Computers and Operations Research 98 (2018) 38–55

Contents lists available at ScienceDirect

Computers and Operations Research

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

Reformulations and branch-and-price algorithm for the Minimum Cost Hop-and-root Constrained Forest Problem

Dilson Lucas Pereira[a,*], Alexandre Salles da Cunha[b,1]

[a] Universidade Federal de Lavras, Departamento de Ciência da Computação, Lavras, Brazil
[b] Universidade Federal de Minas Gerais, Departamento de Ciência da Computação, Belo Horizonte, Brazil

ARTICLE INFO

Article history:
Received 23 May 2017
Revised 12 March 2018
Accepted 11 May 2018
Available online 14 May 2018

Keywords:
Row and column generation
Wireless sensor networks
Vehicle routing
Trees
Parallel heuristics

ABSTRACT

The Minimum Cost Hop-and-root Constrained Forest Problem (MCHCFP) is a combinatorial optimization problem that arises in the design of energy efficient wireless sensor networks with mobile sinks. It aims at providing a communication topology that minimizes energy consumption while controlling network latency and message loss due to communication failure. The problem is defined in terms of a mixed graph $G = (V, E, A)$ with set of vertices V (representing the sensor nodes), edges E, arcs A, and parameters $K, H \in \mathbb{N}$, and $D \in \mathbb{R}_+$. Each edge has an associated length, and each arc has an associated cost. The MCHCFP aims at finding a rooted spanning forest of (V, A) with minimum total cost. The solution must be *hop constrained*, in the sense that the path from any leaf to its root in the forest does not have more than H arcs, as well as *distance constrained*, so that one must be able to find K routes of length at most D, visiting exactly once the chosen roots of the forest. We study four integer programming formulations, two of them are new, the other two come from the literature. The first formulation coming from the literature is a multicommodity flow based model. The second is a Dantzig–Wolfe reformulation of the first. The two new formulations, which are equivalent with respect to their linear programming bounds, arise when the problem is formulated over a layered graph. To evaluate the bounds implied by the formulations other than the network flow model, we develop algorithms based on delayed column and/or row generation. We provide computational experiments showing that the best lower bounds are given by the formulations based on the layered graph. The reformulation coming from the literature provides lower bounds that are very close to those provided by the layered graph formulations, but in significantly less computational time. As it provides the best trade-off between lower bound quality and computational effort, that reformulation is thus chosen as the basis for a branch-and-price algorithm introduced here. Such an algorithm managed to solve instances with up to 60 vertices, a significant improvement over the previous approach in the literature, which systematically solves instances with up to 20 vertices.

1. Introduction

In this paper, we present integer programming formulations and algorithms for the Minimum Cost Hop-and-root Constrained Forest Problem (MCHCFP) (Bechelane, 2009; Bechelane et al., 2009), a hard combinatorial optimization problem that appears in the design of Wireless Sensor Networks (WSN) (Akyildiz et al., 2002).

WSNs are a kind of ad hoc network, based on the collaborative work of sensor nodes, small autonomous processing entities. Sensor nodes are usually equipped with a sensing device, a processor with limited memory and computational processing power, a wireless transmitter-receiver equipment, and limited battery-supplied energy. To overcome their hardware limitations and implement complex tasks, sensor nodes operate in a collaborative and distributed fashion. WSNs usually involve another type of nodes, named sinks. Being much less constrained than sensor nodes in terms of energy availability and computational power, sinks play an important role in WSNs. Through wireless communication protocols, they gather sensed information from sensor nodes, process the collected data, and disseminate operational decisions through the network.

Sinks in WSNs may be fixed or mobile. In the former case, the communication topology is usually a rooted tree, where the sink is the root and the sensor nodes are the other vertices. Therefore, sensed information arrives at the sink by means of message for-

---

[*] Corresponding author.
E-mail addresses: dilson.pereira@dcc.ufla.br (D.L. Pereira), acunha@dcc.ufmg.br (A.S. da Cunha).

[1] Alexandre Salles da Cunha was partially funded by FAPEMIG grant CEX - PPM-00164/17 and CNPq grants 303677/2015-5, 471464/2013-9 and 200493/2014-0.

https://doi.org/10.1016/j.cor.2018.05.013

warding. That means sensor nodes implement their sensing tasks and forward messages received from their neighbors in the tree. Since message forwarding is much more energy consuming than sensing (Aioffi et al., 2011; Akyildiz et al., 2002), WSNs with fixed sinks tend to suffer from low lifetime. In addition, message forwarding implies a non-balanced energy consumption among sensor nodes. As a result, some sensors may have their batteries completely depleted much earlier than others, and the network may suffer from premature disconnectedness.

Mobile sinks are often used aiming to extend the lifetime in WSNs, since they allow a more balanced energy consumption among the sensors. Sinks, drones or other types of vehicles, are allowed to visit some or all sensors and to collect information from them, through radio communication. Message forwarding is reduced and so is energy consumption. A positive side effect of sink mobility is the reduction on the average number of hops in the path between a sensor node and the sink. The probability of losing messages due to transmission failure is also likely to decrease. One of the major drawbacks is the increase in network latency, i.e., the difference between the time the information was collected from the environment and the time it arrived at the sink.

The MCHCFP was introduced attempting to deal with these conflicting quality of service parameters in the design of WSNs. It aims at minimizing the total energy consumption in the network and tries to keep message loss and network latency at low levels, by imposing certain design constraints. The network topology involves $K \in \mathbb{N}$ mobile sinks. Sensor nodes are organized in a rooted H-hop-constrained forest, i.e., the path connecting a leaf in a tree and its root does not involve more than $H \in \mathbb{N}$ hops. The roots, called cluster heads, are the only sensor nodes that establish a direct communication with the sinks. Other nodes are not allowed to do so. By message forwarding, sensed information is sent from sensor nodes to their parent nodes, until it arrives at the root of the tree. Given a constant probability ρ of transmission failure between a pair of sensor nodes, the value H is chosen such that the probability $1-(1-\rho)^H$ of losing a message in a H-hop constrained path ending at the root falls below a certain design threshold. Sinks are initially placed at a central base station or depot, where they recharge their batteries and information consolidation takes place. Each sink visits a set of cluster heads, collects sensed data and returns to the depot. In order to keep network latency below an acceptable threshold, the length of the route implemented by each sink must not exceed a certain design parameter $D \in \mathbb{R}_+$.

The problem can be described in a graph theoretic setting as follows. Let $G = (V, E, A)$ be a mixed graph with a set of vertices V $(n = |V|)$, a complete set of edges E, and a set of arcs A. Sensor nodes are represented by the vertices in V. One particular vertex in that set, denoted by 1, represents a geographical location where a depot or base station lies. The directed graph $G_C = (V, A)$ is referred to as the *communication network* and is used to model the H-hop constrained forest. Weights $\{c_{ij} \in \mathbb{R}_+ : (i, j) \in A\}$ are assigned to the arcs of the communication network. The cost c_{ij} represents an estimate of the energy consumed by a sensor node i when it sends a message (a package of sensed data) to another sensor node j. The undirected graph $G_T = (V, E)$ represents the *translation network*, being used to model the translation of the sinks through the sensed area. Distances $\{d_{ij} \in \mathbb{R}_+ : \{i, j\} \in E\}$ are assigned to these edges.

The MCHCFP is thus a combinatorial optimization problem aiming at finding an H-hop constrained spanning forest in the communication network along with K routes in the translation network. Each sink implements one route and visits only the roots of the trees assigned to that route. Choosing the roots is part of the problem so that each tree in the forest is H-hop constrained and the length of each route is at most D. Each sensor node is assigned to one tree in the forest and each tree's root is visited exactly once by one sink. The exception being the depot, the location where each sink starts and ends its route. The MCHCFP then asks for an H-hop constrained forest with minimum total weight (the sum of the weights of the arcs in the forest) such that exactly K distance constrained routes can be found.

1.1. Literature review

The MCHCFP was first introduced in Bechelane et al. (2009) and was also addressed in Bechelane (2009). The MCHCFP version described here and that in Bechelane (2009) are precisely the same and generalize that in Bechelane et al. (2009). The version studied in Bechelane et al. (2009) considers just one mobile sink, while one or more sinks may be used here. Besides that, instead of a distance constrained route starting and ending at the same depot, the version in Bechelane et al. (2009) looks for distance constrained paths. All the other aspects are shared by both versions.

Solving the MCHCFP requires the integrated resolution of three difficult optimization problems: a location problem (Drezner and Hamacher, 2001), to assign sensor nodes to cluster heads, a vehicle routing problem (VRP) (Toth and Vigo, 2002), to define the routes visiting the cluster heads, and a minimum cost hop constrained spanning forest problem (a special case of the Hop Constrained Minimum Spanning Tree Problem (HMSTP) (Dahl et al., 2006)) to define the communication network. The MCHCFP is thus NP-hard, as discussed in Bechelane (2009); Bechelane et al. (2009).

Two MCHCFP algorithms were proposed in Bechelane et al. (2009). The first is a Branch-and-bound (BB) method based on a compact network flow formulation. The other is a heuristic that involves a construction phase and several local search procedures, aiming to reduce the cost of the forest and the length of the paths. Later in Bechelane (2009), the formulation and algorithms proposed in Bechelane et al. (2009) were extended to the multiple sink case addressed here. Additionally, the heuristics were embedded in a network simulator whose results indicated that the proposed topology significantly improved on network lifetime and coverage, when the same bounds for the probability of losing messages and network latency were imposed to the MCHCFP and to other design strategies in the literature. The BB algorithm in Bechelane (2009) managed to solve some instances with up to 40 vertices in a couple of CPU hours, but failed, at publication time, to solve larger ones within a time limit of 4 CPU hours.

Pereira et al. (2010) introduced two MCHCFP heuristics based on several local search procedures. Making use of parallel programming, those heuristics were capable of providing better results than those obtained by the heuristics in Bechelane (2009); Bechelane et al. (2009), in terms of solution quality and computing times.

A problem closely related to the MCHCFP was later investigated in Romão and Santos (2013); Romão et al. (2013). It is similar to the MCHCFP since the same network topology is sought, but considers only one sink and minimizes a different objective function. In the WSN application where the MCHCFP typically arises, non-leaf nodes in the communication forest have to transmit their own data and also data coming from their descendant nodes. Our model assumes that each node forwards a package of fixed size to its parent node and that the package is large enough to fit all the data collected by the node and by its descendants. A different modeling choice is made in Romão and Santos (2013); Romão et al. (2013). It is assumed that the number and the sizes of these packages are not fixed, thus, energy consumption depends on the sizes of the subtrees rooted at each node in the communication forest. In order to model this different energy consumption model, a quadratic objective function was initially formulated. The linear integer programming formulation in Romão and San-