Computer Networks 57 (2013) 1063–1077

Contents lists available at SciVerse ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Network lifetime maximization for time-sensitive data gathering in wireless sensor networks

Feng Shan [a,e,*], Weifa Liang [b], Jun Luo [c], Xiaojun Shen [d]

[a] School of Computer Science and Engineering, Southeast University, Jiangsu, Nanjing 210096, China
[b] Research School of Computer Science, Australian National University, Canberra, ACT 0200, Australia
[c] School of Computer, National University of Defense Technology, Changsha, Hunan 410073, China
[d] School of Computing and Engineering, University of Missouri-Kansas City, Kansas City, MO 64110, USA
[e] Key Laboratory of Computer Network and Information Integration, Ministry of Education, Nanjing 210096, China

ARTICLE INFO

Article history:
Received 12 July 2012
Received in revised form 25 October 2012
Accepted 11 December 2012
Available online 20 December 2012

Keywords:
Wireless sensor networks
Network lifetime prolongation
Energy optimization
Load-balanced spanning tree
Network flow
Algorithm design

ABSTRACT

Energy-constrained sensor networks have been widely deployed for environmental monitoring and security surveillance purposes. Since sensors are usually powered by energy-limited batteries, in order to prolong the network lifetime, most existing research focuses on constructing a load-balanced routing tree rooted at the base station for data gathering. However, this may result in a long routing path from some sensors to the base station. Motivated by the need of some mission-critical applications that require all sensed data to be received by the base station with minimal delay, this paper aims to construct a routing tree such that the network lifetime is maximized while keeping the routing path from each sensor to the base station minimized. This paper shows that finding such a tree is NP-hard. Thus a novel heuristic called top-down algorithm is presented, which constructs the routing tree layer by layer such that each layer is optimally extended, using a network flow model. A distributed refinement algorithm is then devised that dramatically improves on the load balance for the routing tree produced by the top-down algorithm. Finally, extensive simulations are conducted. The experimental results show that the top-down algorithm with balance-refinement delivers a shortest routing tree whose network lifetime achieves around 85% of the optimum.

1. Introduction

Recent advances in electronic and communication technologies make it possible to build a large scale Wireless Sensor Network (WSN) with hundreds of thousands of sensors. Due to its wide range of applications, from environmental monitoring to mission-critical surveillance [19], WSNs have received tremendous attentions and data gathering as its fundamental function has been extensively studied in the past several years. Sensors in most WSNs are powered by energy-limited batteries, and sometimes it is impossible to recharge or replace these batteries when the network is deployed in harsh or human inaccessible environments such as battlefields or nuclear polluted regions. Therefore, energy conservation in this type of network is of paramount importance in order to prolong the network lifetime. Most existing research focused on maximizing the network lifetime by constructing a load-balanced routing tree for data gathering. However, such a tree may contain long routing paths from some sensors to the base station. In order to meet the need of mission-critical applications that require all sensed data sending their data to the base station with minimal delay, this paper aims to constructing a routing tree rooted at the base station that guarantees to forward the sensed data from

* Corresponding author at: School of Computer Science and Engineering, Southeast University, Jiangsu, Nanjing 210096, China.

E-mail addresses: shanfeng@seu.edu.cn (F. Shan), wliang@cs.anu.edu.au (W. Liang), junluo@nudt.edu.cn (J. Luo), shenx@umkc.edu (X. Shen).

1389-1286/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.comnet.2012.12.005

any sensor along a shortest path while maximizing network lifetime. As we here deal with time-sensitive data gathering, we expect to collect the detailed data from all sensors without any aggregation during the data transfer. We thus assume the energy consumption of each node (sensor) is *proportional* to the number of descendants of the node in the routing tree.

This optimization problem is also applicable for non-mission-critical but large scale WSNs. This is because wireless communication (particularly multi-hop relay communication) is unreliable and long routing paths may cause frequent and repeated re-transmissions that lead to network failures. Therefore, a shorter routing path is highly desirable.

1.1. Related work

Data gathering in a WSN means collecting sensed data from every sensor and forwarding the data to the base station. The most popular paradigm of data gathering is in-network processing that constructs a routing spanning tree rooted at the base station (also referred to as a sink). Depending on different applications, two major models are used in data relaying, namely *the aggregated relay* and *non-aggregated relay* models. Most routing trees adopt the aggregated relay model. Under this model, each relay node aggregates all received data from its children and its own into a fixed-size message. The aggregated data are then transmitted to its parent. This type of data gathering is applicable to applications such as database queries AVG, MIN, MAX, COUNT, and so on. In the non-aggregated relay model, the length of a message transmitted by a relay node depends not only on the length of its own sensed data but also on the lengths of the messages received from its children. We refer to this latter one as the message-length dependent data gathering [15].

An extensively studied data gathering problem under both aggregated and non-aggregated models is to find a routing tree that minimizes the total energy consumption or minimizes the maximum energy consumption among individual nodes. Heinzelman et al. [9] initiated this study under the aggregation model and proposed the clustering protocol LEACH that groups nodes into a number of clusters in a self-organizing manner. Then, a cluster-head serves as the local 'base station' to aggregate the messages gathered from its members and forward the result to the sink directly. Lindsey and Raghavendra [17] presented an improved protocol called PEGASIS, in which all nodes in a cluster form a chain and one of them is chosen as the head responsible for reporting the aggregated result to the base station. Kalpakis et al. [13] attacked this problem by formulating it as an integer program and gave a heuristic solution.

A number of research papers have been published [2,3,7,13,15,22,23] that use various energy saving or balancing models. For example, Goel and Estrin [7] addressed the problem of minimizing the total transmission energy consumption, assuming that the aggregation cost at each relay node is a concave, non-decreasing function. They proposed a hierarchical matching algorithm that delivers an approximate solution within a logarithmic factor of the optimum. Cristescu et al. [3] studied the data correlation problem with an objective of minimizing the total transmission energy consumption. They assumed that each node is cognizant of which nodes it should be merged with so that the merged message has a minimal length. They showed that the data correlation problem is NP-complete, and provided an integer program solution, using the Slepian-Wolf coding approach. Rickenbach and Wattenhofer [21] studied the same problem and provided an improved solution with an approximation ratio of $2(1+\sqrt{2})$, using the *shallow light tree* concept [14]. Buragohain et al. [2] studied the min–max model for the network lifetime maximization problem. Instead of minimizing the total energy consumption, they focused on minimizing the maximum energy consumption among the sensors. They showed that finding an optimal routing tree under this model is NP-complete, and proposed a heuristic solution. Liang and Liu [15] also independently showed its NP-completeness and devised several heuristics that trade off between different energy optimization metrics. Intanagonwiwat et al. [12,11] studied the general data gathering issue by incorporating the semantics of an aggregation query into building an energy efficient routing tree that may not necessarily be a spanning tree. For example, they proposed a data dissemination scheme called *directed diffusion with opportunistic aggregation* [11], where data is opportunistically aggregated at relaying nodes on a low-latency tree. They also explored a greedy aggregation by a novel approach [12] that adjusts aggregation points to increase path sharing and thereby reducing the energy consumption.

With different objectives, a number of algorithms have been proposed to produce different routing (spanning) trees. For example, a Breadth-first search tree in Tiny AGgregation service (TAG) [18] aims at minimizing the transmission delay from each sensor to the root, while a degree-constrained spanning tree [2,25,24] focuses on minimizing the maximum energy consumption by any node. Another kind of spanning tree [21] makes a trade off between the energy cost of a minimum spanning tree and the energy cost of a shortest path tree. It seeks a fair balance between the total energy consumption and the maximum energy consumption among the sensors within each data gathering session. Wu et al. [25] considered the network lifetime maximization problem with the same assumption used in [2] that the size of forwarded data from each relay node is identical and the energy consumption at each node is proportional to the number of its children. They generalized the original algorithm for degree-constrained spanning trees [5] to an algorithm for routing trees in sensor networks by incorporating the residual energy into the design. Wu et al. [24] later further extended their results to a routing forest instead of a routing tree.

In addition to the above mentioned tree construction algorithms, special efforts have also been made by researchers for constructing (energy) load-balanced routing trees. For example, Hsiao et al. [10] introduced the dynamic load-balanced tree for a grid-topology of wireless access networks and developed a distributed algorithm. Dai and Han [4] introduced a *hierarchy-balanced tree* and made use of the Chebyshev sum as a measuring criterion