A new In-network data aggregation technology of wireless sensor networks

XU Jianbo  
College of Computer and Communication, Hunan University, Changsha 410082, China  
Hunan University of Science and Technology, Xiangtan 411201, China  
xjb63@sina.com

ZENG Siliang  
School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan 411201, China

QU Fengjiao  
School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan 411201, China

Abstract

In wireless sensor network, in-network data aggregation can eliminate redundancy, improve data accuracy and thus prolong the lifetime of network. However, most of data aggregation technologies assumed that the aggregator node has been receiving all the data when it performs the aggregation operation, these methods give larger latency to the transmission of network’s data and affect the accuracy of network. Therefore, this paper presents a new data aggregation technology, configuring carefully the aggregator node’s timeouts through making use of a new aggregation timing control scheme, which is to achieve a good trade-off between energy efficiency and data accuracy.

1. Introduction

In wireless sensor network, in-network data aggregation can reduce the number of data packets transmitted and the data conflict, thus raise the data accuracy and data collection efficiency through dealing with the redundant data in-network. But, at the same time, it also brings the largest data latency, thus lowers the precision of the network.

Recently, most of data aggregation technologies adopt the simply fixed timing control scheme, such as Directed Diffusion, LEACH, SCT( Semantic/Spatial Correlation-aware Tree), and so on. In these data aggregation technologies, all aggregator nodes (i.e., intermediate node) would wait for a fixed-period of time before performing aggregation operation, and the period should be long enough. So when the time triggers, the aggregation nodes can receive responses from all of its children. These technologies can save more energy consumption, but bring larger latency to the whole network.

"Cascading timeouts" [1] has been proposed as a highly efficient aggregation timing control scheme. In cascading timeouts, a node’s timeout is set by on the node’s position in the data distribution tree. It is able to achieve energy efficiency without sacrificing data freshness. But, this scheme does not consider the real time data in network.

The WSN’s application-related characteristic has determined that it should adopt proper data aggregation technology to achieve a good trade-off between energy efficiency and data accuracy. Consider most of WSN’s applications use the follow data collection scheme: all nodes periodically report their data detected (non-real time data) to the sink node, and some node report some real time data to sink node at the same time. In this paper, we present a new data aggregation technology with a good timing control scheme for these applications.

2. The new In-network data aggregation technology of WSNs

2.1 Network Model

In this paper, we use efficient cluster-tree-based routing architecture. Its basic ideal is as follows: first we use position information to select some sensors as cluster-heads(CHs), and then we further organize those CHs into a shortest path tree (SPT) topology.

This paper mainly considers the periodic data gathering and event-driven network, that is: after the sink node sends out search query, all of the data source nodes periodically produce sensed data (non-real time data) through the aggregation tree, at the same time, some nodes should instantaneously respond reports if these nodes detect critical data (real time data).

Due to space consideration, we will focus on the timing control scheme of the new data aggregation technology.

2.2 New Timing control scheme

Each intermediate node should make sure the beginning and the lasting of aggregation time, when using in-network data aggregation technology. Generally speaking, the longer lasting time, is the more data packets will be received by sink node, but the longer the network delay is.

In order to deal with the non real time data and real time data in the network at the same time, we propose a new timing control scheme for data aggregation. It adopts differentiated service to non-real time data flow and real time data flow in network.

In this scheme, the network’s data flow can be divided into two kinds of scenes according to the types of detected data packet: (1) Common scene, the network’s data flow is only non real time data, we can use MFS (Multi-level Fusion Synchronization) [2] timing control scheme, mainly
focus on saving energy consumption. (2) Special scene, non-real time data and real time data co-exist, we also use MFS timing control scheme to non-real time data, but adopt immediately transmitted scheme to real time data, focusing on improving the system response speed.

The new data aggregation timing control scheme is as follows:
(1) The sink node sends out search query which includes the periodically interval MAX and delay granularity \( \Delta \).
(2) Each node calculates it timeouts according to the distance between itself and the sink node. For example, the hops between node \( i \) and sink node is \( H_i \), so its timeouts is \( T_i = MAX-H_i \cdot \Delta \).
(3) Each cluster-heads make use of TDMA to gathering all of his children's data. If the received data is non real time data, it is stored in the buffer, the cluster head will aggregate the data in the buffer and send the resultant data to the sink node along the SPT aggregation tree after it timeouts, and each node in the tree waits its own timeouts to aggregate its data in the buffer. If the sensed data is real time data, the cluster head should immediately transmit the data, every intermediate node in the tree transmits the data immediately as the same.

3. Performance Analysis

In the following discussion, \( N \) is the number of the nodes, \( C \) is the number of the cluster-heads, and \( H_i \) is the distance between node \( i \) and sink node.

3.1 Energy Consumption

Through new in-network data aggregation timing control scheme, the number of data packets transmitted in network can divides two kinds circumstance: (1) Common scene that 2.2 section has mentioned, the data flow in the network is only non-real time data flow, each intermediate node aggregates the data received to produce only one data packet. The number of data packets which should be transmitted in each data gathering round is \( \text{IdealPkt/.Round} = N \). (2) Special circumstance, that the non real time data and real time co-exist, suppose the nodes that produce the real time data are \( \{0 < n < N\} \) (under most parts of circumstances, the \( n \) is far smaller than the \( N \)). By using new data aggregation timing control mechanism in the network, the data flow gets different service. The number of data packets which should be transmitted in each data gathering round is

\[
\text{AggPkt/Round} = N + \sum_{i=1}^{N} (H_i - 1)
\]

Without using data aggregation technology in the network, the data packets needed to be transferred is
\[
\text{NoAggPkt/Round} = \sum_{i=1}^{N} H_i
\]

Obviously, \( \text{IdealAggPkt/Round} < \text{AggPkt/Round} \leq \text{NoAggPkt/Round} \). So, the new data aggregation technology has saved a great deal of energy in the whole network.

3.2 Delay

The intermediate nodes should wait timeouts when using in-network data aggregation technology in the network. It will bring a large latency to data flows, then, affect the data accuracy.

Suppose that: \( T \) is the data gathering period, \( H_i \) is the number of hops between node \( i \) and the sink node, \( shd \) is a single hop distance (referred as delay or real distance).

For "Cascading timeouts" \([2]\), each node's timeouts is \( T- (shd*H_i) \). The whole latency of the sensed data is \( T \). The new timing control scheme adopts "Cascading timeouts" to non-real time data flow, meanwhile sends real time data immediately. The whole latency of non-real time in the network is \( T \) with the real time one is \( shd*H_i \), which is far more less than \( T \). Without using data aggregation technology in the network, the latency of all data is \( shd*H_i \).

In the new timing control scheme, the latency of the real time data equals the latency of all data which doesn't use data aggregation technology. So, it improves the real time data's response as well as saving energy consumption.

4. Conclusions

The new in-network data aggregation timing control scheme can achieve a good tradeoff between energy efficiency and data accuracy through differentiated service to non-real time data flow and real time data flow in network, and it is preferred for the network which needs transmitting critical data with high real time response.

This work is supported by Project supported by Hunan Provincial Natural Science Foundation of China,02JJY5006.

5. References