Genetic-Algorithm-Based Construction of Load-Balanced CDSs in Wireless Sensor Networks

Jing He, Shouling Ji, Mingyuan Yan, Yi Pan, and Yingshu Li
Department of Computer Science
Georgia State University, Atlanta, GA, 30303
Outline

– Background
– Motivation
– Problem Definition
– Genetic Algorithm
– Performance Evaluation
– Conclusions
Background

• What is Dominating Set?
• What is Connected Dominating Set?
• Applications of CDSs.
A **Dominating Set** (DS) is a subset of all the nodes such that each node is either in the DS or adjacent to some node in the DS.
A **Connected Dominating Set** (CDS) is a subset of the nodes such that it forms a DS and all the nodes in the DS are connected.
### Applications of CDS: Virtual backbone

<table>
<thead>
<tr>
<th>Virtual Backbone</th>
<th>Flooding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduction of communication overhead</td>
<td>Redundancy</td>
</tr>
<tr>
<td></td>
<td>Contention</td>
</tr>
<tr>
<td></td>
<td>Collision</td>
</tr>
<tr>
<td>Reliability</td>
<td>Unreliability</td>
</tr>
</tbody>
</table>

CDS is used as a virtual backbone in wireless networks.
Applications of CDS: Broadcast

– Only node in CDS relay messages
– Reduce communication cost
– Reduce redundant traffic
Applications of CDS: Unicast

- Only nodes in CDS maintain routing tables
- Routing information localized
- Save storage space
Applications of CDS: Coverage

Area Coverage Problem

CDS provides connectivity
Applications of CDS: Coverage

Target Coverage Problem

CDS provides connectivity
Other applications

– Data collection
– Data aggregation
– Query
– ......
Motivation

• Related Work
• Load-Balanced CDS
• Challenges
Related Work

– Constructing Minimum CDS (MCDS) --- NP-hard

• *Subtraction-based*: begin with the set of all nodes in the network, then systematically remove nodes by some rules to obtain the CDS [Wu04].
Related work

– *Addition-based*: start from a subset of nodes, then include additional nodes to form the CDS
  
  • The MIS-based CDS: obtain CDS by expending the MIS [Wan09]

  • The tree-based CDS: start from a subset of nodes called *initiators* and grows a dominator tree from each of the initiators [Sakai08]
Variety of CDS

- $k$-connected $m$-dominating Set [Kim10] --- fault tolerance
  - $k$-connectivity: between any pair of backbone nodes there exists at least $k$ independent paths
  - $m$-domination: every dominatee has at least $m$ adjacent dominator neighbors

- Minimum Routing Cost CDS [Du10]--- delivery delay
  - It can guarantee that each routing path between any pair of nodes is also the shortest path in the network.

- D-Hop Dominating Sets [Li09]
  - Minimum CDS with bounded Diameters
Our interests

– Load-Balanced CDS

MCDS

LBCDS
Our interests

– Load-Balancedly Allocate Dominatees
Challenges

How to measure the load balance factor of a CDS?
- CDS $p$-norm
- Allocation Scheme $p$-norm

How to find an LBCDS?
- NP-Hard

How to find a proper trade-off between the size of a CDS and the load balance of a CDS?
Problem Definition

- Terminology
- Problem Definition
Terminology

- **CDS $p$-norm**: \( |D|_p = \left( \sum_{i=1}^{m} |d_i - \bar{d}|^p \right)^{\frac{1}{p}} \).

\[
\begin{align*}
|D|_p &= \sqrt{(6-3)^2 + (3-3)^2} = \sqrt{9} \\
|D|_p &= \sqrt{(4-3)^2 + (4-3)^2 + (3-3)^2} = \sqrt{2}
\end{align*}
\]
Problem Definition

Terminology

- Allocated Dominatee Set: \( \{A(s_i)\} \)
- Valid Degree: \( d_i' = |A(s_i)| + 1 \)
- Allocation p-norm: \( |A|_p = \left( \sum_{i=1}^{M} |d_i' - \varepsilon|^p \right)^{1/p} \), where \( \varepsilon = \frac{n-m}{m} \).

\[ |D|_p = \sqrt{(3 - \frac{5}{3})^2 + (1 - \frac{5}{3})^2 + (1 - \frac{5}{3})^2} = \sqrt{2.67} \]

\[ p = \frac{8 - 3}{3} = \frac{5}{3} \]

\[ |D|_p = \sqrt{(2 - \frac{5}{3})^2 + (2 - \frac{5}{3})^2 + (1 - \frac{5}{3})^2} = \sqrt{0.67} \]
Load Balanced CDS (LBCDS) Problem

For a WSN represented by graph $G = (V, E)$, the LBCDS problem is to find a minimum-sized node set $D \subseteq V$, and a dominatee allocation scheme $A$, such that:

1) The induced graph $G[D] = (D, E')$, where $E' = \{ e | e = (u, v), u \in D, v \in D, (u, v) \in E \}$, is connected.

2) $\forall u \in V$ and $u \not\in D$, $\exists v \in D$, such that $(u, v) \in E$.

3) $\min\{ |D|_p, |A|_p \}$.
Genetic Algorithm

- Overview
- Encode Scheme
- Population Initialization
- Fitness function
- Crossover Operations
- Meta-gene mutation operation
Genetic Algorithm

LBCDS-GA

1. Decide $k$, $P_c$, $P_m$, $G$
2. Population Initialization
3. Evaluate fitness of each chromosome in the population
4. Evolution Process
   - Reproduction: choose two parents chromosomes
   - Recombination
     - Crossover
     - Mutation
   - Replacement: replace the worst chromosome
5. Evaluate fitness of each chromosome in the population
6. Stop? $G$
   - yes: Return the fittest
   - no: Repeat from Step 2

--

Georgia State University
Department of Computer Science

Proud to host MILCOM 2011
Encode scheme

A chromosome with meta-genes and genes
Population Initialization

Genetic Algorithm

1 2 3 4

5 6 7 8

C₁: 0 0 0 1 0 0 1 0
C₂: 1 0 0 1 0 0 1 0
C₃: 0 1 0 1 0 0 1 0
C₄: 0 0 1 1 0 0 1 0
C₅: 0 0 0 1 1 0 1 0
C₆: 0 0 0 1 0 1 1 0
C₇: 0 0 0 1 0 0 1 1
Fitness function

– Given a solution, its quality should be accurately evaluated by the fitness value.

\[ f(C_i) = \frac{n - |D|}{w|D_p| + (1 - w)|A_p|} \]

where \(|D|\) is the size of CDS \(D\)
Crossover operations

O = 6
Parents

C_7
0 0 0 1 0 0 1 1

C_8
0 0 1 0 0 1 1 0

O_L = 3

O_R = 6
Parents

C_7
0 0 0 1 0 0 1 1

C_8
0 0 1 0 0 1 1 0

Offsprings

C_6
0 0 0 1 0 1 1 0

C_9
0 0 1 0 0 0 1 1

Offsprings (corrected)

C_6
0 0 0 1 0 1 1 0

C_9
0 0 0 1 0 1 1 1

Offsprings (no correction)

C_{10}
0 0 1 0 0 1 1 1

C_1
0 0 0 1 0 0 1 0

C_{10}
0 0 1 0 0 1 1 1

C_1
0 0 0 1 0 0 1 0
Meta-gene mutation operation
Performance Evaluation

• Simulation Settings
• Simulation results
Simulation Settings

- $n$ nodes are randomly deployed in a fixed area of 300m $\times$ 300m. All nodes have the same transmission range 50m.
- CDS-based aggregation used as the communication mode
- All nodes have initially 1000 unit energy. Receiving and transmitting a packet both consume 1 unit energy.

<table>
<thead>
<tr>
<th>GA Parameters and Rules</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population size ($k$)</td>
<td>50</td>
</tr>
<tr>
<td>Number of total generations ($K$)</td>
<td>100</td>
</tr>
<tr>
<td>Selection scheme</td>
<td>Rank Selection</td>
</tr>
<tr>
<td>Replacement policy</td>
<td>Elitism</td>
</tr>
<tr>
<td>Crossover probability ($p_c$)</td>
<td>1</td>
</tr>
<tr>
<td>Gene mutation probability ($p_m$)</td>
<td>0.2</td>
</tr>
<tr>
<td>Meta-gene mutation probability ($p_i$)</td>
<td>1</td>
</tr>
</tbody>
</table>
Simulation Results

Performance Evaluation
Conclusions

– The first work to study the LBCDS problem in WSNs.

– In order to measure the load balance of a CDS and a dominatee allocation scheme, we define new metrics CDS $p$-norm and Allocation Scheme $p$-norm.

– We propose a GA to solve the LBCDS problem.

– We also conduct simulations to validate our proposed algorithm.
Q & A