Computers & Operations Research 46 (2014) 12–22

Contents lists available at ScienceDirect

Computers & Operations Research

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

Locating sensors to observe network arc flows: Exact and heuristic approaches

L. Bianco [a], C. Cerrone [b], R. Cerulli [b], M. Gentili [b,*]

[a] Dipartimento di Ingegneria dell'Impresa, Università di Roma "Tor Vergata", Via del Politecnico 1, 00133 Roma, Italy
[b] Dipartimento di Matematica, Università di Salerno, Via Giovanni Paolo II n. 132 , 84084 Fisciano, SA, Italy

ARTICLE INFO

Available online 31 December 2013

Keywords:
Sensor location
Traffic network
Branch and bound
Genetic algorithm

ABSTRACT

The problem of *optimally* locating sensors on a traffic network to monitor flows has been an object of growing interest in the past few years, due to its relevance in the field of traffic management and control. Sensors are often located in a network in order to observe and record traffic flows on arcs and/or nodes. Given traffic levels on arcs within the range or covered by the sensors, traffic levels on unobserved portions of a network can then be computed. In this paper, the problem of identifying a sensor configuration of minimal size that would permit traffic on any unobserved arcs to be exactly inferred is discussed. The problem being addressed, which is referred to in the literature as the Sensor Location Problem (SLP), is known to be NP-complete, and the existing studies about the problem analyze some polynomial cases and present local search heuristics to solve it. In this paper we further extend the study of the problem by providing a mathematical formulation that up to now has been still missing in the literature and present an exact branch and bound approach, based on a binary branching rule, that embeds the existing heuristics to obtain bounds on the solution value. Moreover, we apply a genetic approach to find good quality solutions. Extended computational results show the effectiveness of the proposed approaches in solving medium-large instances.

1. Introduction

Monitoring flows on the network is an important topic in the field of traffic management and control. The continuous growth in the demand for private transportation in large urban areas is the cause of severe congestion, pollution, time loss in traffic jams and a deterioration in the quality of life. Monitoring flows on a network allows traffic managers to control and manage these problematic situations. Even though communication technologies for monitoring traffic networks in real time, via sensors and video cameras, are currently available, in most cases we have a very large network which is monitored only in small parts. In this context, partial information on traffic flows (obtained, for example, by located sensors on the network), is often used to estimate flows in the network that are not directly observed. However, recent studies in the literature show that by properly locating sensors on the network it is possible to exactly compute, under the assumption of error free data, the entire set of unobserved flows. In this context it is of paramount importance to design optimal strategies to determine sensor locations, and the problem of locating sensors on the network has been the object of growing interest in the past few years. Problems in this class are differentiated according to the type of sensors that need to be located (counting sensors, path-ID sensors, vehicle-ID sensors or a combination of them) and flows of interest (origin/destination flow volumes, arc flow volumes, route flow volumes, or a combination of them). *Counting sensors* can be considered as all those types of sensors that are able to count vehicle on a lane(s) of a road (for example conventional inductive loop sensors). They can be located on an arc (a vertex) of the network and count the number of vehicles on the arc (on the arcs incident to the vertex) during a particular time interval. *Path-ID sensors* are assumed to be devices that, when located on an arc of a network, can measure the flow volume of each route to which that arc belongs. This is the class of sensors that de-code active transmission provided by tagged vehicles, for example, freight information from trucks or path/schedule information from buses. *Vehicle-ID sensors* are sensors through which a vehicle can be univocally identified while traveling on the network. Licence plate readers or Automatic Vehicle Identification (AVI) readers are examples of sensors that belong to this class.

Following the classification defined in the recent surveys by Gentili and Mirchandani [14,15], two classes of problems can be identified: (i) locating sensors to fully observe flow values (either arc flows, route flows or OD flows) on the network

* Corresponding author. Tel.: +39 089 963444; fax: +39 089 96 3303.
E-mail addresses: bianco@disp.uniroma2.it (L. Bianco), carminecerrone@gmail.it (C. Cerrone), raffaele@unisa.it (R. Cerulli), mgentili@unisa.it (M. Gentili).

0305-0548/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cor.2013.12.013

(Sensor Location Flow-Observability Problems), and (ii) locating sensors to estimate flows (either arc flows, route flows or OD flows) on the network (Sensor Location Flow-Estimation Problems). This dichotomy is derived from the observation that the location of sensors on a network (either on the vertices or on the arcs) can be translated into a system of linear equations where the set of variables corresponds to the unknown flows and the set of equations comes from the deployed sensors. When the resulting system has a unique solution, we say the system is fully observable, and therefore all the flows involved in the system are known (that is, they are *observable*). The resulting location problem consists of determining the optimum deployment of sensors on the network that results in an observable system (Sensor Location Flow-Observability Problems). On the other hand, when the system is underspecified, it admits an infinite number of solutions. The related location problem consists of determining how to optimally deploy sensors on the network so that the derived flow estimates are as good as possible. Generally, underspecified systems arise when one is interested in determining origin-destination flow volumes by locating counting sensors on the arcs of the network. This problem has been extensively studied in the literature (see for example, Chootinan et al. [10], Elhert et al. [12], Yang et al. [18], Kim et al. [19], Lam and Lo [20], and Yang and Zhou [23]). On the other hand, observability problems arise, for example, when either path-ID sensors or vehicle-ID sensors need to be located on the arcs of the network to determine route flow volumes (Gentili and Mirchandani [15,16], Castillo et al. [6–8].

In this paper we focus on the observability problem arising when counting sensors are located on the vertices of the network and arc flow volumes need to be computed. Specifically, we are interested in locating the minimum number of counting sensors on the vertices of a network to compute arc flow volumes on the whole network (we refer to this problem as the *Sensor Location Problem* [*SLP*]).

There is an extensive literature related to facility location on networks and, in particular, to the flow interception problems. The problem addressed in this paper could seem similar to this well known class of problems; however, there is a huge difference. In flow interception problems a set of facilities is to be located on the network (generally on the arcs of the network) to intercept flows such that a given function of the flow is optimized (e.g., total intercepted flow is maximized [17] or total risk reduction is maximized [13]). Any subset of arcs (or vertices) that is selected is feasible for the problem. The only similarity between the SLP and this class of problems is the fact that once a facility (a sensor) is located on the network, a flow is intercepted. The main issue that makes the SLP unique and different from the flow interception problems is that the location of the facility has to be such that from the directly intercepted flows all the non-directly intercepted flows on the remaining arcs of the network can be computed. Hence the network becomes fully observable in terms of arc flows.

The Sensor Location Problem was formally stated by Bianco et al. [5], where two heuristics giving lower and upper bounds on the solution value were presented, and a necessary condition for feasibility was stated. Successively, Bianco et al. [4] developed a combinatorial analysis of the problem and studied its computational complexity considering different special cases. Moreover, some graph classes, where the problem is polynomially solvable, were also presented. Morrison and Martonosi [21] and Morrison et al. [22] addressed the problem on bi-directed trees, giving a necessary and sufficient condition for a subset of vertices to be a feasible solution of the problem. They also defined a matrix reduction procedure to test feasibility of a subset of vertices. Confessore et al. [11] further studied the problem and presented new heuristic algorithms and approximation algorithms. The quality of the solutions provided by the existing algorithms is evaluated by comparison with data-dependent approximation bounds. Indeed, for the problem being considered there does not exist any exact approach providing (at least for small instances) the exact solution value, nor does there exist a mathematical formulation.

In this paper we further extend the study of the problem by (i) developing and testing an exact approach, based on a branch and bound scheme, to optimally solve the problem, (ii) developing and testing a genetic solution algorithm to get near-optimal solutions and (iii) providing a mathematical formulation of the problem based on the concept of MB-paths introduced in [4]. The mathematical formulation provided is a flow based formulation whose optimal solution provides a lower bound on the optimum solution value of SLP. We point out here that this is the first attempt to formulate this problem whose mathematical model was still missing in the literature. We tested our approaches on the set of benchmark instances existing in the literature, on new instances we developed to extend the test cases for the problem and on a real world network. Our results show the efficiency of the proposed approaches in solving medium-large instances.

The paper is organized as follows. Section 2 describes the problem being addressed and details the results in the existing literature that will be used in the present study. Our mathematical formulation is given in Section 3. Section 4.1 describes our genetic approach, while our Branch and Bound algorithm is described in Section 4.2. Computational results are discussed in Sections 5 and 6. Conclusions and further research are the subject of Section 7.

2. Problem description and existing results

In this section we describe in detail the problem being addressed. For this entire discussion we assume that data are error-free. We represent a network by means of a directed graph $G=(N,A)$, where N is the set of n vertices and A is the set of m arcs, the flow on each arc contains subflows that are generated and/or absorbed from different origin/destination pairs. In the discussion to follow the term network and the term graph are used interchangeably. Among the set of vertices, we say v is an *Origin/Destination (OD)* vertex if the flow is generated and/or absorbed by it. The set of OD vertices of G is denoted by $B \subseteq N$. An OD vertex can be either the origin or the destination of flows or both, and it can also be used as a transfer node. Each OD vertex can send flow to or receive flow from any other OD vertex.

For each vertex v that is not an OD one, the flow conservation constraints hold

$$\sum_{w \in FS(v)} f_{v,w} - \sum_{w \in BS(v)} f_{w,v} = 0 \tag{1}$$

where $FS(v)$ and $BS(v)$ are the outgoing and incoming arcs of vertex v, respectively, and $f_{v,w}$ is the flow volume on arc (v,w).

For each OD vertex $v \in B$, we have the following flow conservation constraint:

$$\sum_{w \in FS(v)} f_{v,w} - \sum_{w \in BS(v)} f_{w,v} = S_v \tag{2}$$

where $S_v \neq 0$ is the *balancing flow* at v, that is, a source or a sink flow so that (2) holds. If split ratios[1] associated with the arcs of the network are also known we could define additional relationship involving the flows. Indeed, by using split ratios, we can express the total outgoing flow $F(v)$ as a function of the flow volume of any outgoing arc. Formally, for each $v \in N$ and each outgoing arc (v,w),

[1] The split ratio associated with the outgoing arc (v,w) specifies the fraction $0 \le p_{v,w} \le 1$ of the outgoing flow $F(v)$ that leaves vertex v through the arc (v,w).