Efficient Algorithm for the Convergecast Scheduling Problem on a Square Grid with Obstacles

Adil I. Erzin
Sobolev Institute of Mathematics
Novosibirsk State University
Acad. Koptyug Avenue, 4,
630090 Novosibirsk, Russia
adilerzin@math.nsc.ru

Roman V. Plotnikov
Sobolev Institute of Mathematics
Acad. Koptyug Avenue, 4,
630090 Novosibirsk, Russia
prv@math.nsc.ru

Abstract

In the convergecast scheduling problem, it is required to find a spanning aggregation tree with a root in a base station (sink) in a given communication graph and build a conflict-free (half-duplex, interference-aware) min-length schedule for aggregating data over the edges of the aggregation tree. This problem is strongly NP-hard even in the case of a given aggregation tree. However, if the communication graph is a square grid in each node of which there is a sensor and in which per unit time a data packet is transmitted along one edge of the graph, the problem is polynomially solvable. In this paper, we consider a communication graph in the form of a square grid with rectangular obstacles impenetrable by the messages and propose a polynomial algorithm for constructing a conflict-free schedule for data aggregation. An intensive numerical experiment confirmed the hypothesis that the algorithm constructs an optimal solution in 100% of cases.

1 Introduction

In wireless sensor networks, data collected by sensors is regularly transmitted to the base station or *sink* [Chen et al., 2005]. In networks responsible for security, data collection must occur quickly to ensure a timely response of the system. In wireless networks, the transmission of messages is usually carried out by means of radio communication, which imposes certain restrictions [Incel et al., 2012, Malhotra et al., 2011, Souza & Nikolaidis, 2013]. In particular, if sensors share a frequency, then the situation in which more than one transmitter operates in the receiving zone of the receiver causes *interference* of radio waves, and this situation is called a *conflict* or a *collision* [Tian, 2011, Souza & Nikolaidis, 2013, Wang et al., 2013, Xu et al., 2011]. On the other hand, the communication graph of the wireless sensor network is built according to the criterion of minimizing communication energy consumption [Erzin et al., 2017]; therefore, not all sensors can transfer the packet directly to the base station. Most sensors transmit the collected data in transit through other sensors.

If the collected data can be aggregated (max, min or mean, contamination of the terrain, fire, unauthorized entry into the room, etc.), then each node of the network (sensor) first receives data packets from all vertices

that use it as a transit and then aggregates the data and sends the aggregated package further [Incel et al., 2012, Souza & Nikolaidis, 2013]. For energy efficiency considerations, each vertex sends a data packet once during the data aggregation session. This means that the sending of packets is carried out along the edges of the spanning *aggregation tree* (AT) with the root in the base station [Malhotra et al., 2011, Tian, 2011].

In practice, unit disk graphs are often used, so the time required for the packet to pass through different edges of the communication graph differs insignificantly, and the transit time of the packet along any edge can be set equal to 1 [Wang et al., 2013]. In this case, the time takes integer values and is measured in time rounds (slots). The number of time rounds during which data from all sensors will be aggregated in the base station is called the aggregation time or the *length of the schedule*. In most wireless networks, half-duplex communication systems are used, in which each vertex cannot simultaneously receive and transmit a packet. Moreover, the vertex can receive no more than one message during one time round [Chen et al., 2005, Incel et al., 2012, Souza & Nikolaidis, 2013].

A *schedule* is a function that assigns to each vertex of the communication graph a positive integer, the time round, when this vertex sends the packet. We call a schedule *feasible* if it satisfies the following conditions:

- it does not contain conflicts;

- during one time round, every vertex either receives or transmits a packet;

- during one time round, every vertex can receive at most one packet;

- each vertex sends a packet only once during the aggregation session.

In the *convergecast scheduling problem* (CSP), it is required to find an aggregation tree and a feasible min-length schedule for data aggregation.

The CSP is NP-hard [Chen et al., 2005] even in the case of a given aggregation tree [Erzin & Pyatkin, 2016]. If conflicts are considered only among children of common parents, the problem is equivalent to the phone broadcasting problem [Slater et al., 1981], which is NP-hard in the general case but polynomially solvable in the case of a given AT. The problem is polynomially solvable also in the case in which the communication graph is a unit square grid in each node of which there is a sensor and when the transmission distance is 1 [Gagnon & Narayanan, 2015] or 2 [Erzin, 2017].

In this paper, we consider a special case of CSP on a unit square grid when the transmission distance is 1 (as in [Gagnon & Narayanan, 2015]), in which there are rectangular disjoint obstacles impermeable to radio waves. The lexicographic greedy algorithm (LGA), developed to construct a feasible schedule for a given AT, is proposed, and a hypothesis about the existence of an optimal solution that uses either a "horizontal" (HT) or "vertical" (VT) aggregation tree is formulated. Numerical experiments have been carried out that confirm the hypothesis, but we have not yet managed to prove the hypothesis analytically. If the hypothesis is correct, then to construct the optimal solution for CSP, it is enough to construct two shortest path spanning trees – HT and VT – and apply the LGA. Thus, the problem under consideration would be polynomially solvable.

2 Problem Formulation

The communication graph $G = (V, E)$ is given as a unit square grid of dimension $(n + 1) \times (m + 1)$ with a set of disjoint (tangency allowed) rectangular obstacles $\{O_i, i = 1, \ldots, q\}$ with sides parallel to the axes. At each vertex $(x, y) \in V$, $x = 0, 1, \ldots, n$, $y = 0, 1, \ldots, m$, which is not the interior point of some obstacle O_i, except the origin $(0, 0)$, a sensor is placed. In the node $(0, 0)$ there is a sink (base station). Two vertices (x_1, y_1) and (x_2, y_2) are connected by an edge $e \in E$, if the distance between them is $|x_1 - x_2| + |y_1 - y_2| = 1$ and there is no obstacle between them. During one time round, every vertex $i \in V$ can transmit a packet only along one edge $(i, j) \in E$ incident to it. Obviously, the graph G is connected. Fig. 1 shows an example of a grid with forbidden areas, as well as an aggregation tree (red arrows).

The problem considered in the paper consists in constructing in the communication graph G an aggregation tree and finding a feasible min-length schedule for aggregating the data.

3 The LGA

In the LGA, at each step, a set S of vertices is searched, which send messages during the current time round. In this set, the pendant vertices are included in order of decreasing distance from the sink, starting from the most remote one, if there is no conflict. The vertices that transmit the messages are then deleted together with the incident edges, and the procedure is repeated during the next time round.
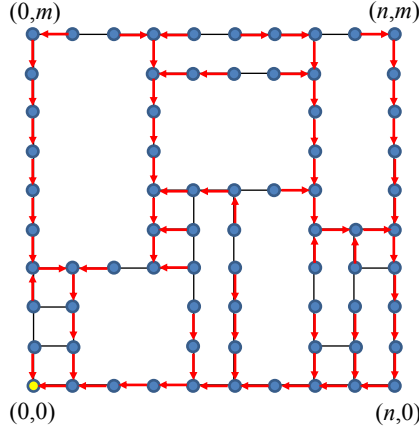
Figure 1: Instance of grid-graph and AT (red arrows).

Let there be given a spanning aggregation tree T of radius R. The pseudocode of the LGA can be written as follows.

Algorithm LGA.
Step 0. Set $t = 0$.
Step 1. Set $t = t + 1$; $k = -1$; $S = \emptyset$.
Step 1.1. Set $k = k + 1$. Find in T the max-cardinality subset of the pendant vertices at a distance $R - k$ from the sink that can send the packets without conflicts taking into account the set of transfers from the nodes in S, and include them in S. If $R - k > 0$, then go to Step 1.1.
Set $S_t = S$; $R = R - 1$. If $R > 0$, then exclude from T the vertices of the set S_t and the edges incident to them, and go to Step 1.
Return S_t, $t = 1, \ldots, L(n, m)$,

where $L(n, m)$ is the length of the schedule, and S_t is the set of vertices sending messages during time round t.

In Step 1.1, it is necessary to find in the current tree the max-cardinality subset of pendant vertices at a distance $R - k$ from the sink that can send the packets without conflicts taking into account the set of transfers from the nodes in S.

This problem can be formulated as follows. Let's construct a graph of conflicts in which a pair of vertices is connected, if vertices cannot transmit simultaneously without conflicts. Then, the problem reduces to constructing the maximal independent set, which is the same problem as maximum clique on the complementary graph and which is NP-hard for arbitrary graph [Garey & Johnson, 1979]. However, if we construct a *shortest path* AT in a square grid, then in the graph of conflicts there will be connected components only in the form of isolated vertices and chains. For the graph of conflicts in this case, the problem of finding the maximal independent set is solved simply with linear complexity.

4   Solution of CSP

It is known that the shortest path tree (SPT) is not always an optimal aggregation tree [Tian, 2011]. This is clearly shown by Fig. 2. If the SPT is used as the aggregation tree (Fig. 2(a)), then the length of the schedule is 4. The minimum length of the schedule is 3 (Fig. 2(b)). Moreover, the length of the schedule on the SPT in Fig. 2(c) is equal to $2k + 1$, which is almost 2 times greater than the minimum length of the schedule equal to $k + 1$ (Fig. 2(d)).

4.1   Horizontal and Vertical Trees

Both the horizontal tree (HT) and vertical tree (VT) are shortest path trees. Because the sink is at the origin, an arbitrary vertex of the grid graph can send the packet along the edges of SPT either to the left or down. If there is an alternative, then in the HT the vertex sends to the left and in the VT it sends down (Fig. 3).
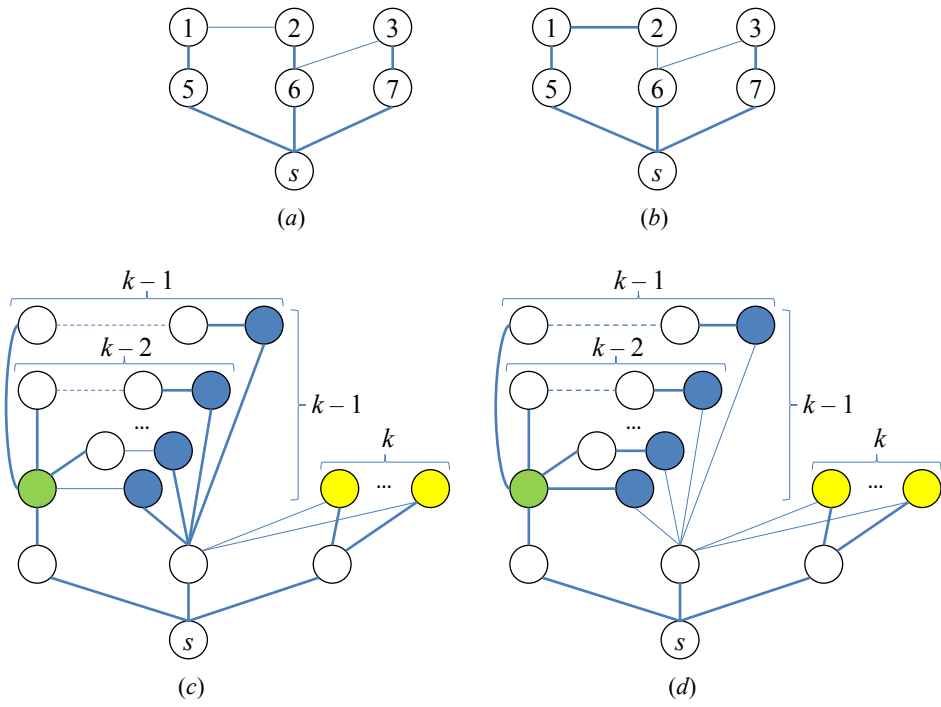
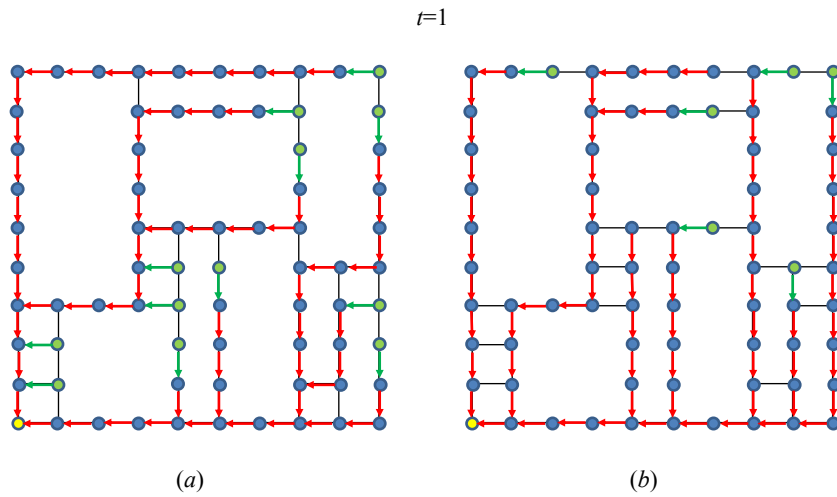Figure 2: (a) SPT, (b) optimal AT, (c) SPT, (d) optimal AT.
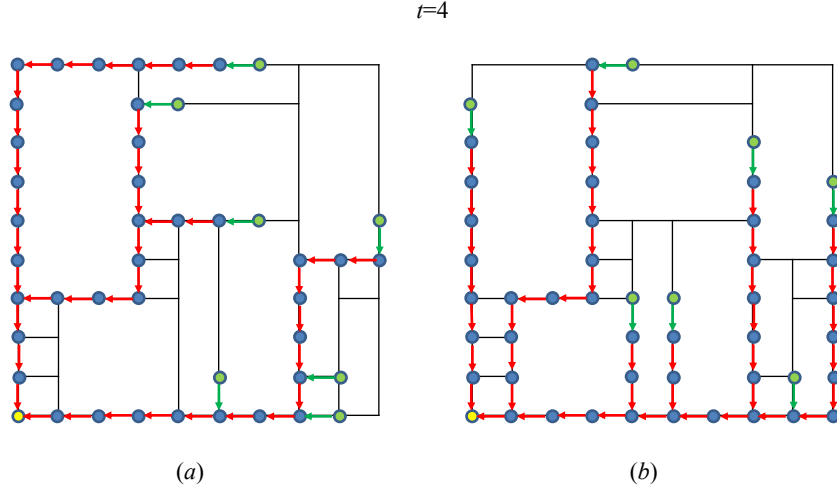

Figure 3: (a) HT, (b) VT.

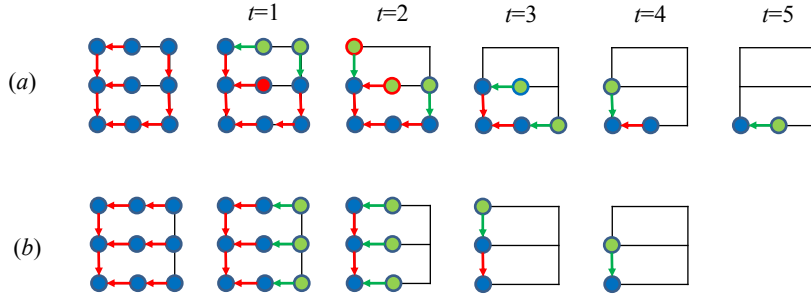Figure 4: Current HT (*a*) and VT (*b*) after the third time round.

Figure 5: Step-by-step work of the LGA for VT (*a*) and HT (*b*).

Apply the LGA to both trees HT and VT. Fig. 3 shows the transfers (green arrows) that are performed during the first time round, and Fig. 4 shows those performed during the fourth time round.

It is easy to show that one of the trees (vertical or horizontal) may not be optimal. Consider the example in Fig. 5.

However, we believe that the following hypothesis is true.

Hypothesis. *In arbitrary $(n + 1) \times (m + 1)$ grid graph with a sink in the origin $(0, 0)$ with any disjoint rectangular obstacles with sides parallel to the coordinate axes, the vertical or horizontal tree is optimal, and the LGA constructs an optimal schedule whose length is equal to the lower bound $n + m$ (the distance to the most remote vertex).*

5   Simulation

We have not yet managed to prove the hypothesis analytically, but we have confirmed it numerically by carrying out numerical experiments. In order to perform these experiments, we implemented the proposed algorithm, the LGA, as well as the construction of HT and VT. We generated 50 random instances for each of the following pairs of n and m: $\{10, 10\}, \{10, 30\}, \{10, 50\}, \{10, 100\}, \{30, 30\}, \{30, 50\}, \{30, 100\}, \{50, 50\}, \{50, 100\}$, and $\{100, 100\}$. In each instance, the length of a schedule constructed by the LGA on either one of the trees (HT and VT) or on both of them was equal to the lower bound $n + m$. This means that our algorithm (LGA on HT and VT) most likely *always* constructs an optimal solution.

6   Conclusion

In this paper, a special case of the problem of minimizing the time of conflict-free data aggregation in a single-channel wireless sensor network, known as the convergecast scheduling problem, is considered. In the case considered, the sensors are located in the nodes of a square grid of dimensions $(n+1) \times (m+1)$ with rectangular disjoint obstacles that are impermeable to radio waves. The transmission range of each sensor is 1, and the base station is at the origin $(0,0)$.

A polynomial algorithm, the LGA, that constructs a feasible schedule for a given aggregation tree is proposed. It is obvious that the length of the schedule depends on the aggregation tree, but it cannot be less than the distance to the most remote vertex of the grid $n + m$. In this paper, in the set of shortest path spanning trees, only two of them are considered: *horizontal* (HT) and *vertical* (VT). We formulated the hypothesis that at least one schedule constructed by the LGA on HT or VT is optimal and that length is $n + m$.

The hypothesis could not be proved analytically, but it was confirmed by a number of randomly generated instances of different dimensions.

References

[Chen et al., 2005] Chen, X., Hu, X. & Zhu, J. (2005) Minimum Data Aggregation Time Problem in Wireless Sensor Networks *X. Jia, J. Wu, and Y. He (Eds.): MSN 2005, LNCS 3794*, 133-142.

[Erzin, 2017] Erzin, A. (2017) Solution of the Convergecast Scheduling Problem on a square unit grid when the transmission range is 2. *LNCS* (to appear)

[Erzin et al., 2017] Erzin, A. I., Mladenovich, N., & Plotnikov, R. V. (2017) Variable neighborhood search variants for Min-power symmetric connectivity problem *Computers & Operations Research, 78*, 557–563. doi:10.1016/j.cor.2016.05.010

[Erzin & Pyatkin, 2016] Erzin, A., & Pyatkin, A. (2016). Convergecast Scheduling Problem in Case of Given Aggregation Tree. The complexity status and some special cases. In *Proceedings of the 11th Int. Symp. On Algorithms and Experiments for Sensor Systems, Wireless Networks and Distributed Robotics.* (Article 16, 6 pages). Prague, Czech Republic: IEEE-Xplore. doi:dx.doi.org/10.1109/CSNDSP.2016.7574007

[Gagnon & Narayanan, 2015] Gagnon, J., & Narayanan, L. (2015). Minimum latency aggregation scheduling in wireless sensor networks. *J. Gao et al. (Eds.): ALGOSENSORS 2014, LNCS, 8847*, 152-168. doi:10.1007/978-3-662-46018-4 10

[Garey & Johnson, 1979] Garey, M. R., & Johnson, D. S. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness.* San Francisco: W. H. Freeman and Co..

[Ghods et al., 2013] Ghods, F., Yousefi, H., Mohammad, A., Hemmatyar, A., & Movaghar, A. (2013) MC-MLAS: Multi-channel minimum latency aggregation scheduling in wireless sensor networks. *Computer Networks, 57*, 3812–3825.

[Incel et al., 2012] Incel, O. D., Ghosh, A., Krishnamachari, B., & Chintalapudi, K. (2012) Fast data collection in tree-based wireless sensor networks. *IEEE Trans. on Mobile Computing. 11(1)*, 86–99.

[Malhotra et al., 2011] Malhotra, B., Nikolaidis, I., & Nascimento, M.A. (2011) Aggregation convergecast scheduling in wireless sensor networks. *Wireless Networks, 17*, 319–335. doi:10.1007/s11276-010-0282-y

[Slater et al., 1981] Slater, P. J., Cockayne, E. J., & Hedetniemi, S. T. (1981) Information dissemination in trees. *SIAM J. Comput., 10(4)*, 692–701.

[Souza & Nikolaidis, 2013] De Souza, J., & Nikolaidis, I. (2013). An exploration of aggregation convergecast scheduling. *Ad Hoc Networks, 11*, 2391–2407. doi:dx.doi.org/10.1016/j.adhoc.2013.06.004

[Tian, 2011] Tian, C. (2011) Neither Shortest Path Nor Dominating Set: Aggregation Scheduling by Greedy Growing Tree in Multihop Wireless Sensor Networks. *IEEE Trans. on Ve- hicular Ttchnolohy, 60(7)*, 3462–3472. doi:10.1109/TVT.2011.2162086

[Wang et al., 2013] Wang, P., He, Y., & Huang, L. (2013) Near optimal scheduling of data aggregation in wireless sensor networks. *Ad Hoc Networks, 11*, 1287–1296. doi:10.1016/j.adhoc.2011.01.003

[Xu et al., 2011] Xu, X., Li, X.-Y., Mao, X., Tang, S., & Wang, S. (2011) A delay-efficient algorithm for data aggregation in multihop wireless sensor networks. *IEEE Transactions on Parallel and Distributed Systems, 22*, 163–175. doi:10.1109/TPDS.2010.80