Tributaries and Deltas: Efficient and Robust Aggregation in Sensor Network Streams

Amit Manjhi, Suman Nath, Phillip B. Gibbons

Presented by
Dhara Dave
Aggregation

- Communication costlier than computation
- *In-network* aggregation more efficient than route-all
- Sensor readings are combined into partial results as message moves towards base station
Tree based Aggregation

- Spanning tree rooted at the base station
- In-network aggregation
- Proceed level by level from leaves
- Answer is exact
- High communication failure rate

“Not uncommon to lose 85% of the readings”
Multipath based Aggregation

- Send partial result to multiple neighbours
- *Ring* topology
  - Nodes divided into levels according to distance from base station
  - Level by level routing
- Energy efficient, robust
- Drawbacks: Approximate answers, long message size for some aggregates
Tributary-Delta Approach

- Dynamically adapt aggregation scheme with message loss rate
- Low loss rates – Tree
- High loss rates – Multi-path
Issues

- Selecting which scheme to use in a node
- Communication between nodes using different schemes
- Converting partial results in transitions between approaches
The Tributary-Delta Approach

- Directed Graph $G$
  - Nodes and base station are vertices
  - Successful transmissions are directed edges
  - Labels $M$ (multipath), $T$ (tree)
  - Edge’s label based on source vertex
- Edge correctness
- Path Correctness
- Switchability
Adaptation Design

- User specified threshold – minimum number of nodes that should contribute to aggregate
- Depending on threshold, BS decides size of delta region
- Epoch synchronisation is an issue when switching
  - M to T: must choose its parents from one of its neighbors in level i-1
  - T to M: transmits to all neighbors in level i-1
Adaptation Strategies

- TD-coarse
  - If the % of contributing nodes is below threshold, all current switchable T nodes are switched

- TD
  - Max, Min nodes from subtree of $M$ node not contributing
  - Subtrees with max problem of robustness targeted for Delta region
  - If oscillations in convergence: reduce frequency of adaptation
Identifying Frequent Items

- Important aggregate (eg- for consensus measure)
- Efficient multi-path algorithms do not exist for this
- Min-total load algorithm presented
Min Total-Load Algorithm

(Terms)

- Conversion function
  - To convert between tree and multi-path results

- Summary (T), Synopsis (M)
  - A set of items and their estimates

- Error tolerance
  - function of height of tree (T)
  - Use “Duplicate Insensitive operator” (M)

- Precision Gradient (T)
  - Error increases with each step
  - User specified guarantee is met as long as last error in the gradient is upper bounded by max error
Experimental Setup

- Implement TD-coarse and TD in the TAG simulator
- Aggregate function used is Sum
- Collect every 100 epochs
- Adapt every 10 epochs
- Threshold of contributing nodes = 90%
- Scenarios – LabData, Synthetic
Results

(a) GLOBAL failure

(b) REGIONAL failure
Conclusions

- Tributary Delta combines the advantages of Tree and Multi-path approaches by using them together
- 2 schemes for adjusting balance between Tree and Multi-path nodes in the network (TD-coarse, TD)
- Efficient than existing techniques for realistic loss rates
  - Even for difficult aggregate - “Finding frequent data”
  - Count - 3 times more accurate
Questions?