Locating Image Sensors on Traffic Networks

Monica Gentili* Pitu Mirchandani†

*Computing Science Department
University of Salerno
P.te Don Melillo, 84084, Fisciano (Sa), Italy
mgentili@unisa.it

†System and Industrial Engineering Department
ATLAS Research Center, The University of Arizona
Tucson, AZ 85721-0020 USA
pitu@sie.arizona.edu

1 Introduction

The problem of locating sensors to detect traffic flow volumes, or other information such as speed, turning ratios, route flow volumes related to them, has relevance in traffic management and control. We can distinguish three kind of sensors: (i) **counting** sensors that count vehicles in order to give traffic volume on a lane or road of a network (e.g., the classical loop inductance detectors); (ii) **image** sensors that provide an "image" of vehicle/traffic flow (using a fixed camera mounted on a pole or a tall building); (iii) **path-ID** sensors (readers) that de-code transmission from vehicles to obtain, for example, freight information from trucks (good carried, its origin, its destination, container weight, custom clearance, fees paid, etc.), route/schedule information from buses (route number, schedule number, passenger count, etc.) and account information from electronic toll tags (toll paid, credit remaining, vehicle ID, etc.).

In this context several location problems arise on where to locate different kinds of sensors to monitor or manage the particular classes of traffic detected. Some problems on locating sensors on a network addressed in the literature are:

(i) the estimation of total flow volumes from all origins to all destinations (O/D matrix estimation) (e.g., [10], [11], [12]);

(ii) the estimation of flow volumes on the non-monitored links of the network (e.g., [1], [2], [3], [6]);

(iii) the estimation of flows on routes from origins to destinations (e.g., [6], [7]).

In this paper we address the problem of locating image sensors on the nodes of a network, where these sensors provide turning ratios at nodes. We focus on the development of two generic locational decision models: (1) "How many and where should image sensors be located to obtain sufficient information on flow volumes on paths?", and (2) "Given that the traffic management planners have already located count detectors (*counting sensors*) on some network arcs, how many and where should image sensors be located to get the maximum information on flow volumes on path?".

We formally state the problem, in the sequel referred to as *IMAGES*. The analysis is developed, as in [7], by considering different problem scenarios: (i) when no counting sensor is already located on the network (*IMAGES-zero*) and (ii) when there are *some* counting sensors located on the network (*IMAGES-par*). We show these problem scenarios are special case of two new combinatorial problems, respectively, the *Full Rank Submatrix with Few Colors Problem* and the *Full Rank Submatrix Extension with Few Colors Problem*. Complexity analysis is developed for these new problems that are proved to be NP-complete. Special instances are presented for which there are polynomial algorithms to find an optimal solution.
In this paper we give a detailed description of both *IMAGES-zero* and *IMAGES-par*. We then describe the relation between *IMAGES-zero* and the *Full Rank Submatrix with Few Colors Problem* and attendant complexity results are provided.

2 Locating Image sensors on nodes

Video vehicle detection system provide non-intrusive vehicle detection through machine vision. Such systems allow users to place virtual detectors in the field of view, rather than physically placing the detectors on the roadway pavement, providing flexible detector placement. An image can be obtained from either (i) a fixed camera mounted on a tall building or a pole and (ii) a moving camera installed on an air-borne platform such as helicopter. By processing the images obtained by fixed or mobile cameras, it is possible to recognize vehicles on the scene and movement of these vehicles. For example, by locating a fixed video camera that takes images of the traffic movement at an intersection of the network we can estimate the turning ratios at the intersection. For instance, locating an image sensors on node 11 of the network in Figure 1, we can measure the turning ratios at this node. More specifically, we are able to measure the proportion of flow that goes to the outgoing arcs a_{16} and a_{17} from each incoming arc a_{15} and a_{19}.

Notations
Let $\Gamma = (N, A)$ be a graph representing the traffic network, where the set N of nodes has size $|N| = n$ and the set of arcs A has size $|A| = m$. A path $Y = \{a_1, a_2, \ldots, a_s\}$ is a sequence of arcs $a_i \in A$ such that $a_i = (v, w)$ and $a_{i+1} = (w, z)$, $\forall i = 1, \ldots, s-1$. Since flow on an arc contains flow from several paths, from different origin-destination pairs, we need to define total flow in terms of path flows. We will simply let this total flow be decomposed into path flows y_i on path Y_i, $i = 1, \ldots, p$, where p is the total number of paths used in the network. We will let f_a be the total flow on arc $a \in A$ for the time interval being considered (note that a counting sensor on arc a measures f_a). Let $B = \{b_{ij}\}$, $i \in \{1, 2, \ldots, m\}$, $j \in \{1, 2, \ldots, p\}$ be the $m \times p$ *arc/paths incidence matrix*, that is $b_{ij} = 1$ if arc a_i belongs to the path Y_j and $b_{ij} = 0$ otherwise. For example, the i-th column of matrix B, $B^i = [0, 1, 1, 0, 1]$ denotes a path with arcs a_2, a_3, a_5 in a five-arc network consisting of arcs $A = \{a_1, a_2, a_3, a_4, a_5\}$. Our aim

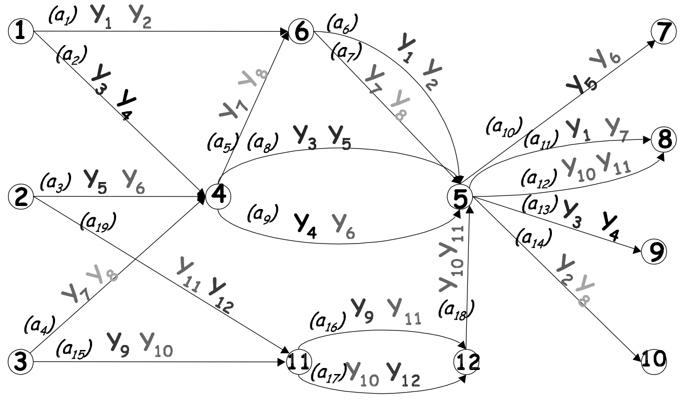

Figure 1: *A network of 12 nodes and 19 arcs.*

is to determine the flow volume y_j of each path Y_j, $j = 1, \ldots, p$, on the network by locating image sensors on the nodes of the network.

Figure 1 shows a network of $n = 12$ nodes and $m = 19$ arcs, where there are $p = 12$ paths. With each arc $a_i \in A$ the set of paths $\mathcal{Y}_{a_i}$ is associated. For example with arc a_8, the set $\mathcal{Y}_{a_8} = \{Y_3, Y_5\}$ denotes the set of paths that contains arc a_8. Let f_{a_i} $i = 1, \ldots 19$ the total flow volume on arc a_i. We can associate with each arc a linear equation which gives the flow volumes of the arc as sum of flow volumes on paths. For example with arc a_8, we can associate the equation $y_3 + y_5 = f_{a_8}$.
Let us suppose counting sensors are located on all the arcs of the network. In this case, we know the entire vector $f = \{f_1, f_2, \ldots, f_m\}$ of the arc flow volumes. To know the path flow volumes we should solve the system of linear equations

$$By = f \tag{1}$$

where y and f are column vectors of p and m components respectively. The general, non-trivial case for system (1) is that $rank(B) = k$ and $k < p$. Thus, if a unique feasible set of path flow volumes exists (that is, all data are consistent), new sensors need to be located to determine path flows. We assume from here on, $min(m, p) = p$ and $rank(B) = k < p$. Under this assumption, the system (1) does not have a unique solution. This means that even locating counting sensors on all arcs of the network we are not able to determine univocally the flow volume of each path Y_j, $j = 1, \ldots, p$. On the other hand, by adding an appropriate set of new equations we may obtain a new matrix that has full rank, and thus a new system having a unique solution.

Indeed, by locating image sensors on a node on the network, we can add new linear equations to system (1). In particular, locating a fixed camera that takes images of the traffic flows at an intersection of the network we can estimate (i) the link flow volume of each arc incident to the node and (ii) the turning ratios at the intersection. See, for example, the simple network in Figure 2 where there are three paths on each arc. By locating an image sensors on node 3, (i) we obtain the following four equations associated with the arcs incident to the node (which

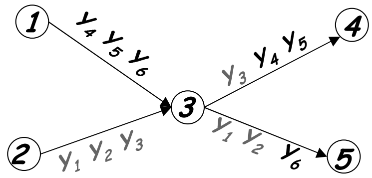

Figure 2: *An image sensor located at node 3 adds 7 equations.*

we can call *external equations*):

$$
\begin{aligned}
Y_4 + Y_5 + Y_6 &= f_{(1,3)} & (2)\\
Y_1 + Y_2 + Y_3 &= f_{(2,3)} & (3)\\
Y_3 + Y_4 + Y_5 &= f_{(3,4)} & (4)\\
Y_1 + Y_2 + Y_6 &= f_{(3,5)} & (5)
\end{aligned}
$$

and also (ii) we detect the turning ratios $t^3_{(1,4)}, t^3_{(1,5)}, t^3_{(2,4)}, t^3_{(2,5)}$. If we know the set of the paths and the arc flow volumes we can define the following four additional equations (which we can call *internal equations*):

$$
\begin{aligned}
Y_4 + Y_5 &= f_{(1,3)} t^3_{(1,4)} & (6)\\
Y_6 &= f_{(1,3)} t^3_{(1,5)} & (7)\\
Y_3 &= f_{(2,3)} t^3_{(2,4)} & (8)\\
Y_1 + Y_2 &= f_{(2,3)} t^3_{(2,5)} & (9)
\end{aligned}
$$

Then, our aim is to answer the following question:

Question 1 *(IMAGES-zero) What is the minimum number of image sensors to locate on the network, and where to locate them in order to add new equations to system (1) that result in a new system having full rank (i.e a unique solution)?*

In *IMAGES-zero*, we do not have any counting sensor. This may not be true in all scenarios. Indeed, in most of the practical applications we might know flow volumes on *some* arcs of the network where counting sensors are already located. In such cases, we suppose that a subset $f^1 = \{a_{i_1}, a_{i_2}, \ldots, a_{i_k}\}$, $i_k < m$, of arc flow volumes, corresponding to the arc subset $A^1 \subseteq A$, are known, resulting in the system:

$$B^1 y = f^1 \qquad (10)$$

where B^1 is a submatrix of B obtained by considering the set of equations associated with the arcs in A^1. The question in these cases is:

Question 2 *(IMAGES-par) What is the minimum number of image sensors, in addition to counting sensors, to locate on the network and where to locate them in order to add new equations to system (10) that result in the new system having a unique solution?*

First, there is an issue on feasibility of the problem. There can be cases where even locating image sensors on *all* the nodes of the network we do not have enough linearly independent equations to determine all path flows (i.e. when the resulting matrix does not have full rank). A sufficient condition for the feasibility of the problem is that the matrix obtained, including all the equations of each node of the network, together with the set of initial equations, has rank equal to p (i.e. the number of paths). Let q_0 be the number of initial equations corresponding to the counting sensors located on the arcs of the network. Let q_i be the number of equations associated with node v_i and $q = \sum_{i=0,1}^{n} q_i$. Let P be the $(q \times p)$ coefficient matrix obtained including the initial set of q_0 equations and all of the equations associated with each node of the network (in the sequel, we refer to this matrix as the *coefficient path matrix*).

Remark 1 *IMAGES is feasible iff* $rank(P) = p$.

3 Complexity Analysis and Mathematical Formulation

Now, we introduce the new problem *Full Rank Submatrix with Few Colors Problem (P1)* and show the relationship between the decision version of *IMAGES-zero* and $P1$. Subsequently, we give the mathematical formulation for $P1$.

The Full Rank Submatrix with Few Colors Problem (P1)
Let L be a 0-1 $(q \times p)$ matrix $(q > p)$ with $rank(L) = p$, $C = \{1, 2, \ldots, n\}$ be a set of colors, and, R be the set of rows of matrix L. Let $c(r) = i$ denote the color assigned to row $r \in R$ of matrix L and $c(S) = \bigcup_{r \in S} c(r)$ be the set of different colors in the subset $S \subseteq R$ of rows. Finally, let K be a positive integer.
Does there exist a set $S \subseteq R$ of linearly independent rows of matrix L such that (i) the corresponding submatrix L_S has $rank(L_S) = rank(L)$ and (ii) $|c(S)| \leq K$ (i.e. the number of different colors assigned to the rows in S is less than or equal to K)?

A subset $S \subseteq R$ of rows of matrix L is feasible for $P1$ it it satisfies condition (i).

When matrix L is the coefficient path matrix of a set of paths on a network, problem $P1$ solves *IMAGES-zero*. Let $\Gamma = (N, A)$ be a traffic network, $|N| = n$, p be the number of paths defined on Γ, and let P be the coefficient paths matrix. Now we assign to each row of P a color and show the relationship between feasible solution of $P1$ defined on P with such a coloring assignment and feasible solution of *IMAGES-zero*.
Each node $v_i \in N$ is associated with the set of rows S_i of matrix P. Let $C = \{1, 2, \ldots, n\}$ be a set of colors. Assign color i to node $v_i \in N$ and to all the rows in S_i associated with v_i. Feasibility of *IMAGES-zero* implies, by Remark 1, $rank(P) = p$. For each subset S of rows of matrix P the number $|c(S)|$ of different colors assigned to rows in S is equal to the number of nodes where we must locate image sensors in order to get the set S of equations. That is, any subset S of rows of matrix P that is feasible for $P1$ and such that $|c(R)| \leq K$, defined on the coefficient path matrix P with the node coloring assignment, corresponds to a set $M \subseteq N$ of nodes with $|M| \leq k$. Since $rank(P) = p$, then the subset M is such that we can determine the flow volumes of all paths p defined on the network Γ.

Related Literature
It can be shown that $P1$ consists in finding a subset $Q \subseteq R$ that minimizes the rank function of a partition matroid M_1 and contains a basis of matroid M_2, both of them being defined on the row set of matrix P. The minimization of a submodular function is a polynomial problem (for a recent survey see for example [5]). However the constraint minimization of a submodular function is in general a NP-complete problem. Polynomial cases are found by restricting the family of sets over which minimizing the submodular function: Grötschel, Lovász and Schrijver [9], Goemans and Ramakrishnan in [8].

Complexity
To prove NP-completeness of $P1$ we consider the special case when each row of matrix L has exactly two non-zero elements. For this particular instance the problem is to look for a spanning *L-forest*[1] of an edge-colored graph with the minimum number of colors. Moreover, polynomial instances are analyzed by considering two main characteristics of the problem:

- the number of rows of matrix L that have the same color (i.e., the size of the sets $R_i = \{r \in R : c(r) = i\}$, for each $i \in \{1, 2, \dots, n\}$);
- the number of non-zero elements in each row of the matrix.

Table 1 summarizes some of our results for $P1$. Each cell of the table defines an instance of $P1$ characterized by (i) the number of non-zero elements in each row (column *NZ*) and (ii) the number of rows of L having the same assigned color i (i.e., columns $R_i = 1$, $R_i = 2$, $R_i > 2$, for each $i \in \{1, 2, \dots, n\}$). In each cell, P indicates that the problem instance is polynomially solvable, while NP means it is NP-complete.

Table 1: Problem $P1$. Complexity Analysis

NZ	$\lvert R_i\rvert = 1$	$\lvert R_i\rvert = 2$	$\lvert R_i\rvert > 2$
1	P	P	?
2	P	?	NP
>2	P	?	NP

The paper will develop the results summarized in the Table 1. Furthermore, it will provide a heuristic when the problem is NP-complete.

References

[1] L.Bianco, G.Confessore, P.Reverberi, "A network based model for traffic sensor location with implications on O/D matrix estimates", *Transportation Science* 35, 1, 50-60 (2001).

[1] An *L-forest* of an undirected graph G is obtained from any forest F of G, by adding at most one edge, to any component of F, that introduces an odd cycle (see [4]).

[2] L.Bianco, G.Confessore, M. Gentili, "Combinatorial Aspects of the Sensor Location Problem", to appear on *Annals of Operations Research.*

[3] G. Confessore, P. Dell'Olmo, M. Gentili, "Experimental Evaluation of Approximation and Heuristic Algorithms for the Dominating Paths Problem", to appear on *Computer and Operations Research.*

[4] M. Conforti, M.R. Rao, "Some new matroids on graphs: cut sets and the max cut problem", *Mathematics of Operations Research* 12, 2, 193-204 (1987).

[5] L. Fleischer, "Recent Progress in Submodular Function Minimization", *Optima* 64, 1-11 (2000).

[6] M. Gentili, New Models and Algorithms for the Location of Sensors on Traffic Networks, Ph.D thesis, Department of Statistic, Probability and Applied Statistics, University of Rome "La Sapienza", (2002).

[7] M. Gentili, P. B. Mirchandani, "Location of Active Sensor on Traffic Network", to appear on *Annals of Operations Research.*

[8] M.X.Goemans, V.S.Ramakrishnan, "Minimizing Submodular Set Functions over Families of Sets", *Combinatorica* 15, 499-513 (1995).

[9] M. Grötschel, L. Lovász, A. Schrijver, "Geometric Algorithms and Combinatorial Optimization", Springer-Verlag (1988).

[10] W.H.K. Lam, H.P. Lo, "Accuracy of O-D estimates from traffic counts", *Traffic Engineering and Control* 31, 358-367 (1990).

[11] H.Yang, Y.Iida, T.Sasaki, "An analysis of the reliability of an Origin/Destination trip matrix estimated from traffic counts", *Transportation Research B* 25, 351-363 (1991).

[12] H. Yang, J. Zhou, "Optimal Traffic Counting locations for Origin-destination matrix estimation", *Transportation Research* 32B, 2, 109-126 (1998).