made up of three phases. In the first phase, the algorithm is employed distributed leader election method [14] to obtain a spanning tree of network graph. The concept of rank in the algorithm refers to level of node in the tree made out that root with the lowest rating level is zero. The second phase starts at the root and terminates at leaf, the node has lowest rank (level node in the tree made) its own black and broadcast a DOMINATOR message then algorithm continues according to the following rule:

- If the first message that a node receives a DOMINATOR message, the node itself is gray, and broadcast a DOMINATEE message.

- If a node received DOMINATEE message of all neighbors with lower ratings, itself is black and broadcast a DOMINATOR message.
This algorithm has a time complexity of \( O(n) \) and message complexity of \( O(n \log n) \) and approximate ratio of \( 8 \, \text{opt} + 1 \).

Thai and colleagues [3] have provided based on MIS an algorithm for finding CDS in DGB. This article is the first article that specifically deals with the DGB. In the first phase for constructing of a MIS, node selected by the highest transmission range and in the second phase using a Steiner tree connecting the nodes to be added to the collection.

In the first phase, all nodes are white. Each node is a sorted list of identifiers (id) all nodes in order to transmission range named Sortlist. Initially, this list is only their IDs node. Each node \( V_i \) broadcast a WHITE message that the message contains id and transmission range of node <id, r/>. After receiving a white message each node update his regular list that it will be adding identifier the message WHITE that according to the transmission range of node to list sorted done. A node that has its own identity on the front (first id) ordered list sees himself black and broadcast a BLACK message because there is ID list is arranged in the sense of having the highest Brad write the neighbors. BLACK message, including node identifier is black. After receiving a BLACK message, WHITE node itself becomes gray and GREY message including node ID and ID BLACK is the message plays after receiving GREY, white node has its own sort lists by deleting the node ID GREY message is sent, arise.

The second phase requires only a Steiner tree are for joining the MIS. All nodes in the MIS are black and the other nodes are gray. Each node has variable is called IDc, that holds black-blue component it belongs. Suppose I have a MIS formed in the first phase. First have \( |I| \) black-blue component (the number of MIS nodes), so first we set IDc black node to node its identifier. Each gray node initially IDc = -1 that show it is not owned by any one of the black blue component. Grey neighboring nodes and nodes that are dark gray-blue are common in the neighborhood; competitors are, so each node gray has list of competitors and a list of black-blue components neighbors (ADJ List).

The first phase of time and message complexity is \( O(n^2) \), which has belonged to the sort list. The second phase is linear time complexity and message complexity is \( O(n \log n) \). Therefore the time complexity and message of this algorithm is \( O(n^2) \). They also showed that the algorithm MCDS approximation of the DGB is fixed (Constant or \( O(1) \)).

The Butenko algorithm [15] First, all nodes are CDS members, the first node with the lowest degree with the help of an algorithm leader will be chosen. If u remove the cause CDS to be discreet, that is divided into two parts, the node u in the set remains black after that node u One of the neighbors with the lowest degree (number of neighbors none black) and law above to remove it, try it and if it was removed from the CDS, and this process will continue until the end. This algorithm has time complexity of \( O(nm) \). A further analysis of this algorithm is not provided.

With CDS, routing easier and adoption of topology changes is faster [16]. CDS play a decisive role in reducing overhead and increase efficiency in its broadcasting [17]. In some ways, such as [18] for the calculation of a CDS, the MIS of a network graph calculate that the proposed algorithm is also used this method.

3 The proposed algorithm

We assume that all the sensors are distributed in a two-dimensional surface. DGB is the model used in the algorithm means that all nodes have a different transmission range. The graph \( G(V, E) \) we use to show network that \( V \) set of nodes and \( E \) set of edges. The proposed algorithm uses a combined weight of the power and transmission range to each node and then by a timer each node with the highest specific
weight and as a member of the CDS chooses. The proposed distributed algorithm consists of two phases. The first phase constructing MIS and second phase connecting MIS nodes by connected nodes and is composed of a CDS.

In this algorithm, each node has a unique identifier ($ID_i$), state ($S_i$), the remaining energy ($E_i$), and transmission range ($R_i$). The algorithm uses combined weight ($W_i$) for each node including energy and transmission range that on the basis of Equation 2 for each node is set up in addition to efficient algorithms in terms of energy, as much as possible to reduce the size of the CDS.

\begin{equation}
W_i = \frac{R_i}{E_i} + E_i
\end{equation}

In every phase select nodes with the highest weight ($W_i$) among neighbors using the timer ($\Delta t_i$). Each node $i$ has a timer ($\Delta t_i$) that is set according to equation (3).

\begin{equation}
\Delta T_i = \frac{1}{W_i} \times T_{MAX}
\end{equation}

$T_{MAX}$ is the maximum time possible for each timer. So each node with the highest $W_i$ ends faster than its neighbors.

3-1. First phase: construction of maximal independent set (MIS)

All nodes are initially white; each node can be in one of the states of White, Black, Grey and Green and can receive BLACK, GREY AND INQUIRY messages. At the end of the first phase nodes are black or gray those black nodes form MIS. Green state is an intermediate state while node in this mode uses a timer to be sent information by an INQUIRY message its neighbor transmission range and at the end of timer if the node still be green and in neighbors green have highest transmission range, black itself and broadcast BLACK message. BLACK message is sent when a node black and GREY message is sent when a node is gray and INQUIRY message is sent when a node is going to be aware from transmission range of other nodes based on the responses received and aware from transmission range of the rest.

At the start of all the nodes are white, the initiator node is itself black. A node that is Black immediately send BLACK message, including its identifier to indicate that a node is MIS. Each White node that receives a BLACK message, gray itself, and broadcast a GREY message that includes the identifier (ID) in the message itself and identity BLACK to show the neighbors that a node is non-MIS. White node receives a GREY message neighboring gray node after can subscribe to the MIS match, so this node is their green and broadcast a INQUIRY message and also set a timer and wait for the response message and remains in this state until the timer is not all. White nodes can also get an INQUIRY message that contains a message identifier and its transmission range that sends to sender of message.

During timer, a green node receives BLACK, GREY AND INQUIRY messages. If a node receives a BLACK message, gray itself and a GREY message sends containing their ID and identity the BLACK message. If receive a message GREY ignores it. If receive an INQUIRY message was sent an answer contains the ID and its transmission range to the sender. The algorithm is shown in Pseudo code 1. First the trigger function executes.

```python
Initiator () {
  Color itself in black and broadcast a BLACK message contained $S_i$ and $ID_i$.
}
```
1. Each white node $v_i$ maintains a transmission range $R_i$, a state $S_i$ and a unique id $ID_i$.

2. Upon receiving a BLACK message, a white node marks itself grey and broadcasts the GREY Message contained $S_i$, $ID_i$ and id in the BLACK message.

3. Upon receiving a GREY message, a white node mark itself green and set a timer and broadcast an INQUIRY message.

4. Upon receiving an INQUIRY message, a white node reply a message contained $R_i$, $S_i$ and $ID_i$.

5. Upon receiving a BLACK message, a green node mark itself grey and terminate the timer and broadcast a GREY message contained $S_i$, $ID_i$ and id in the BLACK message.

6. Upon receiving a GREY message, a green node, ignore the message.

7. Upon receiving an INQUIRY message, a green node reply a message contained $R_i$, $S_i$ and $ID_i$.

8. Ignore the all messages if node is grey or black

Pseudo code 1. The first phase of the proposed algorithm: constructing maximal independent set (MIS)

When the timer expires if a Green node based on the responses sent among all its green neighbors has the highest transmission range, black itself and broadcast a BLACK message including its ID, otherwise the node white itself. This process continues until each node is black or gray. Node in Black and gray states ignore all messages received. After the timer end, every green node runs following function.

Timer-green () {If $R_i$ is largest among green neighbors then

Node marks itself black and broadcast BLACK message contained $S_i$ and $ID_i$.

Else

Mark itself white. }

Theorem 1: The set of black nodes are calculated by the first phase forms a MIS on a network graph.

Proof: black nodes denoted by B. Pseudo code (1) colors the nodes of the graph layer by layer that begin from initiator, a black layer and the next layer of gray, so continue until all the nodes are black or gray. Except initiator, black nodes in each layer select by gray nodes previous layer. So the new black node and nodes that are already black, and exist in B 2-hop distance. So for a middle layer of gray nodes, black nodes can't adjacent to each other and therefore B is independent set (IS). The algorithm ends when the nodes are black or gray, then gray node must have at least one node in the neighborhood is dark gray to black, so if any node to node attributes Independence loses his then B is a MIS.

3.2. Phase II: CDS construction

For the second phase should MIS nodes connected to each other and we have to do is use the minimum number of connecting nodes according to the performance Steiner
tree [3] of this structure is used. Paying attention to Pseudo code 2:

After the implementation of the first phase, all MIS nodes are black and belong to I collection and all other nodes are gray. In this algorithm, each node has a variable $ID_c$ that for nodes Black represents ID of Black-Blue Component that it node belongs. Each black node initializes it with the value of its argument. First for the gray nodes $ID_c = -1$ that indicate the node belongs to no one black-blue Component. Each node is also a list of black-blue components along with $ID_c$ their neighboring ADJ called, keeps a list of competitors also called COMPETITORS. Each gray neighboring node is a black-blue component if the node where black node in the black-blue Component. A gray node $u$ is competitor's $v$ gray node if any two neighboring are common in black-blue components. If the number of black-blue component adjacent any two nodes are the same at this moment, the node with smaller ID is blue. The list COMPETITORS keep competing nodes.

1. $v_j$ Sends a BLACK message contained its $ID_c$.

2. Upon receiving the BLACK message, grey node $v_j$ updates its ADJ and COMPETITORS lists.

3. $v_j$ Sends a GREY message contained its id and its $|ADJ|$.

4. $v_j$ Turns blue if its $|ADJ| > |ADJ|$ of its neighbors and its $|ADJ| > 1$.

5. Each blue node updates it's $ID_c$ to the smallest value in its ADJ list.

6. A blue node then sends a BLUE message contained its new $ID_c$ and new ADJ list.

7. Upon receiving a BLUE message, black node $v_j$ updates it’s $ID_c$ and send a BLACK message.

8. If $|ADJ|$ of a grey node $v_j$ equal to 1, then do nothing.

Pseudo code 2. The second phase of the algorithm: CDS construction

3-3. Theoretical analysis

The following three important properties are listed in [3] for CDS in DGB

Lemma 1: $N_{MIS}(u)$ For a gray node (non-MIS) $u$, refers to the independent neighbors (MIS) in the DGB, where, and

$$ N_{MIS}(u) = \begin{cases} \frac{5}{\ln(\ln(10^{15}))} & \text{if } R = 1 \\ \frac{1}{\ln(2\cos(\pi/5))} & \text{otherwise} \end{cases} $$

$$ R = \frac{r_{MAX}}{r_{MIN}} $$

Lemma 2: In a DGB each node has a maximum of 5 nodes of MIS in the neighborhood.

Lemma 3: The DGB, the size of each MIS, the upper limit is $K_{opt}$ that opt is optimal MIS size.

Theorem 2: In the second phase of the algorithm, the number of blue nodes (connecting) is not more than $K_{opt}$ while $K$ is the same $N_{MIS}(u)$ as in Lemma 1.

Proof: put B for MIS (first phase) and S series blue nodes (connecting in the second
phase). From Lemma 3 we have $|B| \leq K_{opt}$ and the theory 1 and Lemma 2 that the number of MIS neighbors for a blue node is in the range of 2 to 5. Since Steiner tree involved in the second phase in a spanning tree for blue and black nodes, the worst case is when nodes are expanded inline and nodes are alternating black and gray (phase I) after the mode, all gray nodes are blue (the second phase), the number of blue nodes (connecting) can not be more than the number of black nodes (MIS) because other means we have $|S| \leq |B| - 1$ and according to Lemma 3 is $|B| \leq K_{opt}$ after $|S| \leq K_{opt} - 1$. Thus the numbers of blue nodes aren’t more than $K_{opt}$.

**Theorem 3:** The proposed algorithm has a Constant approximation ratio.

**Proof:** The algorithm has two phases; the first phase of MIS construction and the second phase is the construction of CDS. From Lemma 3 approximation ratio of the first phase (MIS) is $K_{opt}$ and from Theorem 2, approximation ratio the second phase is less than $K_{opt}$ then result the approximation ratio of algorithm is less than the 2 $K_{opt}$, thus algorithm has approximately constant approximation ratio.

**Theorem 4:** The proposed algorithm has time complexity and message complexity of $O(n)$.

**Proof:** In the first phase for BLACK, GREY AND INQUIRY messages, each node send a fixed number of these types of messages. Thus, message complexity of the first phase is $O(n)$. Beside, the complexity of the message for the second phase shown [3] is $O(n)$. Time complexity of the first phase in the worst case is $O(n)$ that is the worst when the nodes distributed base on the transmission range in a line, whether ascending or descending, so in this case time complexity is $O(n)$. But as for the second phase that showed [3], time complexity in the worst case is linear, the total of both phases, the proposed algorithm has a time complexity and message complexity of $O(n)$.

**4 Simulation results and their interpretation**

In this section the performance of the algorithm in terms of the size of the CDS, the impact of transmission range, the overhead message and the energy consumed for each node in the constructing of CDS on the network randomly investigated. Algorithm with existing algorithms for the CDS in the DGB [3](TWLZD) compared.

We use MATLAB 8 software for simulation. Network size is from 50 to 200 nodes (up every 25 nodes), which is randomly placed in an environment 100 × 100, to create a connected graph. Transmission range of nodes has been considered different within [10m, 30m] and the results for each chart are 20 times. Energy model presented in [19] we employ in this model to send the L bit in the distance d of the relation 4 and relation 5 is used to receive data.

$$E_{tx}(l, d) = E_{tx-elec}(l) + E_{tx-amp}(l, d) = \begin{cases} lE_{elec} + lgs d^2 & d < d_0 \\ lE_{elec} + l amp d^4 & d \geq d_0 \end{cases}$$

(4)

$$E_{Rx}(l, d) = E_{Rx-elec}(l) = lE_{elec}$$

(5)

We assume the package size broadcast, 250-bit, Update packet size used in TWLZD for the black and blue Component ID) 1000 bits, $T_{max} = 100$ ms and each node of primary energy is 1 Joule.

Figure 2 shows the size of the CDS, along with an increase in the number of nodes. A lower backbone, both in terms of reducing the network overhead and both in terms of increasing lifetime is effective. The algorithm has a better performance than the TWLZD algorithms.
Figure 3 displays the impact of Transmission range changes in CDS size. In this figure $R = \frac{r_{\text{MAX}}}{r_{\text{MIN}}}$ that $r_{\text{MIN}} = 10\text{m}$ and $r_{\text{MAX}}$ from 20 to 60 with an increase of 10 varies. A total of 100 nodes were developed in a 100x100 space. CDS produced by the algorithm, on average, 20.8 percent is smaller than TWLZD.

5 follow Figure 4. For example, $n = 125$, energy consumed by the algorithm is 5.1 times less than the TWLZD.

As a measure for the message complexity (message overhead) examine the number of bytes exchanged. As Figure 4 shows that the algorithm has better performance than is TWLZD. For example, $n = 125$, the number of bytes by algorithm, is 5.7 times less than TWLZD. Note that the proposed algorithm has a message complexity $O(n)$ while TWLZD is $O(n^2)$ complexity.

As a measure of efficiency in terms of energy, the energy consumed in the constructing of CDS measure’s each node. Since energy consumption is directly proportional to the message overhead, Figure 5 follow Figure 4. For example, $n = 125$, energy consumed by the algorithm is 5.1 times less than the TWLZD.

5 Conclusions and future work

In this paper, a distributed energy-aware method with a different transmission range for constructing virtual backbone in wireless sensor network in the DGB was presented that including two phases. In the first phase construction of MIS and the second phase is connected the nodes to MIS and making CDS. The theoretical analyses demonstrate that the proposed algorithm has the message complexity and the time complexity $o(n)$ and ratio is approximately constant. The simulation results showed that the proposed algorithm is more efficient in terms of energy consumption and caused crate a smaller backbone. Future work could be done on the optimization of constructing CDS.
References


