ROLL WG
Internet-Draft
Intended status: Informational
Expires: April 20, 2016

C. Wang
Tongji University
October 18, 2015

Design of Adaptive Data Aggregation Schemes
draft-wang-roll-adaptive-data-aggregation-00.txt

Abstract

Data aggregation is a key energy saving functionality in wireless sensor networks (WSNs) for both data gathering applications and event-based applications, since the communication cost is often the higher order of the computation cost. Through data aggregation, we can reduce the scale of data while maintaining the correctness of data for a set of symmetric functions called divisible perfectly compressible (DPC) functions. Also the achievable minimum data rate among all sensor nodes is limited for random WSNs if we insist data from ALL sensors should be collected. Hence we use gathering efficiency to indicate the number of nodes whose data are gathered. It is intuitive that there exists a tradeoff between the aggregation throughput and gathering efficiency. This document introduces adaptive data aggregation schemes for WSN to consider the tradeoffs between the aggregation throughput and gathering efficiency. Specifically, the adaptive data aggregation schemes includes two protocols, Single-Hop-Length (SHL) Scheme and Multiple-Hop-Length (MHL) Scheme, for different gathering efficiency requirements.

Copyright Notice

Table of Contents

1. Introduction

Data aggregation is an effective technique for conversing communication energy in low-power and lossy wireless sensor network. In wireless sensor networks, the communication cost is often several orders of magnitude larger than the computation cost. Because of the redundancy in raw data collected from sensors, data aggregation can often reduce the communication cost by removing the redundancy and the left information would be compressed compared with the raw data.

Various data aggregation schemes have been proposed, e.g., cluster-based structures and tree-based structures. In data gathering applications, such as environment and habitat monitoring, sensors periodically send the sensed data to the sink. When the traffic pattern and network topology are assumed to be invariable, the structure-based methods need low-maintenance overhead and are thus applicable for such application scenarios.

It's obvious that the number of sensor nodes whose data are collected makes difference to the performance of aggregation throughput. The achievable minimum data rate among all sensor nodes is severely limited for random WSNs if we insist data from ALL sensors should be collected. Here, we denote the gathering efficiency to the ratio of the number of the sensor nodes whose data are gathered successfully to the total number of sensor nodes in the network. A high gathering efficiency means most of sensor nodes' data are collected. Collecting data from a subset of sensor nodes is reasonable because of the potential spatial correlations among sensed environment. There is a tradeoff between aggregation throughput and gathering efficiency.

In this document, we design adaptive structure-based aggregation schemes for WSNs to achieve the optimal tradeoffs between the aggregation throughput and gathering efficiency. In our protocol, for the neighborhood of every node, we will approximately select a portion of nodes and aggregate their data to the sink. Such a sampling scheme will achieve high aggregation throughput while maintaining the spatial coverage by the sampled sensors. Specifically, we design two schemes, Single-Hop-Length (SHL) Scheme and Multiple-Hop-Length (MHL) Scheme, to improve the tradeoffs between throughput and gathering efficiency. Both propsed aggregation schemes are based on lattices.Section 4 will give the details of scheme lattice. And the details of two schemes will be coverd in Section 5 and Section 6, respectively.

This document is organized as follows: Section 3 defines the Terminalogy and Notation in this document. Section 4 will give the network model used in this document and the concept of lattice in two schemes. Section 5 and Section 6 give the details of two schemes.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Terminology and Notation

This document uses the following additional terms: DPC, SHL, MHL.

- DPC functions

  For data gathering, we focus on an important set of symmetric functions called divisible perfectly compressible (DPC) functions, such as the mean, max, and kinds of indicator functions that will be used to compute the data aggregation.

- SHL

  Single-Hop-Length Scheme, in which the routing is nonhierarchical and consists of the links with similar lengths. Selected sensors are at most one hop away from the adjacent nodes. The details will be covered in Section 5.

- MHL

  Multiple-Hop-Length Scheme, in which the routing is hierarchical and consists of the links with various lengths. Backbone stations are at most one hop away from the adjacent nodes, however other sensors with a long link to the backbone station should be removed. The details will be covered in Section 6.

4. System Model

4.1. Network Model

We consider a random WSN, where sensors are deployed on the 2-dimension plane according to a Poisson point process of density r with r belongs [1, n] where n is the number of randomly placed sensors.

4.2. Scheme Lattice

Both proposed aggregation schemes are based on lattices. Definition of Scheme Lattice: Partition a square region A = [0, a] * [0, a] into a lattice consisting of square cells of side length L, we call the produced lattice scheme lattice.

5. Single Hop Length (SHL) Scheme

We design the scheme Single-Hop-Length Aggregation Scheme based on scheme lattice where a = sqrt(n/r). In SHL scheme, The routing is nonhierarchical and consists of the links with similar lengths. By selecting a certain number of sensors in local regions depending on

the given lower bound on gathering efficiency, we can improve the aggregation throughput by deliberating the bottleneck produced by the second limitations, i.e., dense components. Specifically, there are three steps in SHL scheme, local aggregation, horizontal backbone aggregation, and vertical backbone aggregation. Bellow illustrates the SHL scheme:

```
+---------+---------+---------+
| o   o   |  o  o   |  o      |
|   x -------> x -----> X  o  |
|  o   o  | o     o |   ^ o   |
+---------+---------+---|-----+
|  o   o  |   o     |   |  o  |
|   x -------> x -----> x     |
|  o  o   |   o  o  |   ^ o   |
+---------+---------+---|-----+
|   o o   |  o    o |   | o   |
|   x -------> x -----> x     |
| o o     |   o  o  | o   o   |
+---------+---------+---------+
```

5.1. Local Aggregation

In each cell, log2(n) sensors are selected, if applicable. Suppose that each sensor has an associated block of L readins. Then L rounds of measurements from those sensors are aggregated to the aggregation stations by a single hop; all transmissions are scheduled by a 4-TDMA scheme.

5.2. Horizontal Backbone Aggregation

L rounds of data held by each aggregation station are aggregated to the adjacent aggregation stations in the order from left to right in a pipelined fashion; all transmissions are scheduled by a 9-TDMA scheme.

5.3. Vertical Backbone Aggregation

L rounds of data held by each aggregation station are aggregated to the adjacent aggregation stations in the order from bottom to top in a similar pipelined fashion; all transmissions are sheduled by a 3-TDMA scheme.

6. Multiple Hop Length (MHL) Scheme

We design another aggregation scheme MHL scheme based on the scheme lattice where a=sqrt(n/r) and there exists a minimum angle that equals to pi/4 between the boundaries of A and the sides of cells. Choose randomly a sensor from each nonempty cell, called aggregation station, then, we can build the aggregation backbones based on the gathering efficiency required to get the optimal aggregation throughput. The backbone stations, i.e., the stations on the aggregation backbones, are connected by only short links, whereas every peripheral station, i.e., the stations other than backbone stations, can access a specific backbone station node in one-hop transmission. However, the distance between peripheral stations and backbone stations should not exceed a constraint computed to get a high aggregation throughput. Specifically, there are five steps, i.e., Selection, Local Aggregation, Draining Aggregation, Horizontal Backbone Aggregation, and Vertical Backbone Aggregation, to implement MHL scheme. Bellow illustrates MHL scheme:

```
+---------+---------+---------+
| o   o   |  o  o   |  o      |
|   x     |    x    |      o  |
|  o \    | o     o |     o   |
+---- \ --+---------+---------+
|  o   \  |   o     |      o  |
|       ------ x ------ x     |
|  o  o   |    | o  |   | o   |
+---------+----|----+---|-----+
|   o     |  o |  o |   | o   |
|         |    x        x     |
| o       |   o  o  | o   o   |
+---------+---------+---------+
```

6.1. Selection

Choose a subset of cells that contain aggregation stations, denoted by C, in which the aggregation stations are at a certain distance to the corresponding aggregation backbones. The distance is computed based on the gathering efficiency and to get the optimal aggregation throughput

6.2. Local Aggregation

In each cell in C, choose randomly a number of sensors; L rounds of measurements from those chosen sensors are aggregated to the aggregation station by a single hop; all transmissions are scheduled by a 4-TDMA scheme.

6.3. Draining Aggregation

All peripheral stations in C drain the L rounds of data into the corresponding backbone stations by a single hop of a certain distance.

6.4. Horizontal Backbone Aggregation

L rounds of data held by each backbone station are horizontally aggregated to the adjacent backbone stations in the order from left to right in pipelined fashion, until the data are aggregated into the backbone stations on the backbones passed through by the sink node; all transmissions are scheduled by a 9-TDMA scheme.

6.5. Vertical Backbone Aggregation

L rounds of data held by each backbone station in the backbone are aggregated to the adjacent aggregation stations in the order from bottom to top in a similar pipelined fashion; all transmissions are scheduled by a 3-TDMA scheme.

7. Security Considerations

This document has not conducted its security analysis.

8. IANA Considerations

This document does not specified its IANA considerations, yet.

9. Acknowledgments

10. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

Author's Address

Cheng Wang
Tongji University
4800 Cao'an Road, Jiading District
Shanghai
China

Email: cwang@tongji.edu.cn