Allocation of Link Flow Detectors for Origin-Destination Matrix Estimation-A Comparative Study

Torbjörn Larsson, Jan Lundgren and Anders Peterson

Original Publication:

Torbjörn Larsson, Jan Lundgren and Anders Peterson, Allocation of Link Flow Detectors for Origin-Destination Matrix Estimation-A Comparative Study, 2010, COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING, (25), 2, 116-131.
http://dx.doi.org/10.1111/j.1467-8667.2009.00625.x

Allocation of Link Flow Detectors for Origin–Destination Matrix Estimation — a Comparative Study

Torbjörn Larsson
Linköping University
Dept of Mathematics
SE-581 83 Linköping, Sweden
tolar@mai.liu.se

Jan T Lundgren
Linköping University
Dept of Science and Technology
SE-601 74 Norrköping, Sweden
janlu@itn.liu.se

Anders Peterson
Linköping University
Dept of Science and Technology
SE-601 74 Norrköping, Sweden
andpe@itn.liu.se

September 10, 2007

Abstract

Origin-destination (OD) matrices are essential for various analyses in the field of traffic planning, and they are often estimated from link flow observations. We compare strategies for allocating link flow detectors to a traffic network with respect to the quality of the estimated OD-matrix.

First, an overview of allocation strategies proposed in the literature is presented. Second, we give an experimental environment where any allocation strategy can be evaluated, and compared to others, in terms of the quality of the estimated OD-matrix. Third, this environment is used to evaluate and compare three known allocation strategies. Studies are made on the Sioux Falls network and on a network modelling the city of Linköping. Our conclusion is, that the most commonly studied approach for detector allocation, the OD-pair coverage strategy, seems to be unfavourable for the quality of the estimated OD-matrix.

1 Introduction

For most analyses in the field of planning and controlling traffic there is a need for origin–destination (OD) matrices, which specify the travel demand for each pair of origin and destination nodes in the traffic network. Making decisions regarding the design of the traffic system, evaluating the accessibility to city areas, and computing traffic emissions are examples from the broad variety of their uses. OD-matrices are often estimated from traffic link flow observations, and the problem of accurately estimating OD-matrices using this information has been studied thoroughly in the last quarter of a century.

The quality of the estimated OD-matrix depends of many factors. The traffic assignment assumptions and the quality of the link flow observations are important aspects. The dependencies between different link flows, which result from the network topology, must also be considered. We also know that the choice of OD-matrix estimation method will affect the result. Finally, since traffic flow observations typically are available only for some of the network links, the quality of the estimated OD-matrix certainly depends on how these are chosen.

In the literature, several methods for allocating detectors in a traffic network are proposed, with the underlying purpose to provide information to an OD-matrix estimation procedure. All of these methods are intuitively sound and seem to be well motivated for this purpose. Quite surprisingly, they have however not been evaluated and compared with respect to the quality of the estimated OD-matrix. In this paper, we make the first attempt ever to study the problem of allocating link flow detectors in relation to the overall goal of finding a good estimate of the OD-matrix.

There are three major contributions in this paper: First, we present a comprehensive survey and classification of detector allocation strategies. Unlike earlier papers, our focus is on these strategies' allocation objectives, rather than on solution techniques. Second, we design an experimental environment, which is independent of the used allocation strategy and where the impact of different strategies on the quality of the estimated OD-matrix can be evaluated. It can thus be used for evaluating and comparing different allocation strategies, for a certain network. The experimental environment can also be used for studying the role of the traffic assignment procedure. Third, by using this experimental environment, we perform a draft comparative study of different detector allocation strategies and traffic assignment procedures with respect to the quality of the estimated OD-matrix.

The remainder of the paper is organized as follows. In the next section we give the survey of allocation methods proposed in the literature. The methods are classified according to their overall detector allocation objectives. For each method class, we give a representative mathematical programming formulation. In Section 3 we design an experimental environment in which the quality of an estimated OD-matrix can be evaluated. We identify the detector allocation, the traffic assignment, and the OD-matrix estimation procedures as the key components. In Section 4 we present some numerical results, and Section 5 closes the paper with concluding remarks.

2 Overview of detector allocation methods

A detector allocation method aims at finding a subset of network links, where to allocate flow detectors. We consider the case where the only motive for allocating detectors is OD-matrix estimation from link flow observations. Detectors are sometimes also allocated for other purposes, such as the management of traffic signals, different ITS applications, or for example road tolls. The link flow observations provided by such detectors can, of course, be utilized also for OD-matrix estimation. From that point of view, such detectors can however be regarded to be allocated beforehand, and do not need to be included in the allocation problem.

Many proposed allocation methods contain a budget constraint, specifying the number of detectors to be allocated. The allocation problem then solely considers where to place these detectors. Alternatively, the problem is formulated as to minimize the number of detectors which are required as to guarantee some coverage of the total traffic flow. In that case, both the number of detectors and their allocations are determined simultaneously.

Many proposed allocation strategies assume that some information about the traffic in the network is known beforehand. Typically a so-called target OD-matrix, which is the a priori best known estimate of the correct OD-matrix, is available. The corresponding flows on routes and links in the network can then be computed by a traffic assignment procedure. The route flows contain information about which links that are used by travelers in a certain OD-pair, and the flows on any of these links thus provide information about the travel demand in the pair.

If no route choice information is available, one must assume that every possible route from an origin to a destination can be used, by some unknown portion of the travelers in the OD-pair. To be able to provide information about the travel demand in an OD-pair, one must therefore be certain to observe the flow on at least one of the links in every possible route in the pair. Allocation methods based on this requirement are referred to as screen-line based approaches, since such methods try to find screen-lines through the network, completely separating as many OD-pairs as possible. In practice, these methods typically allocate detectors to natural screen-lines, such as railways or rivers.

Our study is focused on how the detector allocation impacts the estimated OD-matrix. Nearly all models for estimation of OD-matrices from link flow observations rely on some assumption on the route choices, or, at least, the target OD-matrix and the traffic assignment principle. Clearly, the route flow information, which is used for estimating the OD-matrix, is available also when the detectors are allocated. Hence, to use a screen-line based approach for the detector allocation would be not to make use of all available information. Furthermore, we primarily consider the case where the traffic is to be covered as well as possible with a limited number of detectors. (For Swedish conditions typically detectors are available only for some five percent of the total number of links in the network.) Requiring a strict screen-line separation of the origin and destination nodes would then result in a very poor coverage.

Next we give an overview of detector allocation methods, and emphasize their differences by classifying them with respect to the fundamental allocation objective. Each strategy is summarized through an integer programming problem, which should be seen as a canonic representation of the strategy, though in some cases the specific integer program cannot be found in the literature.

2.1 Preliminaries

We describe the traffic network by a set of links, A, and a set of OD-pairs, I. The OD-matrix is denoted $g = \{g_i\}$, where g_i is the travel demand in OD-pair $i \in I$. When g is assigned onto the network, it will split up onto the routes in the sets K_i, $i \in I$. Denote the route flow solution by $h = \{h_k\}$, where h_k expresses the flow that is assigned onto route $k \in K_i$, $i \in I$.

By $\Delta = \{\delta_{ak}\}$ we denote the route–link incidence matrix, where element δ_{ak} tells if route $k \in K_i$, $i \in I$, passes link $a \in A$ ($\delta_{ak} = 1$) or not ($\delta_{ak} = 0$). We can then compute the flow on link a as

$$v_a = \sum_{i \in I} \sum_{k \in K_i} \delta_{ak} h_k, \qquad \forall a \in A, \tag{1}$$

and denote by $v = \{v_a\}$ the link flow solution.

A more aggregated description of link flows in terms of travel demands uses the assignment map $P = \{p_{ai}\}$, where element p_{ai} is the proportion of the travel demand in OD-pair $i \in I$ that is assigned onto link $a \in A$. We then have that

$$v_a = \sum_{i \in I} p_{ai} g_i, \qquad \forall a \in A. \tag{2}$$

The detector allocation strategies to be described are based on information about the traffic situation in the network, as follows. We denote by $\hat{g}$ a target OD-matrix, which is typically an out-of-date OD-matrix which shall be refreshed through an OD-matrix estimation procedure. By the use of an assignment procedure we can obtain a corresponding route flow solution, $\hat{h}$, an assignment map, $\hat{P}$, and a target link flow solution $\hat{v}$, which are related through formulas (1) and (2).

In order to state detector allocation problems mathematically, we introduce for $a \in A$ the binary decision variables $x = \{x_a\}$, where

$$x_a = \begin{cases} 1, & \text{if a detector is allocated to link } a, \\ 0, & \text{otherwise.} \end{cases}$$

Letting n be the number of detectors to allocate, we formulate a budget constraint

$$\sum_{a \in A} x_a = n. \tag{3}$$

2.2 Maximum flow coverage

The Maximum Flow Coverage (MFC) strategy, proposed by Lam and Lo (1990), is a quite simple detector allocation rule. As many detectors as are available are allocated according to a ranking with respect to descending target link flows $\hat{v}_a$, $a \in A$. This strategy is formalized by the integer linear program

$$\begin{aligned} \textbf{[MFC]} \quad & \max \sum_{a \in A} \hat{v}_a x_a \\ & \text{s.t.} \sum_{a \in A} x_a = n, \\ & x_a \in \{0,1\}, \qquad \forall a \in A. \end{aligned}$$

We remark that links with large traffic flows are likely to be located along major roads, and since no attention is given to the dependencies among the link flows, it is then likely that many travellers will be counted more than once. The MFC strategy might be important for detecting and controlling bottlenecks in the traffic network, but surely not suitable for providing link flow observations for OD-matrix estimation.

2.3 OD-pair coverage

The OD-Pair Coverage (ODPC) strategy is also proposed by Lam and Lo (1990). The idea is to cover as many OD-pairs as possible, regardless of the travel demand. A link is defined to cover an OD-pair if a minimum proportion of the travel demand in the pair passes the link, and the OD-pair has not been covered before. The OD-pair covering model proposed by Yang et al. (1991) differs from the approach of Lam and Lo in that the number of detectors is variable; the objective is to minimize the number of detectors subject to that all OD-pairs must be covered, at least once.

Yang and Zhou (1998) and Meng et al. (2005) have proposed related approaches, which differ with respect to which links that are eligible to cover a certain OD-pair. Entitled "maximal flow fraction rule", Yang and Zhou formulate a model where only those links are eligible that carry the maximum fraction of the travel demand in the OD-pair. Meng et al. consider a link to be eligible if it is included in an efficient route, that is, a route where the head node of each link along the route is always closer to the destination than the link's tail node, see Dial (1971).

In a version of the OD-pair coverage strategy by Ehlert et al. (2006), it is assumed that the cost for allocating a detector to a certain link is dependent on the link choice. Therefore an allocation cost is associated with each link and the total allocation cost, rather than the number of detectors, is being minimized. The authors also consider an inverse formulation, where a minimum number of OD-pairs are left uncovered, given a budget restriction for the detectors. As an extension, the authors assume that some OD-pairs are covered beforehand.

Logie and Hynd (1990) are the first to use a screen-line based measure. They want to draw screen-lines such that the number of covered OD-pairs is maximized. For breaking ties, double-covering of OD-pairs is minimized, achieving the maximum information value of the counts. Their strategy is not formulated formally.

Chootinan et al. (2005) propose a two-objective screen-line based formulation, where the tradeoff between the coverage of OD-pairs and the number of link flow detectors being used, can be studied. They use a genetic algorithm to generate non-dominated solutions.

Gan et al. (2005) also use a screen-line based approach, which is analogous to the ODPC approach by Yang et al. (1991), apart from the use of a screen-line definition of OD-pair coverage. Yang et al. (2006) further develop this screen-line approach and try to minimize the number of OD-pairs not being separated by any screen-line. As an alternative, they consider minimizing the number of detector links, given that all OD-pairs are separated by at least one screen-line.

Chen et al. (2007) use a screen-line based OD-pair coverage strategy to investigate how a given allocation should be augmented with additional detectors. They model both the situation where the remaining OD-pairs are to be covered with a minimum number of detectors, and the situation where a limited number of additional detectors should be placed as to maximize the number of covered OD-pairs.

To state an ODPC strategy mathematically, we introduce for each $i \in I$ the variable

$$y_i = \begin{cases} 1, & \text{if OD-pair } i \text{ is covered,} \\ 0, & \text{otherwise.} \end{cases}$$

We also need to define which links that can cover a certain OD-pair. Let

$$\hat{\pi}_{ai} = \begin{cases} 1, & \text{if link } a \text{ is eligible to cover OD-pair } i, \\ 0, & \text{otherwise,} \end{cases}$$

and identify $\hat{\Pi} = \{\hat{\pi}_{ai}\}$, where $a \in A$, $i \in I$, as an OD-pair–link incidence matrix.

The key issue here is how to define that a link is eligible to cover an OD-pair. In general, the appearance of $\hat{\Pi}$ is a result of the assignment map $\hat{P}$, which in turn is a result of the assignment procedure used, and the target OD-matrix $\hat{g}$. The complete derivation of $\hat{\Pi}$ from the information contained in $\hat{P}$ requires the use of some formal criterion. An example criterion for a link to be eligible to cover an OD-pair is that a minimum proportion of the travel demand in the pair passes the link. A second example is that a minimum proportion of the links should be eligible for each OD-pair. In Section 3.3 we explain how $\hat{\Pi}$ is defined for our numerical experiments.

We can now formulate an ODPC strategy as the integer program

$$
\begin{array}{lrcll}
\textbf{[ODPC]} & \max \sum_{i \in I} y_i & & & \\
\text{s.t.} & \sum_{a \in A} \hat{\pi}_{ai} x_a & \geq & y_i, & \forall i \in I, \\
& \sum_{a \in A} x_a & = & n, & \\
& y_i & \in & \{0,1\}, & \forall i \in I, \\
& x_a & \in & \{0,1\}, & \forall a \in A.
\end{array}
$$

By introducing weights for the OD-pairs in the objective and costs for the link detectors in the budget constraint, this formulation turns into the "BIP-5" formulation by Ehlert et al. (2006).

2.4 Route coverage

In the Route Coverage (RC) strategy, proposed by Gentili and Mirchandani (2005), the aim is to calculate route flows $\hat{h}$ from flow observations $\hat{v}_a$ on a link subset, using the corresponding equations in (1). We recognize this aim as the problem to cover routes, regardless of the flows on them, by allocating link detectors.

Though Gentili and Mirchandani consider the case where all routes should be covered by using a minimum number of detectors, we give a formulation where the number of covered routes should be maximized, subject to a budget constraint; cf. the formulation of the ODPC strategy above. We introduce for each $k \in K_i$, $i \in I$, the binary variable

$$
\psi_k = \begin{cases} 1, & \text{if route } k \text{ is covered,} \\ 0, & \text{otherwise.} \end{cases}
$$

The RC strategy can then be formulated as

$$
\textbf{[RC]} \quad \max \sum_{i \in I} \sum_{k \in K_i} \psi_k
$$

$$
\begin{aligned}
\text{s.t.} \quad \sum_{a \in A} \delta_{ak} x_a &\geq \psi_k, && \forall k \in K_i, \forall i \in I, \\
\sum_{a \in A} x_a &= n, && \\
\psi_k &\in \{0,1\}, && \forall k \in K_i, \forall i \in I, \\
x_a &\in \{0,1\}, && \forall a \in A.
\end{aligned}
$$

In the approach of Gentili and Mirchandani the sets of routes K_i, $i \in I$, are pre-defined, and can be independent of target OD-matrix $\hat{g}$ and assignment procedure. In practice, the number of possible routes is very large, and it might therefore be necessary to somehow restrict the number of routes considered in the RC strategy.

2.5 OD-demand coverage

Closely related to the ODPC strategy is the OD-Demand Coverage (ODDC) strategy. As in the ODPC strategy, it is pre-defined which links that are eligible to cover the demand in an OD-pair, but instead of maximizing the number of OD-pairs covered, the goal is here to maximize the total demand captured. The approach of Hodgson (1990) is the first of this type, although it assumes that detectors are allocated to the nodes, rather than to the links; this is however of no major importance.

In the ODDC approach of Yim and Lam (1998), coverage of travel demand in an OD-pair can be fulfilled by capturing non-overlapping portions of this demand on several links. Hence, all links are here eligible to contribute to the coverage of an OD-demand (although the contribution can be small).

Our formulation of the ODDC strategy is analogous to that of the ODPC strategy, given above. We only need to weight every OD-pair with its travel demand, to obtain the model

$$
\begin{aligned}
\textbf{[ODDC]} \quad & \max \sum_{i \in I} \hat{g}_i y_i \\
\text{s.t.} \quad \sum_{a \in A} \hat{\pi}_{ai} x_a &\geq y_i, && \forall i \in I, \\
\sum_{a \in A} x_a &= n, && \\
y_i &\in \{0,1\}, && \forall i \in I, \\
x_a &\in \{0,1\}, && \forall a \in A.
\end{aligned}
$$

This model can be generalized by substituting the target demand $\hat{g}$ with some other weights for the OD-pairs, in which case it coincides with the weighted ODPC model by Ehlert et al. (2006).

2.6 Route flow coverage

Analogously to how the ODDC strategy can be described as a weighted form of the ODPC strategy, we can define the Route Flow Coverage (RFC) strategy as a weigted form of the RC strategy. Yang and Zhou (1998) have proposed the "maximal flow-intercepting rule", which states that the detector links should be chosen as to intercept as many route flows as possible. When applying the rule, they maximize the total covered non-overlapping route flows, that is, the routes are weighted with the route flows $\hat{h}$. In the RFC model by Ehlert et al. (2006) the route flow coverage goal is combined with the requirement that all OD-pairs should be covered. Already in 1990 a greedy heuristic for RFC was implemented as a subroutine to the software tool Emme/2 (INRO, 1990).

The formulation of the RFC strategy is a straightforward extension of the RC strategy given above. Given the route flows $\hat{h}$, we can formulate the problem as

$$
\begin{aligned}
\textbf{[RFC]} \quad & \max \sum_{i \in I} \sum_{k \in K_i} \hat{h}_k \psi_k \\
\text{s.t.} \quad & \sum_{a \in A} \delta_{ak} x_a \geq \psi_k, \qquad \forall k \in K_i, \forall i \in I, \\
& \sum_{a \in A} x_a = n, \\
& \psi_k \in \{0,1\}, \qquad \forall k \in K_i, \forall i \in I, \\
& x_a \in \{0,1\}, \qquad \forall a \in A.
\end{aligned}
$$

2.7 Other aspects

Lam and Lo (1990) propose an allocation strategy based on the mean, taken over all OD-pairs, of the relative deviation between the estimated demand and the target. The detectors are allocated with a greedy heuristic, with the goal to minimize the maximal relative deviation. A major drawback with this strategy is that the OD-matrix to be estimated is required as input to the heuristic, and therefore we exclude this strategy from further consideration.

Yang and Zhou (1998) have formulated a "link independence rule", which states that the detectors should be located in such a way that the corresponding rows in assignment map $\hat{P}$ are linearly independent. This rule is not a goal in itself, but of course a reasonable condition.

Bianco et al. (2001) and Confessore et al. (2005) have proposed detector allocation schemes based on node flow observations. By measuring the flow passing a node, and further, by also measuring the turning coefficients, they propose a model where the detectors are allocated in such a way that all link flows can be calculated. The inclusion of turning coefficients is beyond the scoop of our presentation.

	Objective					Base		
	MFC	ODPC	ODDC	RC	RFC	Screen-line based	Old Assignment	Budget Constraint
Lam and Lo (1990): T F[1]	X						X	
Lam and Lo (1990): O-D C[2]		X					X	
Hodgson (1990)			X			[3]	[3]	X
Logie and Hynd (1990)		X				X		
Yang et al. (1991)		X					X	
Yang and Zhou (1998)[4]					X		X	X
Yim and Lam (1998)			X				X	
Chootinan et al. (2005)		X				X		[5]
Gan et al. (2005)		X				X		
Gentili and Mirchandani (2005)				X		[3]	[3]	
Meng et al. (2005)[6]		X					X	
Yang et al. (2006)[7]		X				X		X
Ehlert et al. (2006)[8]		X[9]					X	X
Chen et al. (2007)		X				X		[10]

[1] The "traffic flow" model.

[2] The "O-D covering" method.

[3] How the set of routes for an OD-pair is defined, is not described in the paper.

[4] Here we refer to the problem formulation "TCL-P2". We note that the formulation "TCL-P1" coincides with that by Yang et al. (1991).

[5] The number of detectors is one part in a multi-objective formulation.

[6] This is the same formulation as is proposed by Yang et al. (1991).

[7] Here we refer to the problem formulation "ILP-1". We note that the formulation "ILP-2" coincides with that by Gan et al. (2005).

[8] Here we refer to the problem formulation "BIP-5". We note that the formulations given in "BIP-1" and "BIP-2" coincide with those by Yang et al. (1991) and Yang and Zhou (1998), respectively. "BIP-3" is just a weighted version of "BIP-1", where the allocation cost is taken into account. "BIP-4" and "BIP-6", finally, are just two alternative formulations of "BIP-5", where "BIP-6" assumes that some OD-pairs are already covered.

[9] Each OD-pair is given an individual weight, which, for example, could take the target demand into account. In that case, we should talk about ODDC.

[10] Both cases are covered in different formulations.

Table 1: Characteristics for different allocation strategies.

2.8 Overview of allocation strategies

In Table 1 we summarize and characterize the detector allocation strategies discussed above. The most important characteristic of each allocation method is its overall **Objective**. Another important disjunction is the **Base** for the allocation, that is whether the model formulation is screen-line based or relies on an old assignment. We have also marked those formulations where a budget constraint, like (3), is included.

3 Design of the comparative study

The overall goal for the detector allocation is to find those links on which flow observations enable the very best estimate of the travel demand in the network. For evaluating a certain allocation, we must thus be able to determine the quality of an estimated OD-matrix. Next we design an experimental environment to accomplish this. This environment contains three distinct components: an OD-matrix estimation procedure, a traffic assignment algorithm, and an allocation method.

In real life, a link flow count is, of course, not available unless we actually place a detector on the link, and as a consequence of this, a certain detector allocation can be evaluated only by actually making it. The idea behind the design of our experimental environment is to simulate the flow that could be observed if a detector would be allocated to a certain link. This simulation enables a comparison of different allocations of detectors, and also of different detector allocation strategies.

3.1 The experimental environment

Let $\bar{g}$ be a given, true, OD-matrix, and let $\hat{g} \neq \bar{g}$ be a target OD-matrix, to be used both in the detector allocation method and in the OD-matrix estimation procedure. For experimental purposes we generate $\hat{g}$ from $\bar{g}$ by a random perturbation.

The target OD-matrix $\hat{g}$ is assigned onto the network, according to a traffic assignment principle, which gives a link flow solution $\hat{v}$. Let the route flow solution $\hat{h}$ and the assignment map $\hat{P}$ be consistent with $\hat{v}$, according to the equations in (1) and (2), respectively. The quantities $\hat{g}$, $\hat{P}$, $\hat{h}$ and $\hat{v}$, as well as the OD-pair–link incidence matrix $\hat{\Pi}$ are used as input to a detector allocation strategy, which gives an (optimal) detector allocation x^*. For convenience, we define the subset of links where detectors are allocated, $\bar{A} = \{a \in A \mid x_a^* = 1\}$.

Let $\bar{v} = \{\bar{v}_a\}$, where $a \in \bar{A}$, be the link flows resulting from an assignment of $\bar{g}$. (Note that $\hat{v}$ and $\bar{v}$ are of different dimensions.) These are synthetic observed flows that are used together with the target OD-matrix $\hat{g}$ for performing the OD-matrix estimation, giving a resulting OD-matrix, g^*, which can be compared with the true OD-matrix, $\bar{g}$. In Figure

1 an illustration of the experimental environment is given.

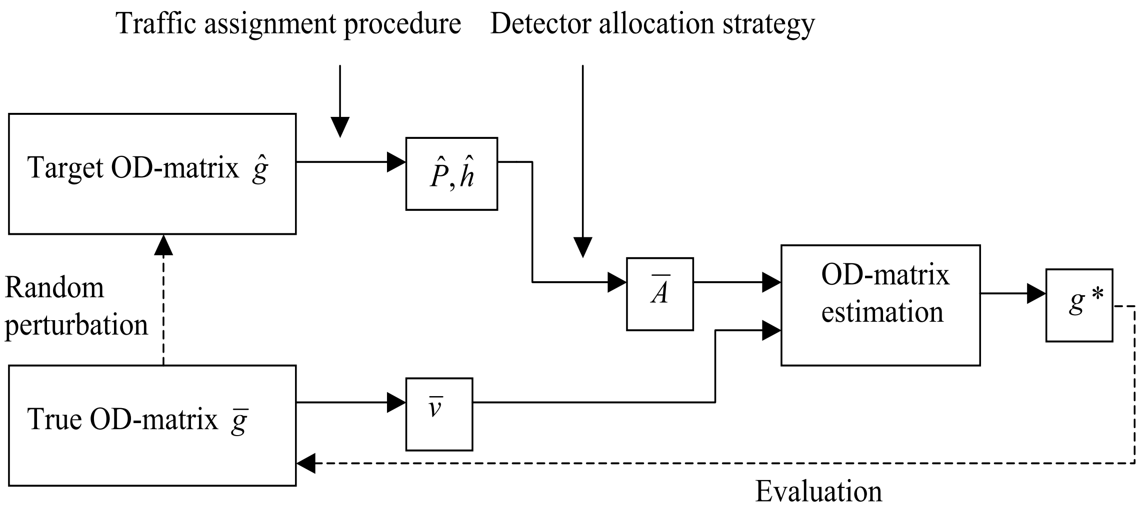


Figure 1: Illustration of the experimental environment.

In the OD-matrix estimation problem, the primary objective is to find a new OD-matrix, which, when assigned onto the network, induces link flows that reproduce the observed link flows as well as possible. Typically, this problem is under-determined. Therefore, as a secondary objective, one also wishes to choose an OD-matrix that is close to the target matrix. For an overview of some models and methods for the OD-matrix estimation problem, see for example Cascetta (2001), Ch. 8.5; Ortúzar and Willumsen (2001), Ch. 12.4; or Bell and Iida (1997), Ch. 7.

When formulating the OD-matrix estimation problem, there are different ways of combining and weighting the information in the target OD-matrix and the link flow observations, respectively. Our purpose of performing the estimation is merely to make a comparative study of different detector allocation strategies, with respect to the quality of the estimated OD-matrix. Therefore we use a well-known and straightforward single-objective formulation of the OD-matrix estimation problem, namely

$$\begin{aligned} \min \quad F(g, \hat{g}) &= \frac{1}{2} \sum_{i \in I} (g_i - \hat{g}_i)^2 \qquad (4) \\ \text{s.t.} \ \sum_{i \in I} \hat{p}_{ai} g_i &= \bar{v}_a, \qquad \forall a \in \bar{A}, \\ g_i &\geq 0, \qquad \forall i \in I. \end{aligned}$$

This problem amounts to finding the OD-matrix, g^*, being closest to the target OD-matrix $\hat{g}$, among those non-negative OD-matrices that exactly reproduce the observed link flows $\bar{v}$. We outline in an appendix a Lagrangian dual method for solving (4) and describe how it can be implemented. Since $\hat{P}$ and $\bar{v}$ arise from different assignments, they might be incompatible in such a way that the problem (4) becomes infeasible; this case can be detected in the dual method. Infeasibility is not likely to occur in the OD-matrix estimation problems that we study, due to the small number of detectors compared to the total number of network links.

OD-matrix estimation problems are inherently under-determined, since the number of OD-pairs is typically much larger than the number of link flow observations. In particular, the estimation problem (4) will therefore in general not be able to recover the true OD-matrix, $\bar{g}$. Not even if all link flows were observed (i.e. $\bar{A} = A$), and these flows would be compatible, can we be certain to recover $\bar{g}$. Furthermore, when many link flows are observed, the problem (4) is likely to become infeasible.

3.2 Traffic assignment algorithms

The traffic assignment problem is well studied; for an overview of models and methods we refer to, for example, Patriksson (1994). The experimental environment could be used in combination with any traffic assignment principle, but a fair comparison fails if $\bar{v}$ and $\hat{v}$ are computed by different principles. We focus on the deterministic user equilibrium principle, which is the most widely used in practice. Whenever the link travel times are monotonically increasing with respect to the flows, the equilibrium link flows are uniquely determined. However, the route flows and the assignment map are typically non-unique and depend on the assignment algorithm and its implementation. These quantities are of decisive importance in the context of OD-matrix estimation and detector allocation, and the choice of assignment algorithm might therefore be crucial for the results.

In our comparative study we have used three assignment algorithms. The Disaggregate Simplicial Decomposition (DSD) algorithm (Larsson and Patriksson, 1992), is chosen because it is known to be quite efficient and to generate an extreme route flow solution, that is, a route flow solution where the number of routes being used is very small. The well-known Frank–Wolfe (FW) algorithm has been included in the study because of its popularity and commercial use. In the FW assignment, many routes are typically used. For a comparison of DSD versus FW route flow solutions, we refer to Tatineni et al. (1998).

The third assignment procedure used is an Incremental Assignment (IA) heuristic. The total demand is then divided into a sum of incremental demands, and each of these is assigned using the all-or-nothing algorithm based on the travel times arising from the previous incremental loadings. The IA heuristic typically terminates with a solution which is an approximate equilibrium.

For solving the traffic assignment problems we use Ciudadsim, which is a traffic network toolbox implemented in Scilab. Ciudadsim is available on Internet at

`http://www-rocq.inria.fr/metalau/ciudadsim/`

and the algorithms are well-documented in the user's manual (Lotito et al., 2003).

3.3 Detector allocation strategies

The experimental environment, described above, can in principle be used to evaluate the effect of any detector allocation. Our comparative study includes the MFC, ODPC and ODDC strategies.

The implementation of the MFC strategy is straightforward; the links are sorted in decreasing order with respect to their target link flows. The first n links in the list define $\bar{A}$. It can be noted that the MFC strategy, thus, is a greedy approach.

The formulation of the ODPC strategy, which is given in Section 2.3, is a combinatorial optimization problem and might be hard to solve, even for medium-sized networks. In our study, we have implemented a simple greedy heuristic for its solution. We start with a set of detector links that is empty, and repeatedly augment it with the link that induces the largest increase in the objective value. At each augmentation, $\hat{\Pi}$ is updated by setting $\hat{\pi}_{ai}$ to zero for all OD-pairs being covered. A similar greedy heuristic has been implemented for the ODDC strategy.

Thus, all our detector allocation strategies are implemented as greedy heuristics. This means that comparing the results of allocating n versus $n+1$ detectors can be interpreted in terms of adding the information given by the flow on one additional link. Further, since the heuristics are greedy, if problem (4) is infeasible when n detectors are allocated, this will also hold when $n+1$, or more, detectors are allocated. If a greedy strategy is not used, this property might not hold.

The optimization problems in the ODPC and ODDC strategies are solved approximately only, wheras the MFC strategy is solved exactly. This should be kept in mind when comparing the numerical results. However, since the same approach is used for both the ODPC and the ODDC strategies, we believe that the results for these two methods are fully comparable.

In the implementation we must specify how the target OD-pair–link incidence matrix $\hat{\Pi}$ is derived from the assignment map $\hat{P}$. We simply define a threshold value $0 < \alpha \leq 1$ and set $\hat{\pi}_{ai} = 1$ whenever $\hat{p}_{ai} \geq \alpha$, and $\hat{\pi}_{ai} = 0$ otherwise. This follows the definition in Ehlert et al. (2006), though we do not concerning the possibility to define α individually for each OD-pair. At an early stage, we made some experiments considering the threshold value α, which was varied between 0.35 and 0.81. The value of α had a very small impact on the results obtained, and we decided to henceforth use $\alpha = 0.51$, which ensures that no alternative path can be more representative for an OD-pair.

4 Numerical results

We here present the results of the comparative study of detector allocation strategies in the context of OD-matrix estimation. The detector allocation strategies and the OD-matrix estimation procedure are coded in Matlab 7.0. Problem (4) is considered to be feasible if all link flows $\bar{v}$ are recovered, by $\hat{P}$ and g^*, to a relative accuracy of 10^{-8}. For all combinations of test network, detector strategy, assignment procedure, and number of detectors being allocated, problem (4) turned out to be feasible, with one exception which is discussed in Section 4.3.

We first present some numerical results for a small test network, with the purpose of verifying the experimental environment's functionality. We then present results for the networks modelling the cities of Sioux Falls and Linköping.

4.1 Verification of the experimental environment

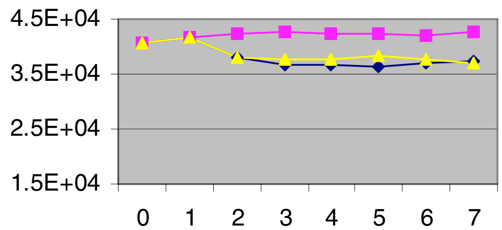
In the initial experiments we used a fictitious test network, constructed by Edwards (2000). The Edwards network consists of 66 nodes, 144 links, and 90 OD-pairs. We have chosen to allocate up to $n = 7$ detectors. The assignment map $\hat{P}$ is generated with the IA (with 25 increments), FW and DSD procedures, respectively.

The random perturbations, used for creating a target OD-matrix $\hat{g}$ (see Section 3.1), are multiplicative and we first choose them from a uniform distribution on the interval $[0.75, 1.25]$, individually for each OD-pair. The expected value for each travel demand $\bar{g}_i$ is $\hat{g}_i$, that is, we have no trend. Figure 2 shows the results for the different assignment procedures and allocation strategies. On the x-axis, the different values of n are shown. The y-axis shows the deviation between the estimated OD-matrix, g^*, which is obtained from the estimation problem (4), and the true OD-matrix, $\bar{g}$; this deviation is calculated as $\frac{1}{2}\sum_{i\in I}(g_i^* - \bar{g}_i)^2$, that is, a measure analogous to that used in problem (4).

In the figure we can see that the results differ relatively little with respect to the choice of assignment procedure. Among the allocation strategies, MFC and ODPC show a slightly better result than ODPC. The overall improvement of the OD-matrix, from $\hat{g}$ towards $\bar{g}$, is relatively small. A plausible explanation is the properties of the problem (4), which are discussed at the end of Section 3.1.

Define the set of uncovered OD-pairs as $\tilde{I} = \{i \in I \mid \hat{p}_{ai} = 0,\ \forall a \in \bar{A}\}$. In the figure, we give the number of uncovered OD-pairs, $|\tilde{I}|$, when seven detectors are allocated, for each of the assignment procedures and allocation strategies. We also give a measure of the uncovered travel demand, calculated as $\frac{1}{2}\sum_{i\in\tilde{I}}(g_i^* - \bar{g}_i)^2$. (Note that, $g_i^* = \hat{g}_i$ holds for all $i \in \tilde{I}$.) The number of uncovered OD-pairs should be compared with the total number of OD-pairs and the uncovered travel demand measurement should be compared with the deviation measurement, depicted on the y-axis. Again, the results differ relatively little with respect to the choice of assignment procedure. Further, the ODPC strategy gives

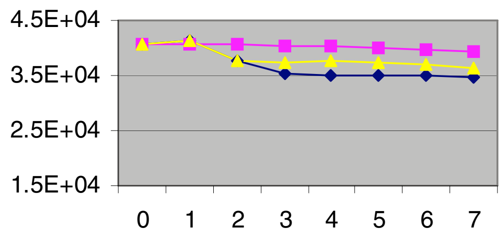
Fig 2a: Edwards, without trend, IA



Uncoverage with 7 detectors.

	OD-pairs	Demand
MFC	46	8.26E+03
ODPC	28	1.78E+04
ODDC	38	1.62E+03

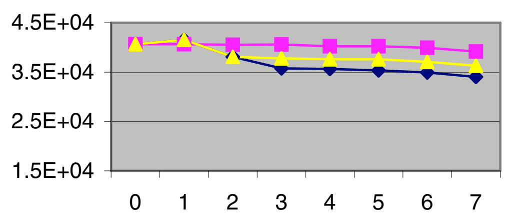
Fig 2b: Edwards, without trend, FW



Uncoverage with 7 detectors.

	OD-pairs	Demand
MFC	44	8.17E+03
ODPC	25	2.01E+04
ODDC	37	1.61E+03

Fig 2c: Edwards, without trend, DSD



Uncoverage with 7 detectors.

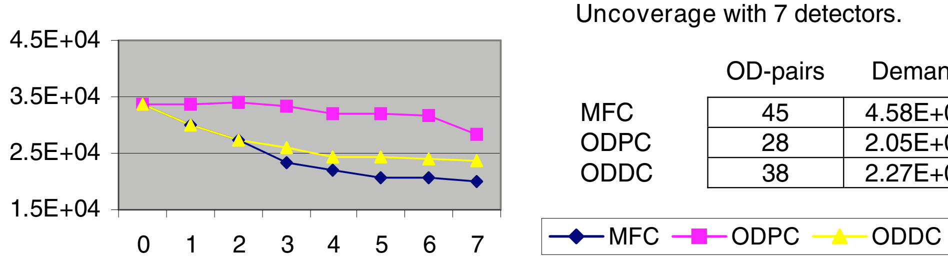
	OD-pairs	Demand
MFC	42	1.19E+04
ODPC	29	2.13E+04
ODDC	38	1.62E+03

Figure 2: Results for the Edwards network, without trend, for the assignment procedures IA (a), FW (b) and DSD (c).

good coverage of OD-pairs, but very poor coverage of the travel demand, while the ODDC strategy gives a very good coverage of the travel demand, but poor coverage of OD-pairs. These results seem quite reasonable. The MFC covers the travel demand fairly well, but it covers few OD-pairs.

In Figure 3 we present the results where instead the random perturbation is chosen from the interval $[1, 1.25]$, that is, with a trend. These results are quite similar to those presented in Figure 2, as to the assignment procedures and the ranking of the detector allocation strategies. The total improvement of the OD-matrix, measured as $\frac{1}{2}\sum_{i\in I}(g_i^* - \bar{g}_i)^2$, is however much smaller when the perturbation has no trend. Our explanation for this is that since travel demand from many OD-pairs passes each link, the positive and negative perturbations without trend are evened out, and the resulting target link flow, $\hat{v}_a$, is very close to the true link flow, $\bar{v}_a$. In such a case, the information provided by detector observations is of little use for estimating the true OD-matrix $\bar{g}$. Therefore, from now on we only consider cases where the perturbation has a trend.

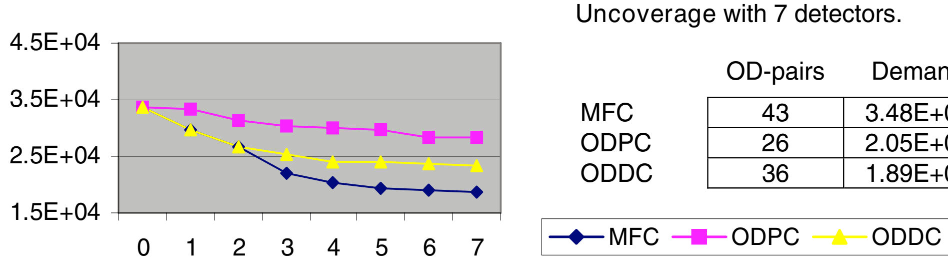
Fig 3a: Edwards, with trend, IA



Uncoverage with 7 detectors.

	OD-pairs	Demand
MFC	45	4.58E+03
ODPC	28	2.05E+04
ODDC	38	2.27E+03

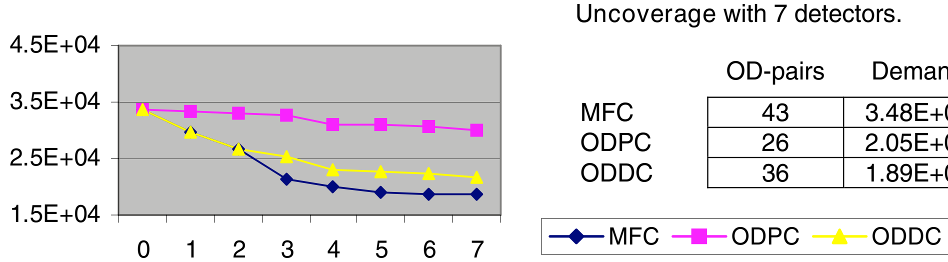
Fig 3b: Edwards, with trend, FW



Uncoverage with 7 detectors.

	OD-pairs	Demand
MFC	43	3.48E+03
ODPC	26	2.05E+04
ODDC	36	1.89E+03

Fig 3c: Edwards, with trend, DSD



Uncoverage with 7 detectors.

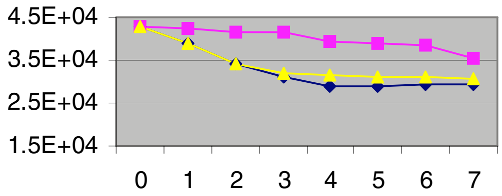
	OD-pairs	Demand
MFC	43	3.48E+03
ODPC	26	2.05E+04
ODDC	36	1.89E+03

Figure 3: Results for the Edwards network, with trend, for the assignment procedures IA (a), FW (b) and DSD (c).

The target OD-matrices are generated randomly, and to illustrate the effect of this we have repeated the experiment, the result of which is presented in Figure 3, but with another random seed. The results are shown in Figure 4.

The deviation $\frac{1}{2}\sum_{i\in I}(\hat{g} - \bar{g}_i)^2$, where the curves start, is more than 25 percent larger in

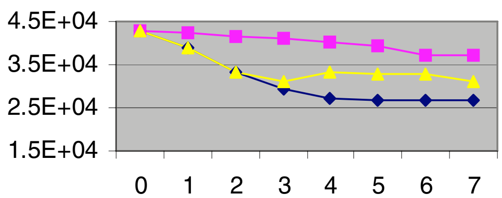
Fig 4a: Edwards, with trend, IA,
2nd random seed



Uncoverage with 7 detectors.

	OD-pairs	Demand
MFC	45	7.45E+03
ODPC	28	2.37E+04
ODDC	34	3.51E+03

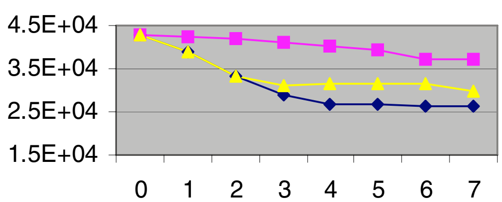
Fig 4b: Edwards, with trend, FW,
2nd random seed



Uncoverage with 7 detectors.

	OD-pairs	Demand
MFC	41	7.25E+03
ODPC	25	2.63E+04
ODDC	36	2.61E+03

Fig 4c: Edwards, with trend, DSD,
2nd random seed



Uncoverage with 7 detectors.

	OD-pairs	Demand
MFC	42	8.79E+03
ODPC	27	2.71E+04
ODDC	38	2.81E+03

Figure 4: Results for the Edwards network, with trend, for the assignment procedures IA (a), FW (b) and DSD (c). Second random seed.

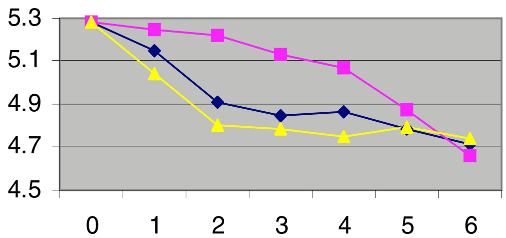
Figure 4 (42,635 compared to 33,683). This difference however, seems to remain unaffected when the number of detectors allocated increases. We conclude that the ranking of the three detector allocation strategies is similar to that in Figure 3. Also with respect to the different assignment procedures, the two random instances behave in the same way. Hence, these two instances lead to same overall conclusions.

We have also performed an experiment with a second random seed for the Edwards network without trend, that is, an experiment comparable to that presented in Figure 2 above. Though we have chosen not to present the results here, this experiment further confirms the independence of the random seed, concerning the evaluation of allocation strategies and assignment procedures.

4.2 Sioux Falls network

The well-known network for the city of Sioux Falls (LeBlanc et al., 1975) has 24 nodes, 76 links and 528 OD-pairs. We allocate up to $n = 6$ detectors. The target OD-matrix $\hat{g}$ is computed as for the previous test networks. We here use 50 increments in the IA algorithm. The results are presented in Figure 5.

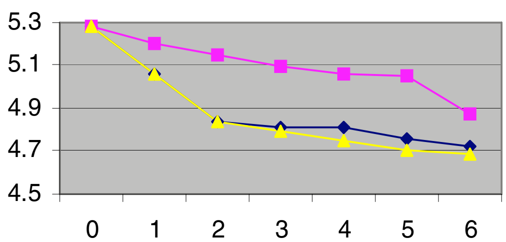
Fig 5a: Sioux Falls, with trend, IA



Uncoverage with 6 detectors.

	OD-pairs	Demand
MFC	338	3.00
ODPC	251	3.18
ODDC	274	3.09

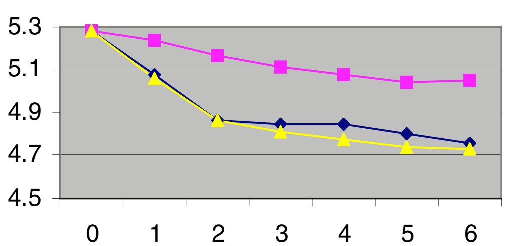
Fig 5b: Sioux Falls, with trend, FW



Uncoverage with 6 detectors.

	OD-pairs	Demand
MFC	222	1.61
ODPC	156	2.12
ODDC	163	1.61

Fig 5c: Sioux Falls, with trend, DSD



Uncoverage with 6 detectors.

	OD-pairs	Demand
MFC	327	3.10
ODPC	274	4.20
ODDC	271	3.25

Figure 5: Results for the Sioux Falls network, with trend, for the assignment procedures IA (a), FW (b) and DSD (c).

Apparently, here the effect of the choice of assignment procedure is larger than for the Edwards network. The IA procedure gives a better result than before, especially in combination with the ODPC allocation strategy. We also note that the results obtained with the MFC and ODDC allocation strategies are quite similar.

4.3 Linköping network

The network that models the city of Linköping in Sweden, has previously been used in, for example, Larsson et al. (2004). The network has 335 nodes, 882 links and $12,372$ OD-pairs. We allocate at most $n = 45$ detectors, and to shorten the computing time, we increase the number of detectors with three at a time. We again use 50 increments in the IA algorithm. The results are presented in Figure 6.

The ODDC strategy together with FW or DSD assignment procedures shows a significantly larger improvement of the OD-matrix than the other methods. The ODPC performs quite poorly, concerning both the OD-matrix estimation and the coverage of travel demand, even though it covers the OD-pairs very well.

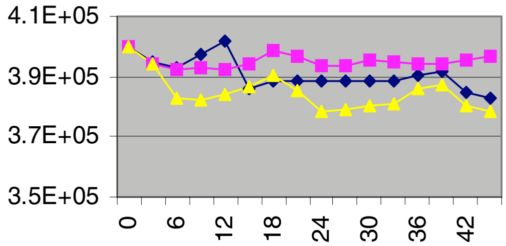
The MFC strategy leaves many more OD-pairs uncovered than the other strategies, while still covering a large amount of the total travel demand. It shows a more irregular behavior here than in the previously presented networks. We explain this by noting that the Linköping network contains some arterial roads and that the MFC strategy tends to create clusters of detectors, so that relatively many detectors appear in small sub-networks (cf. the discussion at the end of Section 3.1). This is believed to cause the estimation problem (4) to have a small set of feasible solutions, or even being near-infeasible, which in turn explains the irregular behavior of the MFC strategy.

Linköping has only about $100,000$ inhabitants and the traffic network is modestly congested. For studying the effect of congestion, we construct a fictitious version of the Linköping network, where the original travel demands are increased by 25 percent (i.e. $\bar{g} := 1.25 \cdot \bar{g}$). Then $\hat{g}$ is generated as in all previous experiments.

The results from this experiment are shown in Figure 7. For the MFC strategy no results are available due to infeasibility in the problem (4). (Infeasibility occured when allocating 3, 9 and 18 detectors in the IA, FW and DSD cases, respectively.) The infeasibility is probably a more pronounced effect of the clustering of detectors, as discussed above.

From the results presented in Figure 7 we see that ODDC strategy dominates the ODPC strategy, independent of the choice of traffic assignment procedure, as in all previous results. Comparing the assignment procedures, the FW algorithm gives the smallest improvement of the OD-matrix. This might be explained by the empirical fact that the FW algorithm produces route flow solutions where the number of used routes is large, and the route flows are small, which makes it difficult to find sufficiently many and representative eligible links.

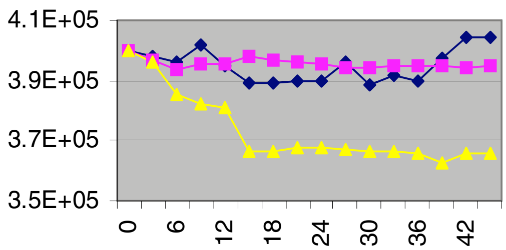
Fig 6a: Linköping, with trend, IA



Uncoverage with 45 detectors.

	OD-pairs	Demand
MFC	3,878	5.32E+04
ODPC	1,110	2.96E+05
ODDC	1,587	1.67E+04

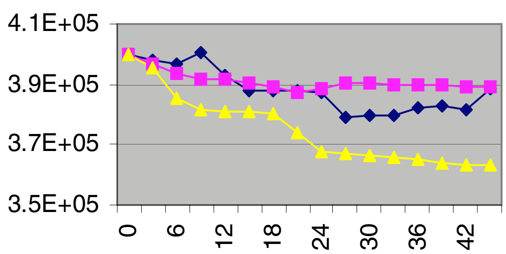
Fig 6b: Linköping, with trend, FW



Uncoverage with 45 detectors.

	OD-pairs	Demand
MFC	3,631	5.12E+04
ODPC	985	2.90E+05
ODDC	1,404	1.49E+04

Fig 6c: Linköping, with trend, DSD



Uncoverage with 45 detectors.

	OD-pairs	Demand
MFC	3,678	5.26E+04
ODPC	1,083	2.96E+05
ODDC	1,448	1.59E+04

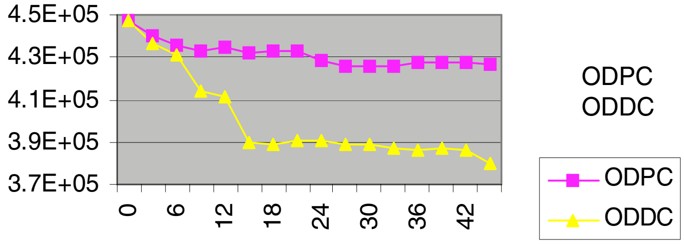
Figure 6: Results for the Linköping network, with trend, for the assignment procedures IA (a), FW (b) and DSD (c).

5 Conclusions and future research

The problem of estimating OD-matrices from traffic link flow observations is well studied. The quality of the estimate depends of many factors; in this paper the impact of the link flow detector allocation strategy was studied.

First, we have given a comprehensive survey of detector allocation methods, which are classified according to their allocation objectives. Second, in order to investigate how the choice of strategy affects the outcome of the OD-estimation procedure, we have designed

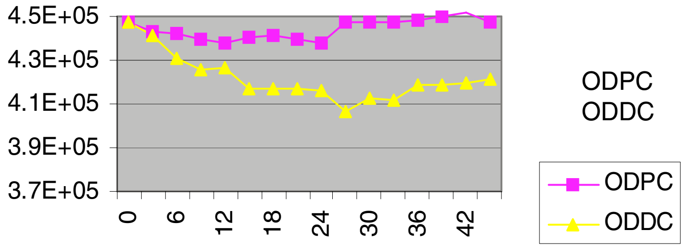
Fig 7a: Linköping congested with trend, IA



Uncoverage with 45 detectors.

	OD-pairs	Demand
ODPC	1,062	2.92E+05
ODDC	1,565	2.22E+04

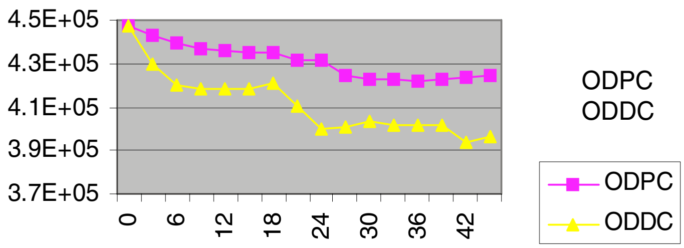
Fig 7b: Linköping congested with trend, FW



Uncoverage with 45 detectors.

	OD-pairs	Demand
ODPC	1,000	1.85E+05
ODDC	1,327	1.90E+04

Fig 7c: Linköping congested with trend, DSD



Uncoverage with 45 detectors.

	OD-pairs	Demand
ODPC	951	1.44E+05
ODDC	1,392	2.15E+04

Figure 7: Results for the 25 percent extra congested Linköping network, with trend, for the assignment procedures IA (a), FW (b) and DSD (c).

an experimental environment. This environment has been implemented, tested and verified. The Lagrangian dual method for the OD-matrix estimation problem (4), described in the appendix, is both robust and able to solve the problem to a high precision.

Third, we have used the experimental environment to study the importance of the choices of detector allocation strategy and traffic assignment procedure for finding accurate estimates of OD-matrices. We conclude from our numerical results that the choice of assignment procedure seems to be of little importance. The maximum flow coverage (MFC) detector allocation strategy, which is the simplest one, gives surprisingly good result in most of the experiments. For the Linköping network, however, which is the most realistic of the networks used, this strategy gives the most irregular results. It is notable that the OD-pair coverage (ODPC) strategy, which is the most popular in the literature (see Table 1), shows the weakest performance in almost all the experiments. The strongest performance is achieved by the OD-demand coverage strategy, which is little studied in the literatur. It could therefore deserve more attention.

In this study, we have only implemented some rather simple greedy heuristics for various detector allocation strategies. For the future it would be interesting to also use the experimental environment to evaluate different algorithms for solving a certain mathematical formulation of a strategy. Another interesting question, of course, is to evaluate additional allocation strategies, such as the route coverage (RC) and route flow coverage (RFC) strategies.

References

Bazaraa, M.S., Sherali, H.D., and C.M. Shetty (1993) *Nonlinear Programming: Theory and Algorithms*, 2nd Edition, Wiley, New York.

Bell, M.G.H., and Y. Iida, Y. (1997) *Transportation Network Analysis*, Wiley, West Sussex, United Kingdom.

Bianco, L., Confessore, G., and P. Reverberi (2001) "A network based model for traffic sensor location with implications on o/d matrix estimation", *Transportation Science* 35, pp 50–60.

Cascetta, E. (2001) *Transportation Systems Engineering: Theory and Methods*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Chen, A., Pravinvongvuth, S., Chootinan, P., Lee, M., and W. Recker (2007) "Strategies for selecting additional traffic counts for improving o–d trip table estimation", *Transportmetrica* 3, pp 191–211.

Chootinan, P., Chen, A., and H. Yang (2005) "A bi-objective traffic counting location problem for origin–destination trip table estimation", *Transportmetrica* 1, pp 65–80.

Confessore, G., Dell'Olmo, P., and M. Gentili (2005) "Experimental evaluation of approximation and heuristic algorithms for the dominating paths problem", *Computers & Operations Research* 32, pp 2583–2405.

Dial, R.B. (1971) "A probabilistic multipath assignment model which obviates path enumeration". *Transportation Research* 5, pp 83–111.

Edwards, H. (2000) "Traffic assignment analysis and o/d matrix calibration", report, KFB & VTI forskning/research 32, Kommunikationsforskningsberedningen, Stockholm, Sweden.

Ehlert, A., Bell, M.G.H., and S. Grosso (2006) "The optimisation of traffic count locations in road networks", *Transportation Research B* 40, pp 460–479.

Gan, L., Yang, H., and S.C. Wong (2005) "Traffic counting location and error bound in origin–destination matrix estimation problems", *Journal of Transportation Engineering* 131, pp 524–534.

Gentili, M., and P.B. Mirchandani (2005) "Locating active sensors on traffic networks", *Annals of Operations Research* 136, pp 229–257.

Hodgson, M.J. (1990) "A flow-capturing location-allocation model", *Gegraphical Analysis* 22, pp 270–279.

INRO (1990) "CNTPOST: A macro for optimal locations of count posts", Emme/2 News 10 October 1990, available at `http://emme2.spiess.ch/e2news/`, downloaded 2007-08-30.

Lam, W.H.K., and H.P. Lo (1990) "Accuracy of o–d estimates from traffic counts", *Traffic Engineering and Control* 31, pp 358–367.

Larsson, T., and M. Patriksson (1992) "Simplicial decomposition with disaggregated representation for the traffic assignment problem", *Transportation Science* 26, pp 4–17.

Larsson, T., Patriksson, M., and C. Rydergren (2004) "A column generation procedure for the side constrained traffic equilibrium problem", *Transportation Research B* 38, pp 17–38.

LeBlanc, L.J., Morlok, E.K., and W.P. Pierskalla (1975) "An efficient approach to solving the road network equilibrium traffic assignment problem", *Transportation Research* 9, pp 309–318.

Logie, M., and A. Hynd (1990) "MVESTM matrix estimation", *Traffic Engineering and Control* 31, pp 454–459.

Lotito, P., Mancinelli, E., Quadrat, J.P., and L. Wynter (2003) "The single class traffic assignment toolbox of Scilab: Ciudadsim. User's manual", available at `http://www-rocq.inria.fr/metalau/ciudadsim/`, last updated February 2003, downloaded 2007-08-30.

Meng, Q., Lee, D.-H., and R.L. Cheu (2005), "Counting the different efficient paths for transportation networks and its applications", *Journal of Advanced Transportation* 39, pp 193–220.

Ortúzar, J. de D., and L.G. Willumsen (2001) *Modelling Transport*, 3rd Edition, Wiley, West Sussex, United Kingdom.

Patriksson, M. (1994) *The Traffic Assignment Problem — Models and Methods*, VSP, Utrecht, The Netherlands.

Tatineni, M., Edwards, H., and D. Boyce (1998), "Comparison of disaggregate simplicial decomposition and Frank-Wolfe algorithms for user-optimal route choice", *Transportation Research Record* 1617, pp 157–162.

Yang, H., Iida, Y., and T. Sasaki (1991) "An analysis of the reliability of an origin–destination trip matrix estimated from traffic counts", *Transportation Research B* 25, pp 351–363.

Yang, B., and E. Miller-Hooks (2002) "Determining critical arcs for collecting real-time travel information", *Transportation Research Record* 1783, pp 34–41.

Yang, H., Yang, C., and L. Gan (2006) "Models and algorithms for the screen line-based traffic-counting location problems", *Computers & Operations Research* 33, pp 836–858.

Yang, H., and J. Zhou (1998) "Optimal traffic counting locations for origin–destination matrix estimation", *Transportation Research B* 32, pp 109–126.

Yim, P.K.N., and W.H.K. Lam (1998) "Evaluation of count location selection methods for estimation of o–d matrices", *Journal of Transportation Engineering* 124, pp 376–383.

van Zuylen, H., and L.G. Willumsen (1980) "The most likely trip matrix estimated from traffic counts", *Transportation Research B* 14, pp 281–293.

Appendix: Solving the OD-matrix estimation problem

For solving the problem (4), that is

$$\begin{aligned} \min \quad F(g,\hat{g}) &= \frac{1}{2}\sum_{i\in I}(g_i-\hat{g}_i)^2 \\ \text{s.t.} \sum_{i\in I}\hat{p}_{ai}g_i &= \bar{v}_a, \qquad \forall a\in\bar{A}, \\ g_i &\geq 0, \qquad \forall i \in I, \end{aligned}$$

we have used a Lagrangian dual algorithm, to be described below. (The reader who is unfamiliar to Lagrangian duality is referred to, for example, Bazaraa et al.; 1993.)

Letting $u=\{u_a\}$, where $a\in\bar{A}$, denote Lagrangian multipliers for the equality constraints, the Lagrangian dual function becomes

$$h(u) = \min_{g\geq 0}\left\{\frac{1}{2}\sum_{i\in I}(g_i-\hat{g}_i)^2+\sum_{a\in\bar{A}}u_a(\bar{v}_a-\sum_{i\in I}\hat{p}_{ai}g_i)\right\}$$

$$= \sum_{a \in \bar{A}} u_a \bar{v}_a + \sum_{i \in I} \min_{g_i \geq 0} \left\{ \frac{1}{2}(g_i - \hat{g}_i)^2 - \left(\sum_{a \in \bar{A}} u_a \hat{p}_{ai} \right) g_i \right\}.$$

For each OD-pair $i \in I$, the minimization problem is solved by

$$g_i(u) = \max \left\{ 0, \hat{g}_i + \sum_{a \in \bar{A}} u_a \hat{p}_{ai} \right\}. \tag{A.1}$$

The dual function is concave and everywhere finite and differentiable, with the partial derivatives

$$\frac{\partial h(u)}{\partial u_a} = \bar{v}_a - \sum_{i \in I} \hat{p}_{ai} g_i(u), \qquad \forall a \in \bar{A}. \tag{A.2}$$

The Lagrangian dual problem is to find

$$\max_u h(u). \tag{A.3}$$

If u^* solves the dual problem, then $g(u^*)$ solves the primal problem (4), that is, $g^* = g(u^*)$.

We have implemented the following steepest ascent algorithm for solving the dual problem (A.3), and thereby the OD-matrix estimation problem (4).

Algorithm

Step 0. Initialization

Set $u^{(0)} = 0$ and let $k = 0$ be the iteration counter.

Step 1. Search direction

Let the search direction be given by $d^{(k)} = \nabla h(u^{(k)})$, as defined by (A.2).

Step 2. Computation of steplength

Find an optimum steplength $t_k \geq 0$, such that $h(u^{(k)} + t \cdot d^{(k)})$ is maximized. Below we derive an analytical expression for t_k.

Step 3. Termination Criterion

The algorithm terminates if $||d^{(k)}||_2$ is small enough, or a maximum number of iterations has been reached. (The latter criterion is used to detect primal infeasibility.)

Step 4. Update

Set $u^{(k+1)} = u^{(k)} + t_k \cdot d^{(k)}$ and $k = k + 1$. Go to Step 1.

Derivation of optimum steplength

Given the search direction $d^{(k)} \neq 0$, we calculate a steplength $t = t_k \geq 0$ such that $\varphi(t) = h(u^{(k)} + t \cdot d^{(k)})$ is maximized. Since the Lagrangian dual function, h, is concave and everywhere differentiable, an optimum steplength is given by $\varphi'(t) = 0$. We note that, since $g(u^{(0)}) = \hat{g} \geq 0$, and since the objective is to minimize the deviation from $\hat{g}$, it is likely that the non-negativity of the OD-matrix is fulfilled throughout the iterations.

This observation leads us to the following steplength selection technique. Suppose that, for all $i \in I$, $g_i(u^{(k)} + t \cdot d^{(k)}) > 0$ holds for all values of $t \geq 0$ that are of interest in the steplength calculation. Then the solution (A.1) simplifies to

$$g_i(u^{(k)} + t \cdot d^{(k)}) = \hat{g}_i + \sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \hat{p}_{ai}, \qquad \forall i \in I,$$

for all relevant values of the steplength t, and we obtain

$$\begin{aligned}
\varphi(t) &= \sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \bar{v}_a + \sum_{i \in I} \hat{g}_i^2 + \\
&\sum_{i \in I} \left\{ \frac{1}{2} \left[\hat{g}_i + \sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \hat{p}_{ai} - \hat{g}_i \right]^2 - \sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \hat{p}_{ai} \left[\hat{g}_i + \sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \hat{p}_{ai} \right] \right\} \\
\varphi(t) &= \sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \bar{v}_a + \sum_{i \in I} \hat{g}_i^2 + \\
&\sum_{i \in I} \left\{ \frac{1}{2} \left[\sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \hat{p}_{ai} \right]^2 - \sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \hat{p}_{ai} \left[\hat{g}_i + \sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \hat{p}_{ai} \right] \right\} \\
\varphi(t) &= \sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \bar{v}_a + \sum_{i \in I} \hat{g}_i^2 + \\
&\sum_{i \in I} \left\{ \frac{1}{2} \left[\sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \hat{p}_{ai} \right]^2 - \sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \hat{p}_{ai} \hat{g}_i - \left[\sum_{a \in \bar{A}} (u_a^{(k)} + t \cdot d_a^{(k)}) \hat{p}_{ai} \right]^2 \right\}
\end{aligned}$$

$$\varphi(t) = \sum_{a\in\bar{A}}(u_a^{(k)} + t\cdot d_a^{(k)})\bar{v}_a + \sum_{i\in I}\hat{g}_i^2 - \sum_{i\in I}\left\{\sum_{a\in\bar{A}}(u_a^{(k)} + t\cdot d_a^{(k)})\hat{p}_{ai}\hat{g}_i + \frac{1}{2}\left[\sum_{a\in\bar{A}}(u_a^{(k)} + t\cdot d_a^{(k)})\hat{p}_{ai}\right]^2\right\}.$$

Hence,

$$\begin{aligned}
\varphi'(t) &= \sum_{a\in\bar{A}} d_a^{(k)}\bar{v}_a - \sum_{i\in I}\left\{\sum_{a\in\bar{A}} d_a^{(k)}\hat{p}_{ai}\hat{g}_i + \left[\sum_{a\in\bar{A}}(u_a^{(k)} + t\cdot d_a^{(k)})\hat{p}_{ai}\right]\sum_{a\in\bar{A}} d_a^{(k)}\hat{p}_{ai}\right\} \\
&= \sum_{a\in\bar{A}} d_a^{(k)}\bar{v}_a - \sum_{i\in I}\sum_{a\in\bar{A}} d_a^{(k)}\hat{p}_{ai}\hat{g}_i - \sum_{i\in I}\left[\sum_{a\in\bar{A}} u_a^{(k)}\hat{p}_{ai}\right]\left[\sum_{a\in\bar{A}} d_a^{(k)}\hat{p}_{ai}\right] - t\cdot\sum_{i\in I}\left[\sum_{a\in\bar{A}} d_a^{(k)}\hat{p}_{ai}\right]^2 \\
&= \sum_{a\in\bar{A}} d_a^{(k)}\bar{v}_a - \sum_{i\in I}\sum_{a\in\bar{A}} d_a^{(k)}\hat{p}_{ai}\underbrace{\left[\hat{g}_i + \sum_{a\in\bar{A}} u_a^{(k)}\hat{p}_{ai}\right]}_{g_i^{(k)}} - t\cdot\sum_{i\in I}\left[\sum_{a\in\bar{A}} d_a^{(k)}\hat{p}_{ai}\right]^2 \\
&= \sum_{a\in\bar{A}} d_a^{(k)}\underbrace{\left[\bar{v}_a - \sum_{i\in I}\hat{p}_{ai}g_i^{(k)}\right]}_{d_a^{(k)}} - t\cdot\sum_{i\in I} + \left[\sum_{a\in\bar{A}} d_a^{(k)}\hat{p}_{ai}\right]^2 \\
&= \sum_{a\in\bar{A}}\left[d_a^{(k)}\right]^2 - t\cdot\sum_{i\in I}\left[\sum_{a\in\bar{A}} d_a^{(k)}\hat{p}_{ai}\right]^2.
\end{aligned}$$

By setting $\varphi'(t) = 0$ we can calculate the optimum steplength t_k as

$$t_k = \frac{\sum_{a\in\bar{A}}\left[d_a^{(k)}\right]^2}{\sum_{i\in I}\left[\sum_{a\in\bar{A}} d_a^{(k)}\hat{p}_{ai}\right]^2}.$$

We note that the denominator in this expression could become zero even though the numerator is not, but only in some rare cases, which did not occur in any of our experiments.

As a consequence of the supposition it is possible that the steplength t_k will be too large, so that $h(u^{(k)} + t_k\cdot d^{(k)}) \not> h(u^{(k)})$ holds. If this happens, we shorten the steplength, by iteratively halving it, until $h(u^{(k)} + t_k\cdot d^{(k)}) > h(u^{(k)})$ holds. Since $d^{(k)}$ is an ascent direction, this will be the case for a sufficiently small value of t_k. (In practice, any shortening of the steplength is seldomly needed.)