An Improved Greedy Construction of Minimum Connected Dominating Sets in Wireless Networks

Ariyam Das
Department of CSE
Jadavpur University, India
AriyamDas@gmail.com

Chittaranjan Mandal
Department of CSE and School of IT
IIT Kharagpur, India
Chittaranjan.Mandal@iitkgp.ac.in

Chris Reade
Department of Informatics and Operations Management
Kingston University, UK
Chris.Reade@kingston.ac.uk

Manish Aasawat
Department of CSE
Jadavpur University, India
ManishkrAasawat@gmail.com

Abstract—A minimum connected dominating set (MCDS) offers an optimized way of sending messages in wireless networks. However, constructing a MCDS is a NP-complete problem. Many heuristics based approximation algorithms for MCDS problems have been previously reported. In this paper, we propose a new degree-based multiple leaders initiated greedy approximation algorithm (PSCASTS) based on the selection of a pseudo-dominating set and an improved Steiner tree construction. We also show that our PSCASTS outperforms existing CDS construction algorithms in terms of CDS size and construction costs. The simulation results show that PSCASTS constructs better non-trivial CDSs for networks with uniform, nearly-uniform and random distribution of sensor nodes. While PSCASTS retains the current best performance ratio of \((4.8+\ln5)\text{opt}+1.2\), \text{opt} being the size of an optimal CDS of the network, it has the best time complexity of \(O(D)\), where \(D\) is the network diameter.

Keywords—connected dominating set (CDS); maximal independent set (MIS); Steiner tree; routing backbone; unit disk graph (UDG)

I. INTRODUCTION

Wireless ad hoc and sensor networks form an important part of the next generation network in providing flexible deployment and mobile connectivity. They consist of static or mobile nodes where each node, having an omni-directional antenna, can broadcast messages to all the nodes within its transmission range. When communicating parties lie outside the single hop radio transmission range, a communicating session between them is established through multi-hop routing using a virtual backbone. Since no fixed infrastructure and centralized management is present in wireless networks, a special connected component of the network, called CDS, is widely used as a virtual backbone for efficient routing and connectivity management in such networks. In remote data gathering applications, a CDS is more particularly used as a data aggregation backbone for in-network data aggregation to optimize network communication, thereby saving communication energy and extending network lifetime [16].

A dominating set is a subset of all the nodes in the network such that any node in the network is either in the set or is adjacent to at least one node in the set. If the nodes in the dominating set induce a connected component, then the latter is known as a connected dominating set (CDS). The nodes in the CDS are called dominators, otherwise dominatees. The CDS ensures that, in the absence of any transmission error, it can receive a packet from any node in the network and can retransmit it to any other remote node. The search space for any route is reduced to the CDS. Thus, during routing, broadcasting responsibility lies only on the CDS nodes, instead of all the nodes in the network. As only the CDS nodes maintain routing information, reduced CDS size effectively saves storage space. Also, a small sized CDS makes routing easier as it can reduce transmission interference and the number of control messages. In other words, to reduce the communication overhead, increase the convergence speed and simplify connectivity management, it is desirable to construct the minimum connected dominating set (MCDS) of the network. However, computing MCDS is a NP-complete problem [1]. So, only distributed approximation algorithms in polynomial time are practical for finding out MCDS in wireless networks. For energy constrained wireless networks, a distributed approximation algorithm should not only construct thinner CDSs but also construct CDSs with low computation and communication costs so as to facilitate speedy switches between disjoint CDSs, thereby extending battery lifetime and optimizing power consumption [17]. Generally, the quality of the CDS is evaluated by its approximation factor, which is the ratio of its size to the size of the MCDS. The construction cost is measured by overall message and time complexities.

A. Prior related research

The different approaches found in the literature on the CDS construction problem can be broadly classified into the following three categories based on the network information they use — centralized algorithms, localized algorithms and distributed algorithms.

Guha and Khullar [2] first gave two centralized greedy algorithms for CDS construction in general graphs having approximation ratio of \(O(\ln\Delta)\), \(\Delta\) being the maximum degree of a node in the graph. The centralized algorithms reported in [2,3] require global information of the complete network, making them unsuitable for wireless networks which do not have centralized control.

In a localized algorithm for CDS construction, Adjih [4] reported an approach based on multipoint relays (MPR) but no approximation analysis of the algorithm is known as yet. Based on the MPR approach several extensions have been reported leading to localized MPR based CDS construction. But the localized approaches, without an approximation factor to ensure an upper bound on CDS size, are not highly effective.

The distributed approach is the best alternative for CDS construction. Distributed algorithms can be based on a single leader or multiple leaders. In the algorithm reported in [5], Wu and Li first constructed a trivial CDS and then redundant nodes are deleted based on two sets of pruning rules. The algorithm requires that each node should know its 2-hop neighbors. The performance ratio of Wu and Li’s algorithm as reported in [6] is \(O(n)\), \(n\) being the network size. The performance ratio of distributed algorithms reported by Stoimenovic et al. in [7] is
yielding $O(n)$ message complexity and $O(\text{factor to } 8|\text{opt}| \text{ in } [11])$. Cardei's distributed algorithm grows in the network. Later, Cardei improved the approximation from a single leader and uses 1-hop connectivity information complexity of $O(n \log n)$, $|\text{opt}|$ being the size of an optimal CDS factor of $8|\text{opt}|+1$, a time complexity of $O(n)$ and a message tree rooted at the leader. The algorithm has an approximation factor of $8|\text{opt}|+1$, a time complexity of $O(n)$ and a message complexity of $O(n \log n)$.

Alzoubi’s multiple leader based distributed approach [6] first constructs an MIS by comparing node IDs within a 1-hop neighborhood without spanning a tree or selecting a leader. In the next phase, the MIS nodes are interconnected to form a CDS. The algorithm has an approximation factor of $192|\text{opt}|+48$. In a similar work [12], the authors reported a distributed algorithm with a performance ratio of 172.

Among all the approximation algorithms for distributed CDS construction in UDGs [1], the best known approximation factor is $(4.8+\ln n)|\text{opt}|+1.2$, achieved by Li’s S-MIS algorithm in [13] and collaborative cover heuristic in [14]. Both the approaches first construct an MIS and then tap the MIS nodes through a Steiner tree construction. In [14], the MIS is constructed using effective coverage as a metric. However, the collaborative cover heuristic [14] has a high message complexity of $O(n \Delta^2)$ and time complexity of $O(n)$.

**B. Motivation and contributions**

A maximal independent set (MIS) is an independent dominating set where no two nodes are adjacent to each other. The recent competitive distributed algorithms [10, 11, 13, 14], achieving constant satisfactory performance ratio, constructs MISs with a specific property stated in [13]. The characteristic property of such MISs is that any pair of complementary subsets of the MIS is separated by exactly two hops. This underlying property assists the interconnection of MIS nodes in the second phase. However intuitively in an MIS, a node can be separated from its nearest node by at most 3 hops. These MISs with lower cardinalities can effectively reduce the CDS size further and improve the ratio of number of connectors to number of independent nodes. The latter ratio has significant impact on the lifetime of the network. We cite an example to demonstrate this point. For the graph given in fig. 1(a) the MIS selection schemes [10, 11, 13, 14] will select an MIS of size 5 comprising of nodes 1, 3, 5, 8 and 10. Each MIS node is separated from its nearest neighbor in the MIS by exactly 2 hops. However, nodes 1, 4 and 9 alone can also form an MIS of the same graph. In the latter case, each MIS node is separated from its nearest MIS neighbor by 3 hops. But to construct such MISs in a distributed fashion, each individual node requires at least 2-hop neighborhood connectivity information. Furthermore Steiner tree construction [13, 14] in the post MIS selection phase will also incur a higher number of message exchanges. Consequently, the communication overhead will be higher.

We also noted that in post Steiner tree construction, some dominators from the MIS can be downgraded to dominatees without any loss in connectivity or coverage of the CDS. This is intuitively true when the neighbors of an MIS node are covered by all the Steiner nodes connecting it. We illustrate this through an example. In the graph given in fig. 1(b) the MIS with minimum size consists of nodes 5, 6 and 7. We will essentially require nodes 2 and 10 to connect the MIS. Thus, nodes 7, 2, 6, 10 and 5 form the CDS. But after CDS construction, node 6 in the CDS becomes redundant. Nodes 7, 2, 10 and 5 alone can also form a CDS. Thus, ideally an MIS should be treated as a pseudo-dominating set (PDS) since post CDS construction some MIS nodes can be removed from the CDS to further reduce CDS size. In other words, a PDS is constructed as an MIS, but it may no longer remain a dominating set after final CDS construction.

We are motivated towards dealing with the issues cited above to reduce the CDS size further at an optimal trade-off in the number of messages exchanged. Thus, in this paper we report a two-phase greedy approximation scheme called PSCASTS (Pseudo-dominating Set Construction And Steiner Tree Spanning) which contributes towards improving the CDS size further than previous approximation algorithms.

The major contributions of this paper are summarized below:

(i) A pseudo-dominating set (PDS) construction which helps in identifying smaller cardinality MISs.

(ii) An improved Steiner tree construction to connect the PDS nodes and selectively remove them to build a better CDS.

(iii) PSCASTS is a multiple-leader based distributed algorithm and has the best time complexity of $O(D)$, $D$ being the network diameter.

(iv) PSCASTS identifies non-trivial CDSs with smaller sizes even when the distribution of sensor nodes in the network is uniform or nearly-uniform.

The rest of the article is organized as follows. In section II, we discuss the network model. Section III explains our algorithm and discusses its distributed implementation. The simulation results are analysed in section IV. Finally, section V concludes the article.

**II. NETWORK MODEL**

In this section we state our assumptions regarding the development of the ad hoc network model:

1) The nodes do not have any geometric or topological information. They do not even have knowledge of their distances to their neighbors.
2) Each node has a unique ID.
3) Nodes exchange hello messages to identify their single-hop neighbors and ascertain their own degree.
4) All the hosts are deployed in a 2-D plane and their maximum transmission ranges are the same.
5) The resultant topology of the network is modeled as a unit disk graph (UDG) in which two nodes are connected by a wireless link if their distance does not exceed the transmission range. All the links are bidirectional.
6) The communication overhead due to interference is negligible.
7) The MAC layer is responsible for scheduling the transmission of messages.
8) The computation is partitioned into rounds, where the nodes receive the messages sent in the previous round, execute local computations and send messages to the neighbors in the next round.

III. PROPOSED APPROACH

Our CDS construction scheme PSCASTS works in two localized phases namely —

A. Pseudo-dominating set or PDS construction
B. Improved Steiner tree construction

A. PDS construction

In this phase, we greedily construct a PDS as an MIS with lower cardinality as compared to other MIS algorithms. Any pair of complementary subsets of our PDS can have a distance of either 2 or 3 hops. Some of the nodes in the PDS (virtual-dominators) may not be included in the CDS after the second phase depending on coverage of the connectors connecting them to the rest of the CDS. The virtual-dominators will act as dominators throughout the CDS construction, particularly during the second phase when connectors are selected to connect both dominators and virtual-dominators. We construct the PDS through a simple degree-based algorithm in linear message complexity. The algorithm requires each node to know its neighbors and neighbors of neighbors. The algorithm is not based on selection of a specific initiator (such as a base station) and may be initiated by multiple leaders yielding an excellent linear time complexity. Thus the first phase of our approach for an optimal CDS construction maintains overall low time and message complexities retaining the basic characteristic of degree-based algorithms and at the same time successfully reduces the CDS size by reducing the cardinality of MIS. Our greedy algorithm, for determining the PDS of a graph is given in Algorithm 1.

Algorithm 1 Determination of PDS of a graph

Input: G(V,E) ← A connected graph where each node has an unique ID; each node in V knows its distance-1 neighbors and distance-2 neighbors and their respective degrees; all the nodes in V are marked white.

Outputs: D ← Set of all black dominator nodes; VD ← Set of all grey virtual-dominator nodes; PDS ← MIS of the graph G(V,E) comprised of black and grey nodes.

1. Initialize PDS, D and VD to Ø.
2. If any node u in V-D has highest degree in its 1-hop and 2-hop neighborhood then
   (i) D ← D U {u} /* Selecting u as dominator. */

(ii) Change color of u from white to black.
(iii) Delete all the adjacent nodes of u and edges incident on them from G.
(iv) Update degree of the remaining nodes in G.
/* In case of a tie, a node which initially had higher degree is selected as dominator. If still a tie persists then the node with least ID is given preference. */
3. Repeat from Step 2 if maximum nodal degree in G>0
4. Each of the undeleted nodes in G, excepting the dominators in D, with degree 0 is added to VD and its color is changed to grey. /* Adding virtual-dominators. */
5. PDS ← D U VD.
/* Dominators & Virtual-Dominators both form the PDS */

In the graph shown in fig. 2 nodes 4, 6 and 10 form the PDS. Nodes 4 and 10 are selected as dominators whereas node 6 is chosen as a virtual-dominator. The recent collaborative cover heuristic [14], which produces smaller MISs than previous MIS selection techniques [10,11], yields an MIS size of 3 or 5 for the same graph in fig. 2 depending on whether node 6 is selected as the initiator or not. In addition, the border effect that arises when dealing with the border nodes such as 3, 5, 9 and 11 can also account for a 67% increment in MIS size as compared to our PDS size. Our scheme does not suffer from leader selection or border effects. We also perform a simple experiment to substantiate that our PDS has smaller cardinality than the MIS selected from other CDS construction schemes. For varying sizes of connected networks in UDGs, we compare the cardinality of our PDSs with that of the MISs obtained from collaborative cover. Each of the approaches was run 100 times. The averaged results are reported in fig. 3. Fig. 3 demonstrates that our PDS has smaller sizes compared to the MIS selected from collaborative cover.

It is important to note that all the neighbors of a virtual-dominator are covered by dominators chosen earlier during PDS construction. The PDS requires the virtual-dominator because there is no other node in the PDS which is adjacent to the virtual-dominator. It is easily verifiable that a virtual-dominator is exactly one hop apart from its nearest dominator. So when connectors interconnect the dominator-virtual-dominator pair, adding virtual-dominators in the CDS may become redundant as explained in the next step. In the following phase, virtual-dominators meeting certain criteria are selectively discarded.

Figure 2. CDS formed by the black nodes (dominators) and the dark grey nodes (connectors) selectively discarding virtual-dominator (light grey).

B. Improved Steiner tree construction

A Steiner tree for a given subset of vertices (terminals) I in a graph G(V,E), is a tree interconnecting (tapping) all the terminals in I using a set of Steiner nodes in |V(G) - I|. In this phase, we tap all the dominators and virtual-dominators in the PDS by selecting Steiner nodes from the dominatees greedily to construct a Steiner tree spanning all the nodes in the PDS.

1603
A property of the UDG [15] states that any node is adjacent to at most five independent nodes. Therefore, any Steiner node can tap at most five terminals from the PDS. Also since a Steiner node is being used as a connector, it essentially connects at least two nodes. According to our approach, a Steiner node will always connect either a dominator-connector, dominator-dominator or dominator-virtual-dominator pair. Based on this, we formulate our approach for the second phase in Algorithm 2.

Algorithm 2  Steiner Tree Construction

Input: G(V,E) ← A connected graph where each node has an unique node ID; each node in V knows its distance-1 neighbors and distance-2 neighbors and their respective degrees; PDS ← Pseudo-Dominating Set of G; VD ← Set of all virtual-dominators in PDS.

Output: CDS ← Connected Dominating Set of the graph G(V,E) formed by black, grey and blue nodes.

1. All dominators and virtual-dominators form separate components.
2. Select u as connector from all rival dominatees in V-PDS, adjacent to components which u covers, according to the following criteria –
   (i) Dominatee which is adjacent to maximum number of separate components is selected.
   (ii) If condition (i) results in a tie then Dominatee with greater degree is chosen.
   (iii) If conditions (i) and (ii) result in a tie then Dominatee with least node ID is preferred.
3. (i) Make connector u and the separate components, that u connects, a single component.
   (ii) Change color of u from white to blue.
4. Repeat from Step 2 until all dominators and virtual-dominators in PDS are in the same component.
5. CDS ← connected component so formed /* PDS nodes are joined to form a single component. */
6. For each virtual-dominator v in VD
   Determine adjacent connectors of v in CDS
   If v is connected to CDS by one connector or any two connectors of v in CDS are adjacent to each other then
   (i) CDS ← CDS – {v} /* Omit virtual-dominator. */
   (ii) Change back color of v from grey to white.

Post PDS construction in the graph shown in fig. 2, node 7 is chosen first as a Steiner node over rival dominatees to tap dominator 4 and virtual-dominator 6. Node 8 then connects dominator 10 to the component 6-7-4. Connectors 7 and 8, connecting the virtual-dominator 6, are neighbors. Therefore, node 6 is discarded from the CDS as shown in fig. 2. The nodes 4, 7, 8 and 10 form a CDS of size 4 even without virtual-dominator 6. The distributed Steiner tree construction in the second phase has linear time and linear message complexities. The CDS constructed by our technique, after discarding the redundant virtual-dominator(s) in the second phase, will still cover the entire network as the coverage of the virtual-dominator(s) will be covered by its adjacent connectors linking the virtual-dominator(s) to the CDS.

We compare our PSCASTS scheme with existing popular CDS construction techniques using the same example network cited in [10]. For the latter example, the dominating tree construction [10] and the 8-approximate CDS algorithm form a CDS of size 6, as shown by the black and grey nodes in fig. 4(a). In contrast, fig. 4(b) shows that our PSCASTS reduces the CDS size to 5 with a PDS of size 3 as opposed to the MIS of size 4 depicted by the black nodes in fig. 4(a).
size and $\Delta$ being the maximum nodal degree in the network. Both the phases in PSCASTS are based on the selection of multiple leaders. This results in a time complexity of $O(D)$ time and $O(D)$ rounds, where $D$ is the network diameter.

IV. SIMULATION RESULTS

In this section we present the simulation results to assess the performance of our PSCASTS scheme and compare it with the existing previous approaches. In the experimental setup, we model the wireless ad hoc network as a set of nodes deployed randomly in a fixed square of dimension $100x100$ square units, known as deployment area $M$, $N$ hosts are randomly generated in $M$ by choosing each of their abscissa and ordinate using a uniform random number generator. We further assume that each node has a uniform transmission range $R$ where $R^2=(d*M)/((\pi*\pi*N))$, $d$ being the network density. The induced graph of underlying network is a UDG. We also ensure that the networks considered are connected. In the following simulations we have considered $R=25$ and used different sets of values of $N$ for different experiments in conjunction with the parameters reported in the previous works. We run the algorithm 100 times for each of the different network sizes. The averaged results are reported in the accompanying figures and table. The complete simulation is carried out in NS-2, a network simulator for wireless networks.

In the first experiment we compare the Steiner nodes required to connect the independent set nodes (i.e. the dominators and virtual-dominators which are not discarded) as a function of network size, using a metric which is the ratio of number of Steiner nodes to the number of independent set nodes (accepted PDS nodes). In this respect we also compare the results with the S-MIS algorithm [13] and collaborative cover heuristic [14] for the network sizes varying from 25 to 225 as given in [14]. The results are reported in fig. 5. The ratio provides a good measure for the average effective connector degree. It is evident from fig. 5 that for large size networks, the ratio for collaborative cover becomes less than 0.3, indicating that on average a Steiner node connects more than 3 independent set nodes. However, for our PSCASTS scheme in large network sizes the ratio tends to be around 0.5, implying that one Steiner node often connects nearly two PDS elements. It is apparent from the results in fig. 5 that the average effective connector degree resulting from PSCASTS scheme is less than that obtained from S-MIS and collaborative cover approaches. This is a noteworthy result and has many practical implications. Less effective degree of a connector indicates that a connector will have to bear less load, which in turn enhances the battery life and hence lifetime of the network.

Next we analyse through simulation the significance of selectively ignoring virtual-dominators from PDS in the CDS construction. We vary the network size from 25 to 250 and determine the fraction of total virtual-dominators being discarded and its impact in the reduction of CDS size. The illustration in fig. 6 shows that nearly all of the virtual-dominators are discarded post Steiner tree construction, occasionally at most one virtual-dominator is retained as a connector for bridging two disjoint components of the CDS. For large network sizes this accounts for around 10% reduction in CDS size.

Next we compare the performance of our PSCASTS scheme with the CDS construction techniques reported in [10,11,13,14]. We determine the CDS size for randomly connected distribution of nodes with the network size $N$ as 20, 50 and 100. The result is presented in fig. 7. Fig. 7 demonstrates that PSCASTS outperforms the dominating tree construction technique [10], 8-approximate CDS algorithm [11], S-MIS approach [13] and collaborative cover heuristic [14] in identifying a smaller size CDS. The results reveal that our approach reduces the CDS size by 16% compared to the previous best collaborative cover heuristic. Furthermore, the S-MIS [13] algorithm involving Steiner tree construction and the degree-based 8-approximate CDS algorithm [11] result in 29% and 34% higher CDS sizes respectively when compared with PSCASTS. Interestingly, for ideal uniform distributions and nearly uniform distributions of nodes, PSCASTS achieves significantly better results in optimizing CDS size than collaborative cover heuristic [14]. Illustrations provided in [14] show that the coverage based heuristic can identify optimal sub-structures and produce CDS of smaller sizes than previously reported degree-based schemes for uniform and nearly uniform node distributions. However, experiments, involving uniform hexagonal distribution and nearly uniform distribution with slight random shifting of nodes, indicate that PSCASTS produces even better CDSs than collaborative cover for both the distributions.

Finally we discuss about the message exchanges needed for CDS construction in PSCASTS. The message complexities of 8-approximate degree-based CDS algorithm [11] and collaborative cover heuristic [14] are $O(n)$ and $O(n\Delta^2)$ respectively, while that of PSCASTS is $O(n\Delta)$, where $n$ is the number of nodes in the network and $\Delta$ is the maximum nodal degree in the network. These theoretically deduced upper bound on the number of messages exchanged indicate that the mean number of messages broadcast in our CDS construction is very close to that of earlier degree-based approach [11] and better than the recent collaborative cover heuristic [14]. Thereby, PSCASTS offers an acceptable trade-off by producing non-trivial CDSs of significantly smaller sizes than all previous approaches at a slightly higher expense of number of messages exchanged as compared to previous degree-based CDS construction techniques.

<table>
<thead>
<tr>
<th>Network Size</th>
<th>S-MIS</th>
<th>Collaborative Cover</th>
<th>PSCASTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5. Performance comparison of number of Steiner nodes and number of independent nodes.
Our algorithm is fully localized which makes it practical for situations where the topology often changes unpredictably. Our future work is to investigate further the maintenance of the CDS when nodes have mobility. It is our interest to further study the effectiveness of using PSCASTS in deploying a CDS as an aggregation backbone in wireless networks for in-network data aggregation through aggregation based energy model. Although our work presented in this paper is limited to UDGs, our reported algorithm is applicable for CDS construction when hosts in a network have different transmission ranges.

V. CONCLUSION

In this paper, we study the problem of constructing a connected dominating set in wireless networks through a new multi-leader initiated degree-based greedy approximation algorithm PSCASTS. The performance ratio of our reported algorithm is $(4.8+\ln 5)\text{opt} + 1.2$, where $\text{opt}$ is the size of any optimal CDS. We have constructed the CDS in two steps, firstly Pseudo-Dominating Set (PDS) selection and then through an improved Steiner tree construction for connecting the different components in the PDS. Eventually, we selectively discarded the virtual-dominators from the CDS. The simulation results comparing our approach with the approaches that were surveyed in terms of CDS size and time complexity.

Our algorithm is fully localized which makes it practical for situations where the topology often changes unpredictably. Our future work is to investigate further the maintenance of the CDS when nodes have mobility. It is our interest to further study the effectiveness of using PSCASTS in deploying a CDS as an aggregation backbone in wireless networks for in-network data aggregation through aggregation based energy model. Although our work presented in this paper is limited to UDGs, our reported algorithm is applicable for CDS construction when hosts in a network have different transmission ranges.

REFERENCES