Theoretical Computer Science 509 (2013) 122–139

Contents lists available at ScienceDirect

Theoretical Computer Science

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

Optimal time data gathering in wireless networks with multidirectional antennas

Jean-Claude Bermond [a], Luisa Gargano [b,*], Stephane Perénnes [a], Adele A. Rescigno [b], Ugo Vaccaro [b]

[a] MASCOTTE, joint project CNRS-INRIA-UNSA, 06902 Sophia-Antipolis, France
[b] Dipartimento di Informatica, Università di Salerno, 84084 Fisciano (SA), Italy

ARTICLE INFO

Keywords:
Data gathering
Personalized broadcasting
Multidirectional antennas
Sensor networks
Radio networks
Interference

ABSTRACT

A wireless network consists of a large number of devices, deployed over a geographical area, and of a base station where data sensed by the devices are collected and accessed by the end users. In this paper we study algorithmic and complexity issues originating from the problem of data gathering in wireless networks. We give an algorithm to construct minimum makespan transmission schedules for data gathering under the following hypotheses: the communication graph G is a tree network, the transmissions in the network can interfere with each other up to distance m, where $m \geq 2$, and no buffering is allowed at intermediate nodes. In the interesting case in which all nodes in the network have to deliver an arbitrary non-zero number of packets, we provide a closed formula for the makespan of the optimal gathering schedule. Additionally, we consider the problem of determining the computational complexity of data gathering in general graphs and show that the problem is NP-complete. On the positive side, we design a simple $(1+2/m)$-factor approximation algorithm for general networks.

1. Introduction

Technological advances in very large scale integration, wireless networking, and in the manufacturing of low cost, low power digital signal processors, combined with the practical need for real time data collection have resulted in an impressive growth of research activities in Wireless Sensor Networks (WSNs). Usually, a WSN consists of a large number of small-sized and low-powered sensors, deployed over a geographical area, and of a base station where data sensed by the sensors are collected and accessed by the end users. Typically, all nodes in a WSN are equipped with sensing and data processing capabilities; the nodes communicate with each other by means of a wireless multi-hop network.

A basic task in a WSN is the systematic gathering at the base station of the sensed data, generally for further processing. Due to the current technological limits of WSN, this task must be performed under quite strict constraints. Sensor nodes have low-power radio transceivers and operate with non-replenishable batteries. Data transmitted by a sensor reach only the nodes within the transmission range of the sender. Nodes far from the base station must use intermediate nodes to relay data transmissions. Data collisions, that happen when two or more sensors send data to a common neighbor at the same time, may disrupt the data gathering process. Another important factor to take into account when performing data gathering is the *latency* of the information accumulation process. Indeed, the data collected by a node of the network can frequently change, thus it is essential that they are received by the base station as soon as it is possible without being delayed

* Corresponding author. Tel.: +39 089969725; fax: +39 089969600.
E-mail addresses: lg@dia.unisa.it, lgargano@gmail.com (L. Gargano).

0304-3975/$ – see front matter
http://dx.doi.org/10.1016/j.tcs.2013.03.017

by collisions [16]. The same problem was posed by France Telecom (see [6]) on how to bring internet to places where there is no high speed wired access. Typically, several houses in a village want to access a gateway connected to the internet (for example via a satellite antenna). To send or receive data from this gateway, they necessarily need a multiple hop relay routing.

All these issues raise unique challenging problems towards the design of efficient algorithms for data gathering in wireless networks. It is the purpose of this paper to address some of them and propose effective methods for their solutions.

1.1. The model

We adopt the network model considered in [1,2,9–11,14]. The network is represented by a node-weighted graph $G = (V, E)$, where V is the set of nodes and E is the set of edges. More specifically, each node in V represents a device that can transmit and receive data. There is a special node $s \in V$ called the *Base Station (BS)*, which is the final destination of all data possessed by the various nodes of the network. Each $v \in V - \{s\}$ has an integer weight $w(v) \geq 0$, that represents the number of data packets it has to transmit to s. Each node is equipped with a half-duplex transmission interface, that is, the node cannot transmit and receive at the same time. There is an edge between two nodes u and v if they can communicate. So $G = (V, E)$ represents the graph of possible communications. In fact one gets a symmetric digraph (the transmissions are directed), which is modeled by an undirected graph. Some authors consider that two nodes can communicate only if their distance in the Euclidean space is less than some value. Here we consider general graphs in order to take into account physical or social constraints, like walls, hills, impediments, etc. Simple graphs modeling urban situations are paths, stars, and grids. Although they are not representative of real networks, we study in this article trees as they contain paths and stars as a special cases and they are the first cases where the complexity of the gathering problem is unknown. Furthermore, many protocols of transmission use a tree of shortest paths for routing.

Time is slotted so that a one-hop transmission of a packet (one data item) consumes one time slot; the network is assumed to be synchronous. These hypotheses are strong ones and suppose a centralized view. The values of the completion time we obtain will give lower bounds for the corresponding real life values. Said otherwise, if we fix a value on the completion time, our results will give an upper bound on the number of possible users in the network.

Following [10,12,16], we assume that no buffering is done at intermediate nodes and each node forwards a packet as soon as it receives it. One of the rationales behind this assumption is that it frees intermediate nodes from the need to maintain costly state information.

Finally we use a binary model of interference based on the distance in the communication graph. Let $d(u, v)$ denote the distance (that is, the length of a shortest path) between u and v in G. We suppose that when a node u transmits, all nodes v such that $d(u, v) \leq m$ are subject to the interference of u's transmission and cannot receive any packet from their neighbors. This model is a simplified version of the reality, where a node is under the interference of all the other nodes and where models based on SNR (Signal-to-Noise Ratio) are used. However our model is more accurate compared to the classical binary model ($m = 1$), where a node cannot receive a packet only in the case one of its neighbors transmits. We suppose all nodes have the same interference range m; in fact m is only an upper bound on the possible range of interferences, since due to obstacles the range can be sometimes lower (however, see also [15] for a critique of this model).

Under the above model, simultaneous transmissions among pairs of nodes are successful whenever transmission and interference constraints are respected. Namely, a transmission from node v to w is called collision-free if, for all simultaneous transmissions from any node x, the following holds:

$$d(v, w) = 1, \qquad d(x, w) \geq m + 1.$$

The gathering process is called *collision-free* if each scheduled transmission is collision-free. Therefore, the collision-free data gathering problem can be stated as follows.

> **Data Gathering.** *Given a graph $G = (V, E)$, a weight function $w : V \to N$, and a base station s, for each node $v \in V - \{s\}$ schedule the multi-hop transmission of the $w(v)$ data packets sensed at node v to the base station s so that the whole process is collision-free and the makespan, i.e., the time when the last packet is received by s, is minimized.*

1.2. Gathering vs. personalized broadcasting

Actually, we will describe the gathering schedule by illustrating the schedule for the equivalent personalized broadcast problem, since this latter formulation allows us to use a simpler notation and to get easier proofs.

> **Personalized broadcast:** *Given a graph G, a weight function $w : V \to N$, and a BS s, for each node $v \neq s$ schedule the multi-hop transmission from s to v of the $w(v)$ packets destined to v so that the whole process is collision-free and the makespan, i.e., the time when the last packet is received at the corresponding destination node, is minimized.*

We notice that any collision-free schedule for the personalized broadcasting problem is equivalent to a collision-free schedule for data gathering. Indeed, let $\mathcal{T}$ be the last time slot used by a collision-free personalized broadcasting schedule; any transmission from a node v to its neighbor w occurring at time slot k in the broadcasting schedule corresponds to a transmission from w to v scheduled at time slot $\mathcal{T} + 1 - k$ in the gathering schedule. As the graph is symmetric, when two