Scheduling data gathering with variable communication speed

*Joanna Berlińska**
Adam Mickiewicz University in Poznań, Poznań, Poland

Keywords: scheduling, data gathering networks, variable communication speed

1 Introduction

Gathering data from remote processors is an important stage of many applications. Running computations in distributed systems requires collecting results obtained by many workers. Wireless sensor networks collecting data find environmental, military, health and home applications [1]. Specific communication protocols have been designed for wireless sensor networks to increase data gathering efficiency [5, 6, 9]. General scheduling algorithms for data gathering were proposed in [2, 3, 4, 7]. It was assumed in these papers that the network parameters, such as the speed of communication and processing, are constant. However, in reality the communication speed often changes because of sharing communication links with other users, maintenance activities etc. Hence, in this work we study scheduling for data gathering networks with variable communication speed.

2 Problem formulation

We analyze a star network consisting of m nodes $P_1, P_2, \dots, P_m$ and a single base station. Node P_i has to transfer data of size α_i directly to the base station, possibly in many separate messages. Only one node can communicate with the base station at a time. We assume the linear model of communication, i.e., communication capabilities of node P_i are characterized by a single parameter called *communication rate* (inverse of speed). Thus, if node P_i communicates with rate C, then it transfers data of size x in time Cx. According to the methodology of *divisible load theory* [8], we assume that data size x is a rational number.

It is assumed that the communication rate of a link between node P_i and the base station changes in negligible time, when another application starts using it, and then remains constant for some period of time. In other words, it is a piecewise constant function of time. Let $t_0 = 0$ be the time when data gathering starts. The communication rates change at n moments $t_j > 0$, $1 \leq j \leq n$, $t_1 < t_2 < \cdots < t_n$

* Speaker, email: `Joanna.Berlinska@amu.edu.pl`

and $t_{n+1} = \infty$. The communication rate of node P_i in interval $[t_j, t_{j+1})$ will be denoted by $C_{i,j}$ for $j = 0, 1, \ldots, n$.

Problem DG-VS (scheduling data gathering with variable communication speed) consists in scheduling the communications between the nodes $P_1, P_2, \ldots, P_m$ and the base station so that the whole data is transferred in the shortest possible time T.

3 Offline algorithm

Theorem 1. *The offline version of DG-VS can be solved in polynomial time.*

Proof. Suppose that $T \in [t_k, t_{k+1})$ for given $k \in \{0, 1, \ldots, n\}$. Then, T can be found by solving the following linear program:

$$\text{minimize} \quad T \tag{1}$$

$$\sum_{j=0}^{k} x_{i,j} = \alpha_i \quad \text{for} \quad i = 1, 2, \ldots, m \tag{2}$$

$$\sum_{i=1}^{m} C_{i,j} x_{i,j} \leq t_{j+1} - t_j \quad \text{for} \quad j = 0, 1, \ldots, k-1 \tag{3}$$

$$\sum_{i=1}^{m} C_{i,k} x_{i,k} \leq T - t_k \tag{4}$$

In the above program, $x_{i,j}$ ($1 \leq i \leq m, 0 \leq j \leq k$) are rational variables representing the amount of data sent by node P_i in interval $[t_j, t_{j+1})$. We minimize the data gathering completion time T. By constraints (2) each node transfers all its data to the base station. Inequalities (3) and (4) guarantee that the communications fit in the time intervals where they are assigned. Linear program (1)-(4) has $m(k+1) + 1 = O(mn)$ variables and $m + k + 1 = O(m + n)$ constraints, and hence it can be solved in polynomial time.

In order to solve DG-VS one can use binary search to find the smallest k for which program (1)-(4) has a solution. The number of binary search iterations is $O(\log n)$. The optimum communication schedule can be obtained from the values of variables $x_{i,j}$. Namely, in each interval $[t_j, t_{j+1})$ we schedule consecutively communications from nodes $P_1, P_2, \ldots, P_m$ of sizes $x_{1,j}, x_{2,j}, \ldots, x_{m,j}$ correspondingly, starting at time t_j. Thus, problem DG-VS can be solved in polynomial time. □

4 Online algorithm

Let us assume that although we do not know the exact ranges of communication speeds changes, the relative range of communication rate changes is known. Namely, if $C_i^{\max}$ and $C_i^{\min}$ are the maximum and the minimum communication rate of node P_i, then $\frac{C_i^{\max}}{C_i^{\min}} \leq \Delta$ for some given $\Delta > 1$. Such a situation may arise, e.g., when using a network with *QoS Percentage-Based Policing* [10].

Observation 1. *Any online scheduling strategy for DG-VS which does not introduce idle times in communication is* Δ*-competitive, since no communication can be more than* Δ *times slower than in the optimum schedule.*

Observation 2. *If no additional information is given, no online algorithm* A *consisting in reacting to changing communication speeds can be better than* Δ*-competitive.*

Proof. Let $m = 2$, $\alpha_1 = \alpha_2 = 1$, $C_{1,0} = C_{2,0} = 1$. We can assume without loss of generality that the first sender chosen by algorithm A is P_1. Now, let $t_1 = 1$, $C_{1,1} = \frac{1}{\Delta}$, $C_{2,1} = \Delta$. The schedule length $T = 1 + \Delta$ obtained by algorithm A is Δ times larger than the optimum schedule length $T^* = 1 + \frac{1}{\Delta}$. □

Since by Observations 1 and 2 it is not possible to construct a better than trivial online algorithm without additional knowledge, let us now assume that the network is homogeneous, i.e. $\alpha_i = \alpha$, $C_i^{\min} = C^{\min}$ and $C_i^{\max} = C^{\max}$ for all i.

Theorem 2. *There exists a* $\frac{1+(m-1)\Delta^2}{1+(m-1)\Delta}$*-competitive online algorithm solving problem DG-VS for a homogeneous network.*

Proof. Consider algorithm A that always chooses as the sender the node with the smallest current communication rate. Let $\mathcal{S}$ denote the schedule of length T constructed by A and let $\mathcal{S}^*$ be the optimum schedule of length T^*. Let P_i be the last sender in schedule $\mathcal{S}$. The total length of intervals when P_i transfers data in schedule $\mathcal{S}^*$ will be denoted by T_i^*. Let $T_{other}^* = T^* - T_i^*$. Note that it is possible to send data from P_i in schedule $\mathcal{S}$, whenever P_i sends data in schedule $\mathcal{S}^*$. Hence, in the corresponding time intervals the communication in $\mathcal{S}$ is not slower than in $\mathcal{S}^*$ and the size of sent data is at least α. The remaining data, of size at most $(m-1)\alpha$, are sent in $\mathcal{S}$ in time at most ΔT_{other}^*. Thus,

$$T \leq T_i^* + \Delta T_{other}^*. \tag{5}$$

Furthermore, we have $T_{other}^* \leq (m-1)\Delta T_i^*$, and since $T_i^* + T_{other}^* = T^*$, we get $T_{other}^* \leq \frac{T^*(m-1)\Delta}{1+(m-1)\Delta}$. Hence, we obtain from (5) that $\frac{T}{T^*} \leq \frac{1+(m-1)\Delta^2}{1+(m-1)\Delta}$. □

5 Future research

In this work, we analyzed minimizing data gathering time in a network with variable communication speed. We proposed a polynomial-time offline algorithm solving problem DG-VS and a $\frac{1+(m-1)\Delta^2}{1+(m-1)\Delta}$-competitive polynomial-time online algorithm solving DG-VS in a homogeneous network. Future research may concern the construction of non-deterministic online algorithms for DG-VS.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, Wireless sensor networks: A survey, *Computer Networks*, **38** (2002), 393–422.

[2] J. Berlińska, Communication scheduling in data gathering networks with limited memory, *Applied Mathematics and Computation*, **235** (2014), 530–537.

[3] J. Berlińska, Scheduling for data gathering networks with data compression, *European Journal of Operational Research*, **246** (2015), 744–749.

[4] K. Choi, T. G. Robertazzi, Divisible load scheduling in wireless sensor networks with information utility, *IEEE International Performance Computing and Communications Conference 2008*, IEEE, 2008, pp. 9–17.

[5] S. C. Ergen, P. Varaiya, TDMA scheduling algorithms for wireless sensor networks, *Wireless Networks*, **16** (2010), 985–997.

[6] S. Kumar, S. Chauhan, A survey on scheduling algorithms for wireless sensor networks, *International Journal of Computer Applications*, **20** (2011), 7–13.

[7] M. Moges, T. G. Robertazzi, Wireless sensor networks: scheduling for measurement and data reporting, *IEEE Transactions on Aerospace and Electronic Systems*, **42** (2006), 327–340.

[8] T. G. Robertazzi, Ten reasons to use divisible load theory, *IEEE Computer*, **36** (2003), 63–68.

[9] L. Shi, A. Fapojuwo, TDMA scheduling with optimized energy efficiency and minimum delay in clustered wireless sensor networks, *IEEE Transactions on Mobile Computing*, **9** (2010), 927–940.

[10] T. Szigeti, C. Hattingh, R. Barton, K. Briley, *End-to-End QoS Network Design: Quality of Service for Rich-Media & Cloud Networks*, Cisco Press, 2013.