Distributed Indexing and Data Dissemination in Large Scale Wireless Sensor Networks

Yiwei Wu and Yingshu Li
Georgia State University

Outline

• Motivation
• Introduction
• Connected dominating set Based Index (CBI) scheme
• Simulation
• Conclusion and Future work
Motivation

• Many data dissemination techniques proposed for wireless sensor networks may not work well in a large scale sensor network
  – where a huge amount of sensing data are generated.
• Design goals
  – providing timely responses to queries
  – Ensures scalability and load balance

Introduction

• Local Storage (LS)
Introduction

• External Storage (ES)

• Data-Centric Storage (DCS)
Introduction

- Index
  - to provide low average query and storage communication
  - Quad-tree approach

- Index
  - grid approach
CBI data dissemination scheme

• For a graph $G(V,E)$, a Dominating Set $S$ of $G$ is defined as a subset of $V$ such that each node in $V \setminus S$ is adjacent to at least one node in $S$.

• A Connected Dominating Set (CDS) $C$ of $G$ is a dominating set of $G$ which induces a connected subgraph of $G$.

CBI – Cont’d

• A k-hop dominating set $D$ in $G$ is a set of nodes with the property that every node in $G$ is at most $k$ hops away from at least one of the nodes of $D$. 
CBI – Cont’d

• Network hierarchy

CBI – Cont’d

• Storage Nodes
  – K-hop dominating set of the whole network
  – Sensing data are collected and stored at the nodes close to the sensing nodes.
  – can combine data from different sources by using functions such as suppression (eliminating duplicates), Min, Max and Average
CBI – Cont’d

• Index Nodes
  – Connected m-hop dominating set
    • Dominates all storage nodes only
  – We are not interested in any particular form of the index structure used in a single node, e.g. quad-tree, B-tree, etc.
    • gives us more flexibility to process range and binary queries.
  – each index node only stores one copy index of its dominatees (storage nodes).
    • the query will be flooded to all the index nodes to get the query result
    • this flood overhead is much lower since the size of index node set is smaller enough compared with the total number of nodes in the whole network.

CBI – Cont’d

• An example
Simulation

- Simulation Setting
  - K = 3 and m = 5
  - 2000 nodes
  - 150x150 square
  - Transmission range is 10
  - Sensing range is 5
  - The size of data message $S_d = 80$
  - The size of query message $S_q = 10$
  - The size of an index update message $S_i = 10$
  - 10 mobile targets randomly move whose velocities are 0.25
  - Assume that the result of one target is returned for each query
  - Simulation duration time is 100

Simulation

- *Comparison of the storage nodes and index nodes with different k and m*
Simulation

- Comparing performances of different data dissemination schemes
  - total message complexity
    - total number of messages generated in the whole network.
  - hotspot message complexity
    - the maximum number of messages sent by one single node
  - total traffic complexity
    - the amount of data sent by all nodes
  - hotspot traffic complexity
    - the maximum amount of data sent by each node.
Simulation

![Graph showing performance comparison between CBI, ES, LS, and DCS schemes]

Conclusion

- We proposed an integrated distributed Connected Dominating Based Indexing (CBI) data dissemination method to support scalable handling of large amount of sensing data in large scale wireless sensor networks.
  - CBI can provide timely responses to queries.
  - CBI data dissemination framework ensures scalability and load balancing.
- Simulation results show that the CBI scheme outperforms the ES, LS and the DCS schemes in overall performance.
Future work

- Our future work is how to maintain our data dissemination framework in presence of network dynamic changes.

Q & A

Thanks