Practical Limits on Achievable Energy Improvements and Usable Delay Tolerance in Correlation Aware Data Gathering in Wireless Sensor Networks

Yujie Zhu, Karthikeyan Sundaresan
and Raghupathy Sivakumar
GNAN Research Group, Georgia Tech
http://www.ece.gatech.edu/research/GNAN

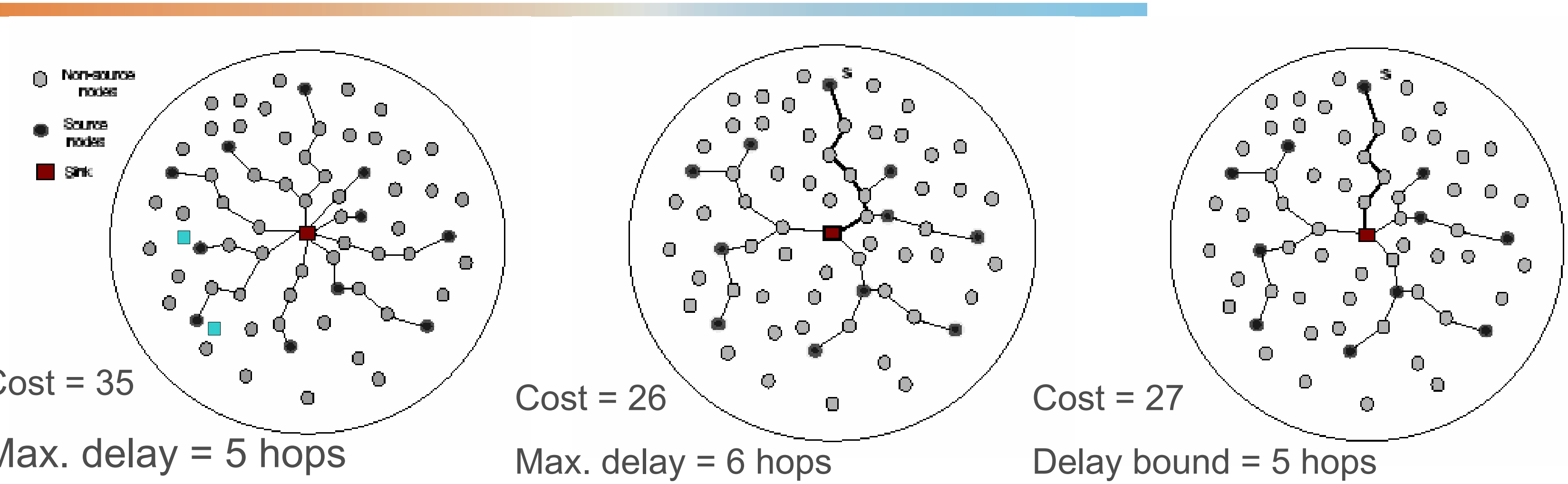
Introduction

- **Data Gathering:** the collection of sensor data from the sensors in the field to the sink for processing.

- Study data gathering from sensors with correlated data
  - ☞ Leverage the correlation by fusing the data inside the network to the best extent possible ⇨ Aggregation.

- Correlation degree (ρ) : A measure of the degree of information redundancy between sensor messages.
  - $\rho = 1$: two messages are perfectly correlated
  - $\rho = 0$: no information redundancy between the two messages
  - $0 < \rho < 1$: two messages are partially correlated.
  - We focus on scenarios where $\rho = 1$ in this presentation

Correlation Aware and Unaware Data Aggregation



- Correlation unaware tree
  - ⇨ Shortest path tree (SPT)
    - Minimize the delay of data gathering
    - Opportunistic aggregation
- Correlation aware tree
  - ⇨ Steiner minimum tree (SMT)
    - Optimize the message cost for data gathering when (ρ = 1)
    - Explicit Aggregation

- Impact of network parameters on aggregation efficiency:
  - Delay tolerance ⇨ Results in less efficient aggregation structure

Goals and Contributions

Goal: Investigate how the energy efficiency of correlation aware aggregation structures is impacted by network parameters:

✓ Node density ✓ Source density ✓ Source distribution
✓ Correlation degree ✓ Delay bound

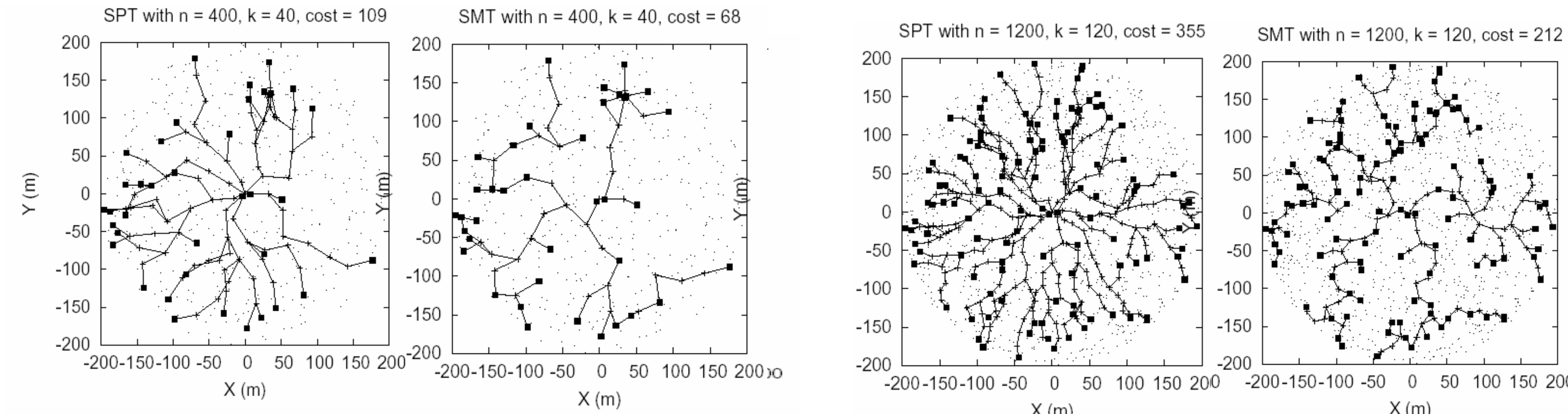
Two fundamental questions:

- Is there a practical limit on the achievable energy improvement by adopting a correlation aware aggregation structure?
  → The energy improvement in using correlation aware aggregation is not as significant as expected and tends to saturate
- Is there a maximum useable delay bound that can deliver the maximum achievable energy improvement?
  → The maximum useable delay bound is only a small constant times the delay along the longest path on the shortest path tree

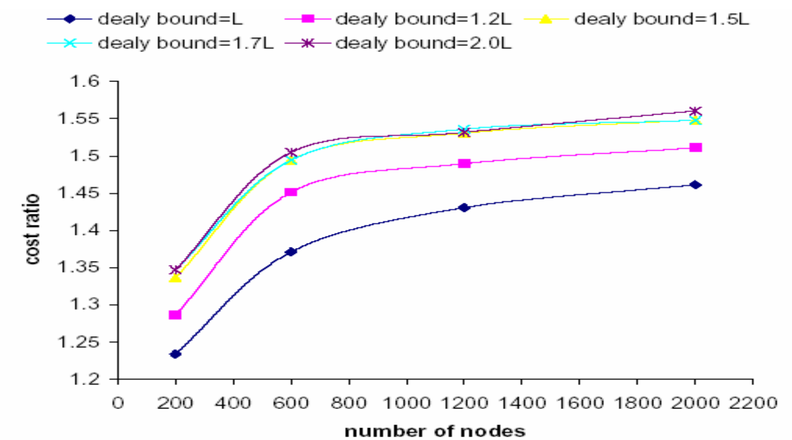
Simulation Model

- Custom-built simulator written in C++
- Evaluation Metrics
  - *cost ratio:* the ratio of the cost of the correlation unaware tree to that of the correlation aware tree over the same set of sources and sink.
- Evaluation Environment
  - *n* sensors randomly distributed in a disk of radius *R*
  - The same transmission range
  - *k* randomly chosen sensors report data to the sink
- Methodology
  - Start from a shortest path tree
  - Calculated DB-SMTs for different delay bounds
  - Vary network parameters such as node density, source density etc.
  - Compare the costs of DB-SMTs and SPTs
- Algorithms
  - SPT: Dijkstra's algorithm
  - DB-SMT: BSMA (bounded shortest multicast algorithm)

Performance Analysis: Node Density



(a) $SPT, n = 400$ (b) $SMT, n = 400$ (c) $SPT, n = 1200$ (d) $SMT, n = 1200$



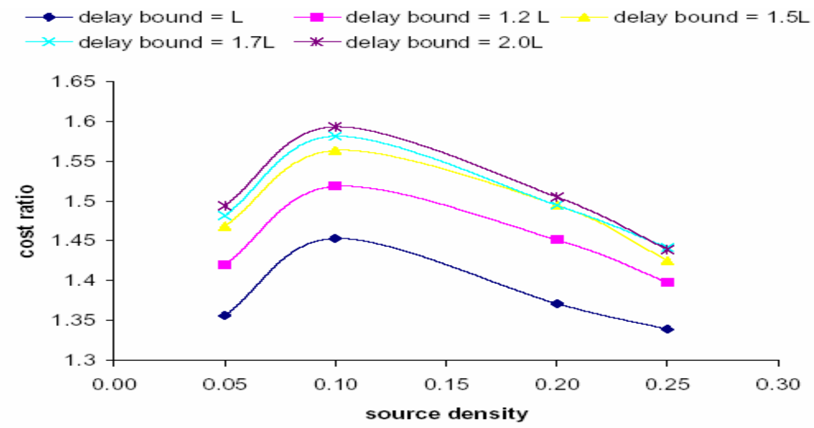
(c) $source\ density = 1/5$

- The cost ratio of SPT over DB-SMT increases with node density
- At low node density, correlation aware data gathering does not bring significant cost improvement.
- The cost ratio tends to saturate when node density is high

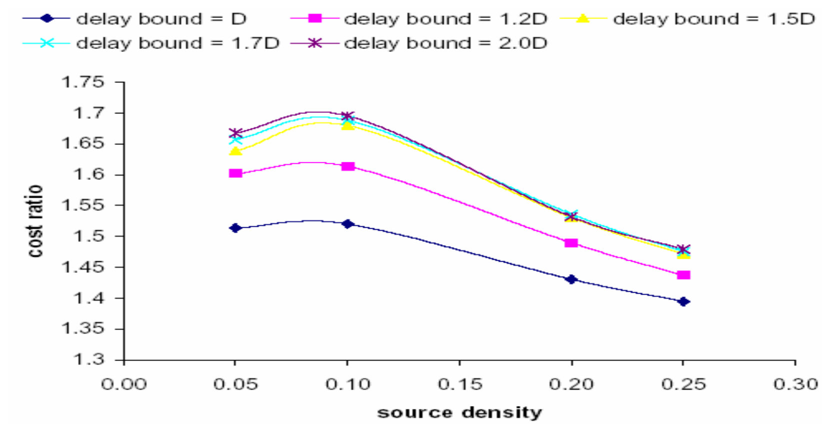
Georgia Institute of Technology

GNAN Research Group

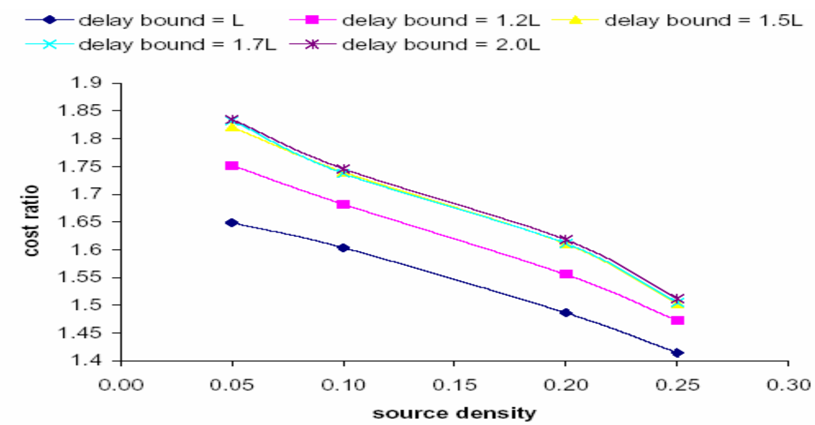
Performance Analysis: Source Density



(a) $n = 600$



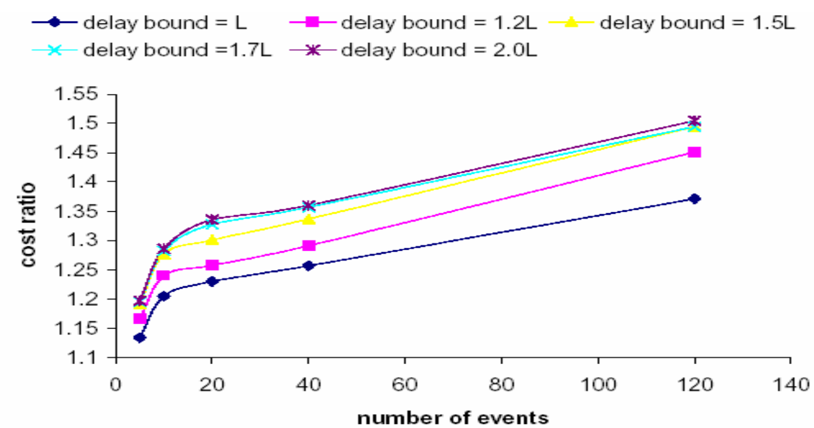
(b) $n = 1200$



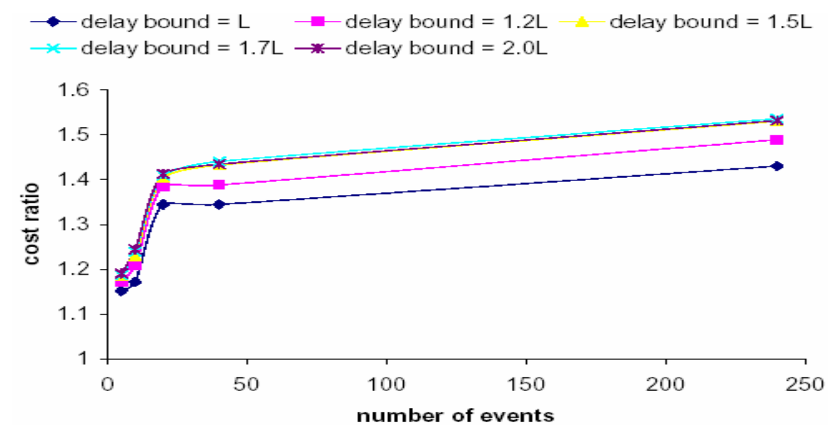
(c) $n = 2000$

- Medium source density ensures the best cost improvement
- The optimal source density for energy improvement reduces as node density increases

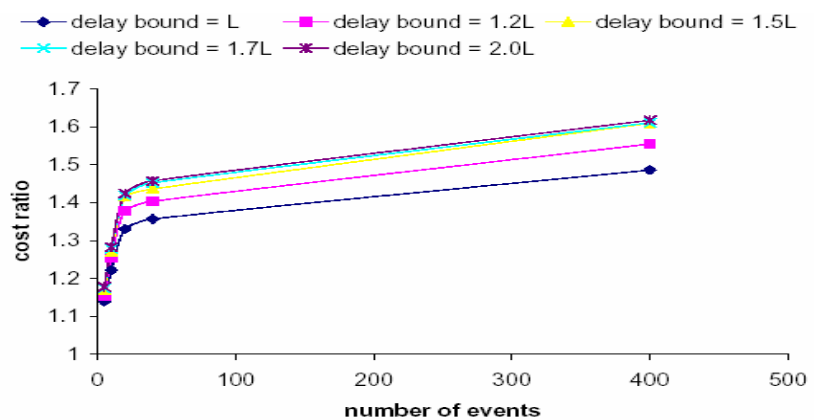
Performance Analysis: Source Distribution



(a) $n = 600$



(b) $n = 1200$



(c) $n = 2000$

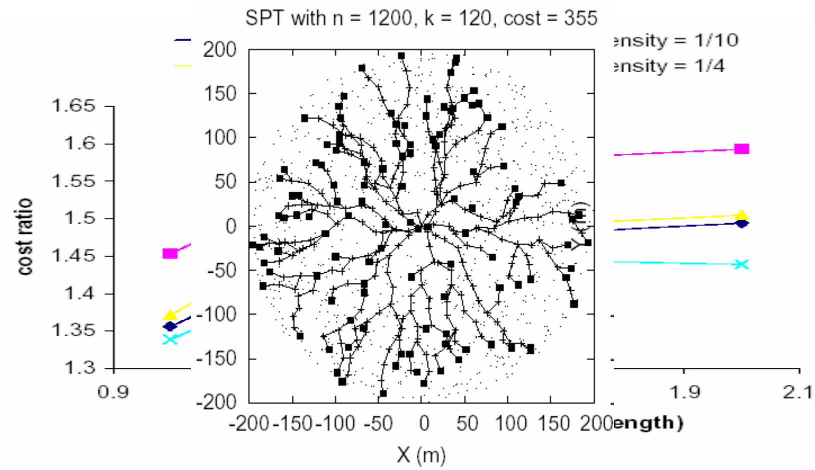
- Non-uniform source distribution

- Cost ratio of SPT over DB-SMT increases as the distribution of source nodes tends towards uniform distribution

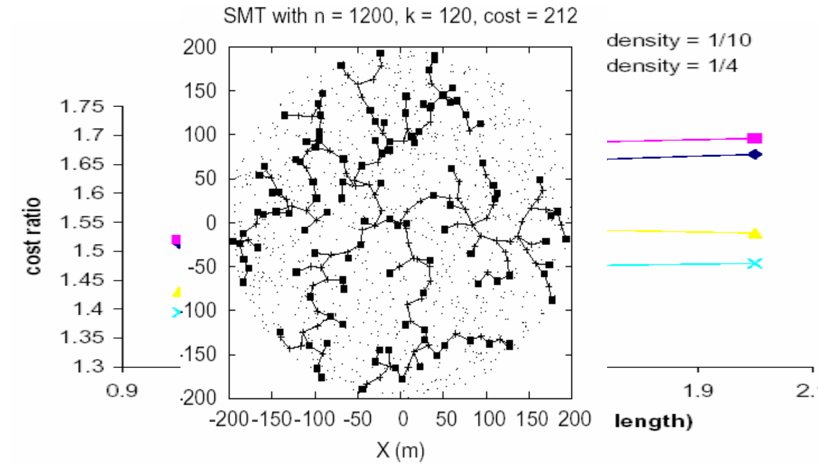
Georgia Institute of Technology

GNAN Research Group

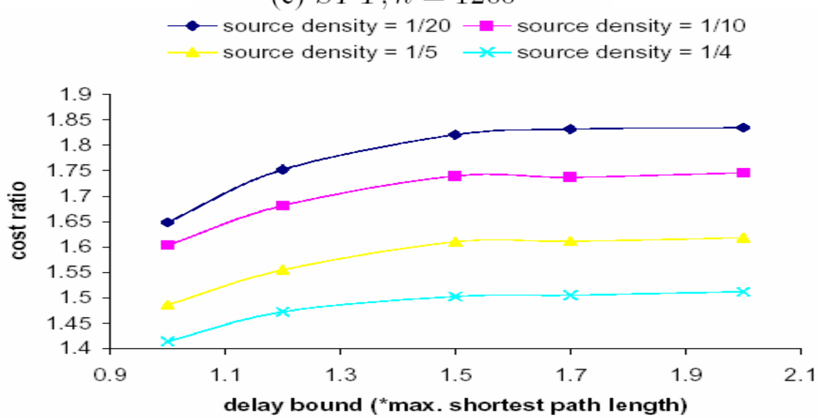
Performance Analysis: Delay Bound



(c) $SPT, n = 1200$



(d) $SMT, n = 1200$



(c) $n = 2000$

- The cost ratio of SPT over DB-SMT increases as delay bound increases.
- Delay bounds beyond twice the maximum shortest path length do not help reduce the DB-SMT cost further.

Georgia Institute of Technology

GNAN Research Group

Major Observations and Practical Implications

- The cost ratios of SPT over DB-SMT scales very slowly (tends to saturate) with respect to node density.
  - For correlation aware aggregation:
    - The energy improvement is limited
    - Explicit communication is usually required for construction
  - For correlation unaware aggregation:
    - Can usually be established in a distributed fashion
    - Proactive establishment is possible without information about the sources and their locations
  - ☞ Correlation aware data gathering may not always be a good choice

- Increasing delay bound beyond a (small) constant order of the longest shortest path length does not help reduce aggregation tree cost further.
  - An application does not have to be designed with large delay tolerances to ensure maximum energy efficiency

Related Work and Conclusions

- Related work:
  - On correlation aware data gathering trees
    - [Cristescu 04]: Proposed two heuristic data gathering structures
    - [Intanagoniwawat 02]: Proposed an approximation of SMT called Greedy Incremental Tree(GIT)
  - On data aggregation tree efficiency
    - [Krishnamachari 02]: Compared data-centric routing with address-centric routing
    - [Pattem 04]: Compared routing-driven compression (RDC) and compression-driven routing (CDR)

- Conclusions:
  - Studied the impact of network parameters on the energy efficiency of correlation aware aggregation trees in wireless sensor networks
  - Drew two major observations
  - Investigated the practical implications