

Paper presented:

- Suman Nath, Phillip B. Gibbons, Srinivasan Seshan, Zachary R. Anderson. Synopsis diffusion for robust aggregation in sensor networks . Proceedings of the 2nd international conference on Embedded networked sensor systems, pp: 250 - 262, 2004

This paper considers the problem of aggregating sensor data. Most previous work have formed trees as virtual topologies to aggregate sensor data as this method ensure that data is not double counted. However trees are not fault tolerant in anyway since a single link or node failure can cause a significant amount of data to be lost.

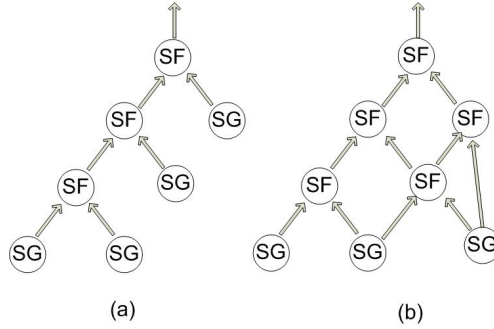
This paper addresses the problem by allowing multiple copies of sensor data to arrive at a queering source from multiple paths, but it is done so in such a way that a repeated reception of the data does not cause any data to be double counted. This paper gives a general framework to show how this can be done for many different aggregate functions. Concurrent work addresses the issuing of dealing with duplicates if the aggregation is in the form of a summation.

In the proposed framework, data is aggregated in a series of three steps. First the data is transformed into a synopsis which will be transmitted over the network. As the data is being aggregated, synopsis at a node may be combined together, in an operation called synopsis fusion. At the final destination, the requested data is extracted from the synopsis. These three steps are named Synopsis Generation  $GS(\cdot)$  , Synopsis Fusion  $SF(\cdot, \cdot)$  , and Synopsis Evaluation  $SE(\cdot)$ .

These functions are required to satisfy four properties.

- $SG(\cdot)$  will generate the same synopsis for any two pieces of data the specified aggregation operation would consider duplicates of each other
- $SF(\cdot, \cdot)$  is commutative. i.e.  $SF(a,b)=SF(b,a)$  for all valid a,b
- $SF(\cdot, \cdot)$  is associative. i.e.  $SF(a, SF(b,c)) = SF(SF(a,b),c)$  for all valid a,b,c
- $SF$  is same-synopsis idempotent. i.e.  $SF(a,a) = a$

With these properties it can be shown that the final synopsis is independent of any variation in order or any duplication that may occur. This can be seen by considering a canonical form, the deep left subtree, in part (a) of the figure below and noting any general tree such as the one in figure (b) can be transformed into it using the four properties. The transform is done



by duplicating nodes that has more than one output. After this the properties allow for nodes to swap places in the tree, or be shown to be redundant and removed.

The synopsis along with the function  $SF$  form a semi-lattice. This means they have an order of sorts in that if  $F(x,y)=z$  then  $F(x,z)=z$  and  $F(y,Z) = Z$ . This is a very useful property since it allows nodes to use passive acknowledgment of message reception despite the fact the next hop node does not retransmit the message. To verify reception, a node snoops on the next hop node, and performs a fusion on the message sent by the next hop node and its own synopsis. If the result of the fusion is equal to the synopsis transmitted by the next hop, its proof the message was received or only contained information redundant to other messages the next hop neighbor did receive.

The paper gives some results showing the improvement this framework can provide over regular tree structures. To show the effectiveness of the framework they propose three aggregation schemes that use it. One forms rings around a query source where nodes report their aggregated synopsis to nodes one hop closer to the source. A variation of this scheme is also given where a node is allowed to change which ring it is in based on how successful its communications are. Also proposed is a flooding scheme were every node sends its synopsis to every other node in radio range. The simulations assumed that probability of a message failure was proportional to the distance between nodes.