Gossip-based Data Fusion Framework for Radio Resource Map

Jin Yang
Department of Computer Science and Automation, Ilmenau University of Technology
P.O.Box 100565, D-98694 Ilmenau, Germany
Email: jin.yang@tu-ilmenau.de

I. INTRODUCTION

Wireless sensor networks are used as efficient technology to detect changes in environment and resources. In disaster scenarios, it is especially useful to form a global view of resource availability map based on locally sensed and measured data, which will support further decision making in system discovery and rescuing process. Each sensor node can provide the measurement and local information in its own sensing and measurement coverage, contributing as a piece of “resource map”, so that the sink can generate the whole picture of “resource map” on resource availability of the entire area. The generation of radio resource map is the research focus of this paper, since in disaster scenarios with the knowledge about availability of radio frequency band, it is possible to estimation the base station location and the its damage status.

During the process of forming the global view of radio resource availability map based on partial view, the data harvesting and data fusion techniques should be explored to guarantee the flexibility as well as energy efficiency requirements in disaster scenarios. Considering the fact that in sensor networks, the communication cost is often several orders of magnitude higher than the computation cost, various structured [1] and unstructured approaches [2] to fuse data along the data transmitting and harvesting path have been proposed for aggregation in data gathering applications where all nodes periodically report to the sink. These techniques are quite suited for such applications as fusion with statistic index (average, maximum, minimum) on single data (temperature, power index, humidity). However, there are several special requirements resulted from generating “radio resource map” in disaster scenarios which can not be easily met using these aggregation protocols. First, the measurement of radio resource is described as $<\text{Azimuth, Elevation, Signal to Noise Ratio, Quality}>$, and the fusion process depends on interplay of these 4 variables among nearby sensor nodes accoring to radio propagations, which increase the difficulty in fusing process. Second, the data harvesting strategy should be robust enough to suffer the disaster scenario itself, which necessitates the self-organization design. Third, energy-efficiency demands close coupling of the information exchange protocol with the radio fusion process.

This PhD thesis work explores the potential of gossip-based data harvesting and data fusion to build a global radio resource map for sensor network in disaster scenarios. The problem considered is basically a multi-source single-sink problem, where the data collection and fusion are tightly coupled with query processing in sensor systems since many queries need fusion methods to process sensor data. The multi-source multi-sink solution which can be divided into query seperately to each sinks, thus a combination of multi-source single-sink problem, may provide further load-balancing characteristics in the resource map query, but is out of the scope of this paper. Three steps are necessary in order to generate the radio resource map: 1. Locally self-organization information exchange to triangulate the sensor nodes in the network; 2. Partial resource map generation on triangle; 3. Harvesting partial resource map toward the sink. There are two main challenges in these three steps. First, the generation of the global view of resource map is not only a simple process of combining and adding all the partial radio resource map together like merging several data table into a long one, the conflicts among the partial views should be solved, and estimations are necessary to fill in the blanks where the measurement data is not available. This is relatively complex data fusion on real radio resources in the sense of map merging. Second, proper self-organization strategy in information exchange which leads to robust and decentralized map generation should be adopted in both the triangularization process and the data collection process, and coupled with the radio data fusion scheme to reduce unnecessary messages as many as possible.

II. GOSSIP BASED RADIO RESOURCE MAP GENERATION

This section describes the above steps in detail:

A. Gossip-based interest publishing and triangularization

There are in general two ways to trigger data collection in sensor networks: the event-driven (source triggers) and the application-driven approach (sink triggers). Due to the need to generate the resource map at the sink, publishing interest will result in two advantages: first, only relative resource information will be collected and fused, it will greatly reduce the computation and communication cost in disaster scenarios; second, the relative location and position from sink to source will be set up, which is necessary in our scenario where no geographic location information is available to all nodes. Otherwise, GPS or other localization protocols will be necessary.
The interest message including the TTL (starting from 0 at the sink) should be gossiped from sink to source. In order to facilitate triangulization process, the number to randomly forward the message is set to be 2, in this way, the triangular are automatically formed. Comparing with the existing naive method which uses flooding to publish an interest (as in directed diffusion), the gossip protocol only chooses random number (in this triangulization process, only 2) of its neighbors to transmit the interest. It can save the bandwidth consumed, reduce the transmission delay efficiently while triangulating the network system.

### B. Partial resource map generation on local triangle

The reason to triangulate the sensor network system for generating local radio resource map is that with measurement of at least three sensor nodes, the observation position can be accurately derived and the radio strength on each position within the triangular area can be calculated too, which results in complete radio resource map generation. In the scenario of radio strength, the measurement data from each sensor node will be \( <\text{Azimuth, elevation, Signal to noise ratio, signal quality}> \). The three points on the triangle can exchange this information and calculate the partial resource map at any point which has extra power or computation resources. And the local conflict resolution is also implemented at this stage.

### C. Harvesting and fusing on reversed gossip tree

As experiment result shows in [2], tree structure has the highest aggregation efficiency but is costy in maintenance and construction. Using gossip to maintain a tree structure can be low-cost and robust: it adapts to dynamic changes in the sensor network, while taking the advantage of high aggregation rate of tree during the data collection process. The map data quality can be justified with improved accuracy during gossiping, the process will stop once the expected map data quality is achieved.

Once the gossip message on interest data is published by the sink and received by the other sensor nodes, the nodes which can provide data will invoke the first round of forwarding data. Besides the neighbors which they random choose, they also need to add to their list the sensor nodes from which they received the interest published. The sensor nodes need to start two threads, one for statistics on counting how many nodes has forwarded the information; the other thread working on processing the data, including fusing the data, getting rid of redundant messages, and again adding the upstream nodes into the gossip list and forward the data towards plus the counts on sub nodes to the gossipping list. The fusion executed on the subroot of the gossip tree to estimate the Signal to Noise Ratio within the triangle is:

\[
SNR_{base} = SNR_{i_1,j_1}d_1^2Q_1 + SNR_{i_2,j_2}d_2^2Q_2 + SNR_{i_3,j_3}d_3^2Q_3
\]

(1)

The reverse forwarding tree is built by each node selecting the neighbor nodes which provides the top \( N \) counts, which defines the width of the reverse forwarding tree. After the selection, the node belong to the sup layer will send a inform message to the selected lower layer node, the tree structure will be formed from the data source till the data sink. The data will follow the tree structure to be delivered and fused towards sink. The reason to use the count field in constructing the tree is to guarantee that nodes with the most submodes being selected into the tree structure. This is based on the observation that packets need to be aggregated early on their route to the sink for efficiency. In the case of a failure happens, the tree structure can rebuild again by choosing the top \( N \) neighbor nodes from the new available neighbor list. This maintenance will again be achieved by using low-cost gossiping among the subtrees.

### III. Evaluations

The thesis plans to use theoretical model, extensive simulations and experiments in verifying the effects of gossip-based map generation process, in comparison with flooding, and structure-free [2] map generations. In theoretical modeling, the thesis is to use queuing and graphic theory to model the performance in terms of total number of packets transmitted in network and fusion efficiency. In simulation part, the focus is on fusion efficiency and robustness. Fusion efficiency is measure by normalized number of transmissions in relation to the different application requirements such as maximum delay (time critical applications), event speed and event size, and node density (deployment requirements). And robustness will look into the situation when part of the nodes failed due to disaster scenario, the time to recover the system. In experiments, the measurement data using directed antenna and omni antenna mobile nodes at local area of Ilmenau is being collected. We hope to use the data to verify the map quality and accuracy using different map generation frameworks in the future.

### IV. Conclusion and future work

The thesis proposes gossip-based data fusion and harvesting techniques to generate the global “radio resource map” in disaster scenarios. Efficient generation of map requires low-cost and robust data exchange combined with map fusion. We proposed gossip protocol to triangulate the network, and to form and maintain the tree-structure for aggregation. In the future, we will compare this strategy using simulation and experiment with naive solutions of flooding with TAG, etc, to further prove these techniques can reduce the communication cost as well as providing fault-tolerance and flexibility in disaster scenarios. Overall, the work expects future approaches will not only regard the data fusion as a separate technique with statistics of local situation, but an efficient way coupled with data harvesting processes in digesting and summarizing more complex local information.

### References
