Estimation of dynamic origin-destination matrices using linear assignment matrix approximations

Tomer Toledo and Tanya Kolechkina

***Abstract*—This paper presents a general solution scheme for the problem of off-line estimation of dynamic Origin-Destination (OD) demand matrices using traffic counts on some of the network links and historical demand information. The proposed method uses linear approximations of the assignment matrix, which maps the OD demand to link traffic counts. Several iterative algorithms that are based on this scheme are developed. The various algorithms are implemented in a tool that uses the mesoscopic traffic simulation model Mezzo to conduct network loadings. A case study network in Stockholm Sweden is used to test the proposed algorithms and to compare their performance to current state-of-the-art methods. The results demonstrate the applicability of the proposed methodology to efficiently obtain dynamic OD demand estimates for large and complex networks and that, computationally, this methodology outperforms existing methods.**

***Index Terms*—Dynamic traffic assignment, Origin-Destination matrix estimation, Assignment matrix**

I. Introduction

In recent years, there have been significant advances in the development of Dynamic Traffic Assignment (DTA) and traffic simulation models, which predict time-dependent traffic conditions on a road network. An important input to these models is the demand for travel, which is commonly represented by Origin-Destination (OD) demand matrices. Collecting OD information directly by conducting surveys is time consuming and cost expensive. Moreover, the measurements may quickly become outdated. Therefore, OD demand matrices are commonly estimated using traffic counts, collected from the links of the network and effectively combined with available OD information (e.g., derived from direct measurements or from previous estimates).

OD estimation methods for the static problem (e.g. [1]-[5]) predict trip rates over a long period of time (such as a peak period), within which conditions on the network are assumed to be stationary. Dynamic OD Estimation (DODE) models (e.g. [6]-[9]) relax the assumptions of stationary demand, represent traffic dynamics and incorporate stochasticity in these elements. The estimated OD matrices are therefore more suitable as inputs to DTA and traffic simulation models. While the problem formulation has been well-established in the literature, there is still a need for efficient algorithms for its solution in large-scale congested networks. This is demonstrated, for example, in recent work by Cipriani *et al.* [10]. The authors present a modification to the Stochastic Perturbation Simultaneous Approximation (SPSA) algorithm, a state-of-the-art solution approach to the DODE problem. Their algorithm required 15 hours to estimate an OD demand with four time slices for the Calgary network (734 links, 221 nodes and 77 centroids).

This paper presents a general solution scheme for the DODE problem. A critical construct in DODE is an assignment matrix, which maps OD demand flows to traffic counts at sensor locations. In congested networks, the assignment proportions depend on the unknown time-dependent OD demands. The methods proposed in this paper are based on use of linear approximations of the assignment matrix in the optimization iterations. Several iterative algorithms, based on this scheme, that differ in the search direction they use are developed. The algorithms are tested using the mesoscopic traffic simulation model Mezzo [11] on a network in the Stockholm area. The case study demonstrates the computational efficiency of these algorithms compared to current state-of-the-art approaches for large-scale networks.

The rest of this paper is organized as follows. In the next section, the DODE estimation problem is formulated mathematically. Algorithms for the solution of this problem that have been proposed in the literature are presented in section III. The solution scheme proposed in this work is presented in section IV. The details of this algorithm and its implementation are presented in sections V and VI, respectively. A case study demonstrating the proposed algorithms is presented in section VII. Finally, conclusions are presented in Section VIII.

Manuscript received June 17, 2012. This research was supported in part by the Israel Science Foundation (grant 1218/07) and Israel Ministry of Industry, Trade and Labor. It also benefited from participation in the EC MULTITUDE COST Action (TU0903).

T. Toledo is with the Faculty of Civil and Environmental Engineering at the Technion – Israel Institute of Technology, Haifa 32000 Israel (Corresponding author, phone: +972-4-829-3080; fax: +972-4-829-5708; e-mail: toledo@ technion.ac.il).

T. Kolechkina was with the Technion – Israel institute of Technology, Haifa 32000 Israel. She is now with the Department of Civil and Environmental Engineering, Carlton University, Ottawa ON, K1S 5B6 Canada. (e-mail: Tanya.kolechkina@gmail.com).

II. PROBLEM FORMULATION

Consider a transportation network represented by a directed graph $G(C,L)$, where C is a set of nodes and L is a set of links. $L' \subseteq L$ is a subset of links equipped with sensors. The OD matrix $X = \{x_{nr}\}$ defines the demand for travel for each OD pair $n \in N$ in each departure time period $r \in R$. N and R are the number of OD pairs and departure time periods, respectively. The data available for the demand estimation includes both direct and indirect information. A-priori information on the OD matrix (e.g., from direct measurements, previous studies or application of other, possibly static, traffic assignment models) is captured by a target (historic) OD matrix, $X^H = \{x_{nr}^H\}$. The indirect measurements include link flow observations, $\tilde{Y} = \{\tilde{y}_{lt}\}$ on a subset of link in the network $l \in L'$ during time interval $t \in T$. It is assumed that T and R describe the same period of interest, but their decompositions to time intervals may be different. It is also assumed that X, X^H and $\tilde{Y}$ are arranged as column vectors.

Reference [6] formulated the DODE problem as a constrained optimization problem. The objective function for this problem $Z(X)$ includes two parts. The first part measures the distance between the estimated OD matrix and the historic OD matrix. The second part measures the distance between the observed link flows and those predicted by the model when the estimated OD demand flows are assigned to the network. This formulation is given by:

$$\min_{X \geq 0} Z(X) = w_1 F_1\left(X, X^H\right) + w_2 F_2\left(Y(X), \tilde{Y}\right) \quad (1)$$

F_1 and F_2 are distance functions. w_1 and w_2 are weighting factors that reflect the relative uncertainty in the information contained in X^H and $\tilde{Y}$, respectively. This uncertainty may result from sensor measurement errors and from modeling errors. Given multiple measurements of the traffic counts and demands, the weight factors could be estimated by the inverse of the standard deviations of these measurements. $Y(X) = \{y_{lt}\}$ are the link flows predicted by an assignment (or loading) of the demand to the network. The assignment model may be an optimization problem in itself. In this case, the overall formulation becomes a bi-level optimization problem.

The mapping of the demand to the traffic counts, $Y(X)$, may be expressed as a proportion of the OD demand that passes through a count location:

$$Y = A(X)X \quad (2)$$

$A = \{a_{nr,lt}\}$ is the assignment matrix. $a_{nr,lt}$ is the fraction of the OD demand x_{nr} that passes link l during observation period t. These fractions depend on the route choices and on the travel times from the origins to counting points. In turn, these depend on the congestion in the network, which depend on the demand for travel. Thus, the assignment matrix depends on the demand to take into account the effect of congestion.

The functional form of F_1 and F_2 depends on the assumptions made regarding the structure of the errors in the traffic counts and travel demands. See [6] and [12] for alternative formulations. Assuming that errors are independently and identically normally distributed, and incorporating (2), the DODE problem becomes:

$$\min_{X \geq 0} Z(X) = \frac{w}{2}\left\|X - X^H\right\|^2 + \frac{1}{2}\left\|A(X)X - \tilde{Y}\right\|^2 \quad (3)$$

$w = \frac{w_1}{w_2}$ is a weighting factor that captures the relative variability in the information contained in the historic OD matrix compared to that in the traffic counts. Note that the optimal solution is insensitive to scalar multiplications, and so is not affected by the division by w_2 or by a factor of two, which simplify the notation and the derivative expressions that will follow.

III. SOLUTION ALGORITHMS

Several methods to solve the DODE problem in (3) have been proposed. Assuming that the assignment matrix is fixed and exogenously known, [6] and [13] solved the DODE problem as a quadratic optimization problem using standard gradient methods. However, in most cases assignment matrices are not a-priori known and are not fixed, due to the effects of congestion that depends on the travel demand. In congested network, changes in the demand affect travel times. In turn, travel times affect the route choices and travel times from origins to counting points that determine the assignment fractions. Several authors (e.g. [7], [8], [14]-[17]) applied algorithms that iterate between two basic steps: (i) An assignment step, in which a given demand matrix (the solution of the previous iteration) is assigned to the network to yield an assignment matrix, and (ii) an optimization step, in which an auxiliary OD demand solution is obtained through the optimization of a quadratic objective function. This quadratic sub-problem results from fixing the assignment matrix in (3) to its value in the current iteration. The new OD matrix is obtained by moving from the current solution in a search direction, which is defined by the difference between the auxiliary solution and that of the previous iteration. The step size in this search direction may be pre-defined by simple functional iterations, the method of successive averages (MSA) and similar weighing methods. Alternatively, exact or approximate line search methods that minimize the value of the objective function along the descent direction may be used.

Lundgren *et al.* [18] pointed out that the assumption of a constant proportional assignment may not correctly capture the

marginal effect of a change in the OD matrix on traffic counts, which may reduce the efficiency of the iterations. They propose an estimation method that relaxes this assumption in the step size calculation using an adaptation of the approach proposed in [5] and [19] for the deterministic and stochastic static cases, respectively. In their method, the auxiliary solution is computed, as before, under the assumption that the assignment matrix is fixed at its current values. The auxiliary demand is then assigned to the network to obtain the link counts at measurement locations. The two sets of demands and corresponding counts (from the current and auxiliary demand solutions) are then used to derive linear approximations of the link flows as functions of the OD demands. Substituting these approximations in the objective function, they obtain a modified form of the objective that accounts for the local dependence of the assignment matrix on the demand. The new iteration solution is obtained by a steepest descent algorithm. The search direction is defined, as usual, by the difference between the current and auxiliary solutions. The optimal step size is calculated as a closed form expression derived from the modified objective function. This algorithm overcomes the difficulty associated with assuming a constant assignment matrix. However, it requires an additional assignment (of the auxiliary solution) in every iteration. This additional assignment reduces its attractiveness for large-scale problems where the assignment step is computationally expensive. In section VI modified version of this algorithm that avoids the additional assignment step is presented as a special case of the solution strategy proposed here.

Several authors proposed meta-heuristic approaches that do not rely on the assignment matrix at all. These methods include evolutionary algorithms ([20]-[24]) and simulated annealing [25]. Genetic algorithms use a population of candidate solutions in each generation (iteration). New candidate solutions are generated by functions that mimic biological processes of reproduction and survival. Simulated annealing is a stochastic optimization method that probabilistically advances from a current solution to a randomly generated neighboring one. An important advantage of these methods is that they are generally able to find global and not only local optima. However, they usually require a large number of objective evaluations (and there assignments), which in the context of DODE can be computationally expensive.

The most studied assignment matrix-free method is the Simultaneous Perturbation Stochastic Approximation (SPSA) method ([26]-[31], [10]). This algorithm has a similar structure to the steepest descent method in that it calculates a negative gradient search direction and then determines the step size in this direction.

$$X^{(k+1)} = X^{(k)} + \theta^{(k)} d^{(k)} \tag{4}$$

$X^{(k)}$ is the OD demand in iteration k. $\theta^{(k)}$ and $d^{(k)}$ are the step size and search direction in iteration k, respectively.

In finite differences, the gradient is approximated by perturbing each element in the OD demand, one at a time. For central differences, this requires $2NR$ objective function evaluations. In SPSA the search direction is found by perturbing all elements in $X^{(k)}$ simultaneously instead of one-at-a-time. Thus, the elements of the search direction are calculated as:

$$d_i^{(k)} = \frac{Z\left(X^{(k)} + c^{(k)}\Delta^{(k)}\right) - Z\left(X^{(k)} - c^{(k)}\Delta^{(k)}\right)}{2c^{(k)}\delta_i^{(k)}} \tag{5}$$

$d_i^{(k)}$ is the ith element of the search direction d in iteration k. $c^{(k)}$ is a small positive perturbation constant. $\Delta^{(k)} = \left\{\delta_i^{(k)}, \ldots, \delta_i^{(k)}, \ldots, \delta_i^{(k)}\right\}$ is a random perturbation vector.

This way, only two assignments are required in order to compute the search direction in each iteration regardless of the problem dimension. The original algorithm uses pre-defined step sizes, but line search minimization methods have also been used. While the algorithm simplicity is appealing, it requires two assignments in each iteration compared to only one for most algorithms described above. Furthermore, the performance of the algorithm depends on the values of several parameters that control the perturbation constant and the step size. However, useful parameter values are problem-specific and may be difficult to find [27].

IV. Proposed Solution strategy

The work presented in this paper relies on linear approximation of the assignment matrix in every iteration, which relaxes the assumption of constant assignment proportions and explicitly accounts for congestion effects. Unlike [18] the linear assignment approximation is used not only in the step size calculation, but also to obtain the descent directions. Using the first order Taylor expansion around the current solution, the elements of the linear assignment function in iteration k are defined by:

$$\begin{aligned} a_{nr,lt}(X) &= a_{nr,lt}\left(X^{(k)}\right) + \sum_{n'r'} \frac{\partial a_{nr,lt,n'r'}}{\partial x_{n'r'}}\left(X^{(k)}\right)\left(X - X^{(k)}\right) \\ &= \gamma_{nr,lt}^{(k)} + \alpha_{nr,lt,n'r'}^{(k)\prime} X \end{aligned} \tag{6}$$

$\gamma_{nr,lt}^{(k)}$ is a constant assignment proportion at the current solution. $\alpha_{nr,lt,n'r'}^{(k)}$ is a vector whose entries capture the marginal effect of $x_{n'r'}$ on the assignment proportion $a_{nr,lt}$ at the current solution. The re-writing of the linear approximation using α and γ will be useful later, when these parameters will be estimated via regression.

The linear approximation above involves $NR+1$ parameters for each entry in the assignment matrix. The total of $(NR \cdot LT)(NR+1)$ parameters in the assignment matrix is too large to realistically be estimated. However, by neglecting the effect of changes in the demand in other OD pairs on the

assignment fractions of OD pair nr, the number of parameters is reduced to $2(NR \cdot LT)$. This is equivalent to relaxing the standard assumption of constant assignment fractions only for the effect of the demand on the same OD pair, which can be expected to have the largest marginal effect on the assignment fractions. This assumption also makes the assignment functions separable in the OD demands, which will be useful in estimating their parameters. Under this assumption, the assignment function is given by:

$$a_{nr,lt}(X) = \gamma_{nr,lt}^{(k)} + \beta_{nr,lt,nr}^{(k)} x_{nr} \quad (7)$$

and in matrix form:

$$A(X) = \Gamma + \mathrm{B}\,diag(X) \quad (8)$$

$\Gamma = \{\gamma_{nr,lt}\}$ is the matrix of assignment proportion constants. $\mathrm{B} = \{\beta_{nr,lt} \equiv \alpha_{nr,lt,nr}\}$ is a matrix that has the same dimensions as the assignment matrix. $diag(X)$ is a diagonal matrix, with entries corresponding to the elements of X.

By substituting (8) in the objective function (3) it is approximated, in each iteration, by:

$$\min_{X \geq 0} Z(X) = \frac{w}{2}\left\| X - X^{H} \right\|^2 + \frac{1}{2}\left\| \left(\Gamma + \mathrm{B}\,diag(X)\right) X - \tilde{Y} \right\|^2 \quad (9)$$

A generic scheme based on iterative solution of the approximated problem above is shown in

Fig. 1 and defined by the following steps:

Step 0. *Initialization*: Set an initial estimate of the OD matrix $X^{(0)}$. Set the iteration counter $k = 0$.

Step 1. *Assignment*: Load $X^{(k)}$ to the network to obtain $Y^{(k)}$ and $A^{(k)}$.

Step 2. *Assignment matrix approximation*: Estimate the parameters $\Gamma^{(k)}$ and $\mathrm{B}^{(k)}$.

Step 3. *Search direction*: Use the approximated objective function (9) to calculate a search direction $d^{(k)}$.

Step 4. *Step size*: Use the approximated objective function (9) to compute a step size $\theta^{(k)}$ in the search direction.

Step 5. Update: Calculate the new OD matrix estimate $X^{(k+1)} = X^{(k)} + \theta^{(k)} d^{(k)}$.

Step 6. *Convergence test*: If the termination criteria holds then stop. Otherwise, set $k := k+1$ and go back to step 1.

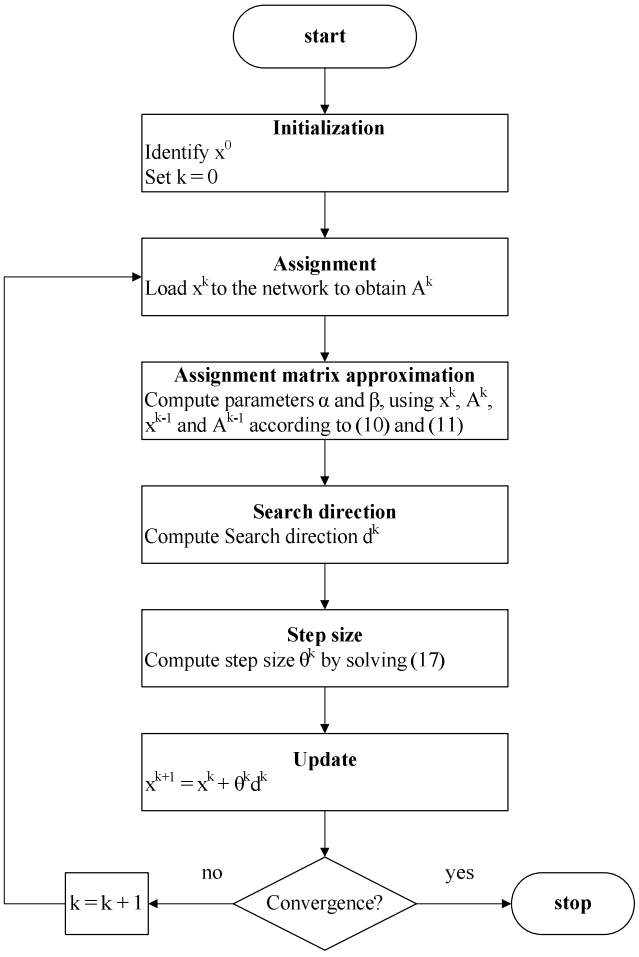

Fig. 1. Generic algorithm based on the proposed solution strategy

V. Algorithmic Details

The solution scheme proposed above, based on the linear assignment approximation, may be implemented using a wide range of algorithms that would differ in the specifics of the search direction and step sizes. The various steps of the solution algorithm are detailed next.

A. Assignment matrix approximation

The assignment matrix is approximated with the linear function in (8), which requires the estimation of the unknown parameter matrices Γ and B. As noted above, this system of equations is separable in the OD demands. Therefore, given the assignments of the demands from *p* previous iterations, for each element in the assignment matrix, the pair $\gamma_{nr,lt}^{(k)}$ and $\beta_{nr,lt}^{(k)}$ may be obtained separately solving the following systems:

$$a_{nr,lt}^{(k-p)} = \gamma_{nr,lt}^{(k)} + \beta_{nr,lt}^{(k)} x_{nr}^{(k-p)} + \varepsilon_{nr,lt}^{(k-p)} \quad \forall p = 0,1,...,P-1 \quad (10)$$

$\varepsilon_{nr,lt}^{(k-p)}$ is an error term. P is the number of iterations that are used to estimate the assignment matrix.

These systems of equations may be solved in the least squares sense for $P \geq 2$ by minimizing the following objective function:

$$\min_{\gamma_{nr,lt}^{(k)},\beta_{nr,lt}^{(k)}} \sum_{p=0}^{P-1}\left(\varepsilon^{(k-p)}\right)^2 = \sum_{p=0}^{P-1}\left(a_{nr,lt}^{(k-p)} - \gamma_{nr,lt}^{(k)} - \beta_{nr,lt}^{(k)} x_{nr}^{(k-p)}\right)^2 \quad \forall nr, lt \tag{11}$$

The closed-form solution to this problem is given by:

$$\beta_{nr,lt}^{(k)} = \frac{P\sum_p a_{nr,lt}^{(k-p)} x_{nr}^{(k-p)} - \sum_p a_{nr,lt}^{(k-p)} \sum_p x_{nr}^{(k-p)}}{P\sum_p \left(x_{nr}^{(k-p)}\right)^2 - \left(\sum_p x_{nr}^{(k-p)}\right)^2} \quad \forall nr, lt \tag{12}$$

$$\gamma_{nr,lt}^{(k)} = \frac{1}{P}\left(\sum_p a_{nr,lt}^{(k-p)} - \beta_{nr,lt}^{(k)} \sum_p x_{nr}^{(k-p)}\right) \quad \forall nr, lt \tag{13}$$

The expressions above cannot be used in the first iteration. In this iteration, a constant assignment matrix may be estimated by assigning the initial OD matrix to the network. Alternatively, the available historic OD matrix X^H, which is used as the initial demand estimate $X^{(0)}$, may be modified (e.g. multiplied by a scalar) to obtain an additional solution $X^{(1)}$. The assignment of this solution to the network will enable the assignment matrix approximation according to (12) and (13). It should also be noted that the assignment fractions used in the approximation are obtained as the results of a DTA or traffic simulation model. These models may be stochastic, in which case, multiple replications will be needed for more accurate estimates.

B. Search direction

Various methods may be used to define a search direction for the approximated problem defined in (9). These methods would typically involve the use of the gradient and Hessian of the objective function at the current solution. The elements of the gradient and the Hessian are given by:

$$\frac{\partial Z}{\partial x_{nr}} = w\left(x_{nr} - x_{nr}^H\right) + \sum_{lt}\left(\left(\sum_{n'r'}\left(\gamma_{lt,n'r'} + \beta_{lt,n'r'} x_{n'r'}\right)x_{n'r'} - y_{lt}\right)\left(\gamma_{lt,nr} + 2\beta_{lt,nr} x_{nr}\right)\right) \tag{14}$$

$$\frac{\partial^2 Z}{\partial x_{nr}^{\ 2}} = w + \sum_{lt}\left(\begin{array}{l} 2\beta_{lt,nr}\left(\sum_{n'r'}\left(\gamma_{lt,n'r'} + \beta_{lt,n'r'} x_{n'r'}\right)x_{n'r'} - y_{lt}\right) \\ +\left(\gamma_{lt,nr} + 2\beta_{lt,nr} x_{nr}\right)^2 \end{array}\right) \tag{15}$$

$$\frac{\partial Z^2}{\partial x_{nr}\partial x_{n'r'}} = \sum_{lt}\left(\gamma_{lt,nr} + 2\beta_{lt,nr} x_{nr}\right)\left(\gamma_{lt,n'r'} + 2\beta_{lt,n'r'} x_{n'r'}\right) \tag{16}$$

Or in matrix form:

$$\nabla Z = w\left(X - X^H\right) + \left(\Gamma + 2\mathrm{B}diag\left(X\right)\right)'\left(\left(\Gamma + \mathrm{B}diag\left(X\right)\right)X - \tilde{Y}\right) \tag{17}$$

$$H = wI + 2diag\left(\mathrm{B}'\left(\left(\Gamma + \mathrm{B}diag\left(X\right)\right)X - \tilde{Y}\right)\right) + \left(\Gamma + 2\mathrm{B}diag\left(X\right)\right)'\left(\Gamma + \mathrm{B}diag\left(X\right)\right) \tag{18}$$

C. Step size

After the search direction has been determined, a step size in this direction needs to be calculated. Using the linear approximation to the assignment matrix, an optimal step size may be calculated as the solution of a one-dimensional minimization problem:

$$\theta^{(k)} = \arg\min_{\theta} Z\left(X = X^{(k)} + \theta d^{(k)}\right) = \frac{w}{2}\left\|X - X^H\right\|^2 + \frac{1}{2}\left\|\left(\Gamma^{(k)} + \mathrm{B}^{(k)}diag\left(X\right)\right)X - \tilde{Y}\right\|^2 \tag{19}$$

$\theta^{(k)}$ and $d^{(k)}$ are the step size and search direction in iteration k, respectively.

The line search problem above may be solved exactly by finding the root of the derivative of the objective function, which is a cubic function of θ. Inexact solutions may also be calculated, for example using the Armijo rule.

VI. Implementation

The proposed algorithm was implemented in MATLAB, and using the Mezzo model [11] for traffic assignment. Mezzo is a mesoscopic event-based traffic simulation model. It models vehicles individually, but does not represent their movement in detail. The travel demand in Mezzo is represented by a time-sliced OD matrix. Vehicle generation is done for each OD pair separately with inter-arrival times that follow a negative exponential distribution. The model incorporates an iterative dynamic traffic assignment procedure, which updates the set of routes, and the travel times after each loading. Mezzo is an open-source model, which facilitated its use within the DODE framework. However, clearly, the solution algorithms are independent of any specific traffic simulation model.

The structure of the DODE implementation using MATLAB and Mezzo is shown in Fig. 2. MATLAB serves as the DODE engine, which reads and writes Mezzo files, converts formats, performs all the estimation steps and calls Mezzo as a subroutine. Thus, other DTA models may also be used with changes only to the reading, writing and DTA calling functions.

Several specific algorithms were implemented within this framework. As discussed above they differ only in the search direction they utilize.

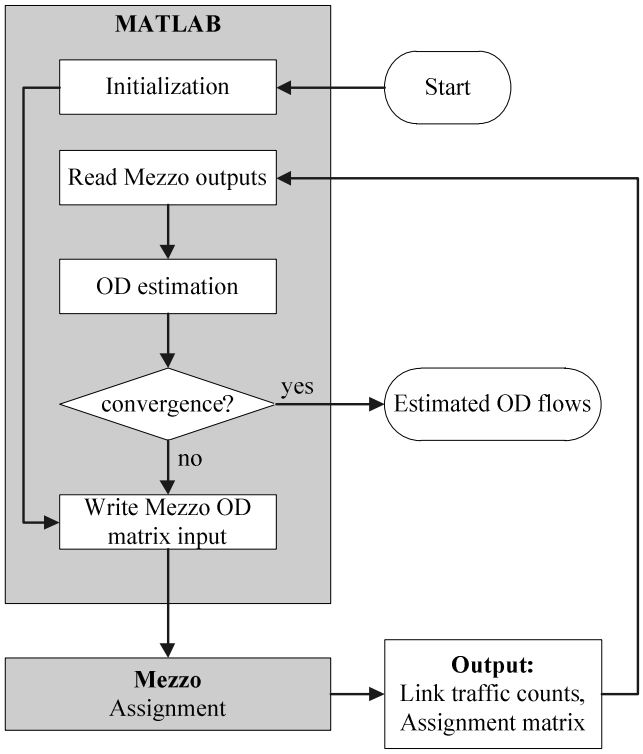

Fig. 2. Structure of the MATLAB-Mezzo DODE implementation

A. Gradient and relative gradient methods

These algorithms are based on the steepest descent method, which uses the gradient as the search direction:

$$d^{(k)} = -\nabla Z\left(X^{(k)}\right) \tag{20}$$

$\nabla Z\left(X^{(k)}\right)$ is given by (14) or (17).

The relative gradient variant is adapted from the algorithm proposed by Spiess [32] for the static OD estimation problem. It avoids drastic changes in the demand and has been found to work well in practice (e.g. [33]). In this method, the gradient direction is modified through entry-wise (Hadamard) multiplication by the current solution to capture the relative change in the demand:

$$d^{(k)} = -X^{(k)} \circ \nabla Z\left(X^{(k)}\right) \tag{21}$$

B. Modified Lundgren et al.'s method

This algorithm is based on the one proposed in [18], but with changed assumptions regarding the assignment matrix. The search direction is that of Newton's method, but using a diagonal approximation of the Hessian in order to simplify its inversion:

$$d^{(k)} = -\hat{H}^{-1}\left(X^{(k)}\right)\nabla Z\left(X^{(k)}\right) \tag{22}$$

$\hat{H}^{-1}\left(X^{(k)}\right)$ is the inverse of the diagonal approximation of the Hessian with the second order derivatives $\frac{\partial^2 Z}{\partial x_{nr}^{\ 2}}$ at the current solution, as given in (15).

Note that this modified algorithm requires only one simulation evaluation in each iteration, compared to two evaluations in the original algorithm.

C. Quasi-Newton methods

In these methods, an approximation of the inverse of the Hessian is used with Newton's direction:

$$d^{(k)} = -B^{(k)}\nabla Z\left(X^{(k)}\right) \tag{23}$$

$B^{(k)}$ is a positive definite matrix that approximates the inverse of the Hessian. It is adjusted in every iteration, for example, with the BFGS update:

$$B^{(k)} = \left(I - \frac{P^{(k)}Q^{(k)\prime}}{Q^{(k)\prime}P^{(k)}}\right)B^{(k-1)}\left(I - \frac{Q^{(k)}P^{(k)\prime}}{Q^{(k)\prime}P^{(k)}}\right) + \frac{P^{(k)}P^{(k)\prime}}{Q^{(k)\prime}P^{(k)}} \tag{24}$$

$$P^{(k)} = X^{(k)} - X^{(k-1)} \text{ and } Q^{(k)} = \nabla Z\left(X^{(k)}\right) - \nabla Z\left(X^{(k-1)}\right)$$

As with the steepest descent method, a relative variant following [33] may also be used:

$$d^{(k)} = -X^{(k)} \circ B^{(k)}\nabla Z\left(X^{(k)}\right) \tag{25}$$

VII. Case Study

The proposed solution algorithms were applied in a case study using the Mezzo model of the Sodermalm network in Stockholm, Sweden. The network is shown in Fig. 3. It consists of 577 nodes and 1101 links. The OD matrix contains 347 OD pairs with travel demand specified for four 15-minute departure time intervals (1388 demand values in total) during the morning peak between 7AM and 8AM. The overall demand in the network in this 1 hour period is 69,008 vehicle trips. There are 55 link count locations (yielding 220 15-minutes counts in total) on this network. Thus, $|C| = 577, |L| = 1101, |N| = 347, |R| = |T| = 4, |L'| = 55$.

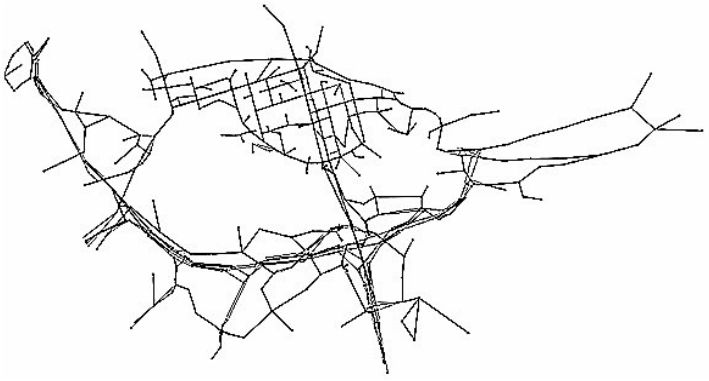

Fig. 3. Sodermalm network in Mezzo

An estimate of the OD matrix for this network was available from previous studies. This matrix was used as the "true" OD matrix. The corresponding traffic counts were unknown. They were obtained by assignment of the "true" demand using Mezzo. The historic (seed) OD demand matrix was then obtained by multiplication of the "true" OD matrix by a factor of 0.6. The DODE problem was solved for four values of the weighting factor $w = \{0.001, 0.01, 0.1, 1\}$. Within assignment steps, three Mezzo replications were conducted. The solutions of three previous iterations ($P = 3$) were used in the assignment matrix approximation step.

Fig. 4 through Fig. 7 show the progress of the objective function values through the iterations for $w = 0.001, 0.01, 0.1, 1$, respectively. The results are presented for three algorithms that implement the proposed scheme: based on the relative gradient method (21), a modified Lundgren *et al.*'s method (22) and a quasi-Newton method (23) , and, for comparison, two current state-of-the-art algorithms that were described above: Lundgren *et al.* [18] and SPSA [28]. The results indicate that the algorithms that use the linear approximation of the assignment matrix significantly outperform the SPSA and Lundgren *et al.*'s algorithms throughout the iterations. Within the algorithms based on the linear approximation, the ones that use the relative gradient and the quasi Newton search directions have the best performance. It may be that the approximation of the Hessian matrix with its diagonal elements within the modified Lundgren *et al.*'s algorithm does not yield good search directions. Furthermore, the numbers of iterations needed in order to achieve convergence are significantly lower with the algorithms based on the linear assignment matrix approximation, as shown in Fig. 8. For the purpose of this figure, convergence was defined as reaching an objective function value that is twice the lowest value obtained (by any of the algorithms) within 250 iterations. Note that the SPSA method did not reach convergence within 250 iterations in any of the cases, which seems to indicate that the search directions that the algorithm produces are increasingly inefficient as the algorithm progresses. Lundgren *et al.*'s algorithm did not converge in two of the four cases. Fig. 9 shows run times to convergence in minutes. The algorithms based on the linear approximation of the assignment matrix require one Mezzo simulation evaluation in each iteration, compared to two simulation evaluations in each iteration for the SPSA and Lundgren *et al.*'s algorithms. As a result, the gain in terms of run times obtained by the use of the linear assignment matrix approximation is even larger compared to the gain in terms of numbers of iterations. Furthermore, this gain is expected to increase for larger and more complex networks as the proportion of the overall computational effort taken up by the simulation evaluations should also increase.

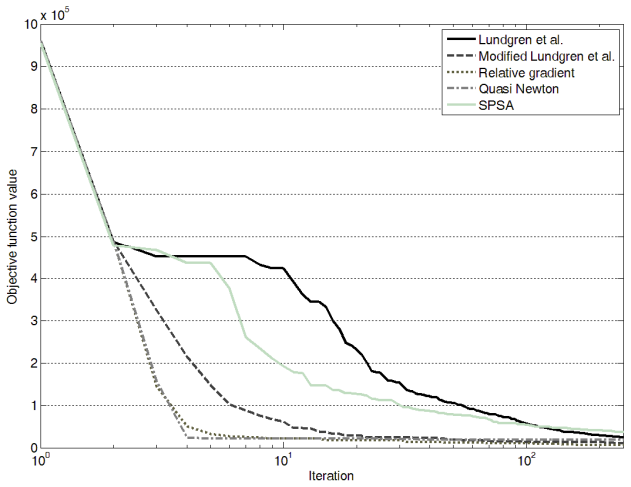

Fig. 4. Performance of various algorithms for $w = 0.001$

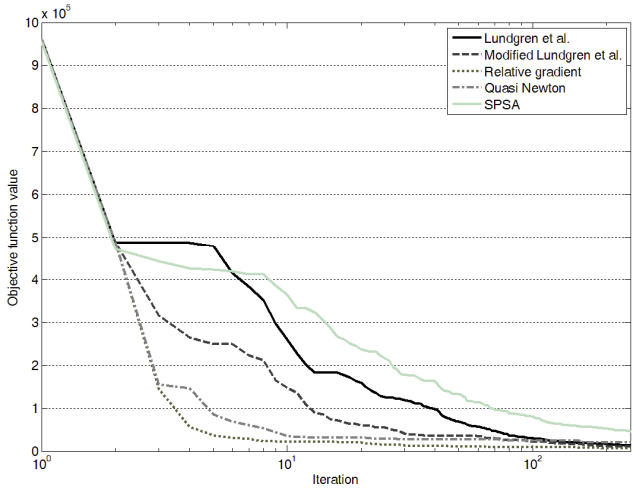

Fig. 5. Performance of various algorithms for $w = 0.01$

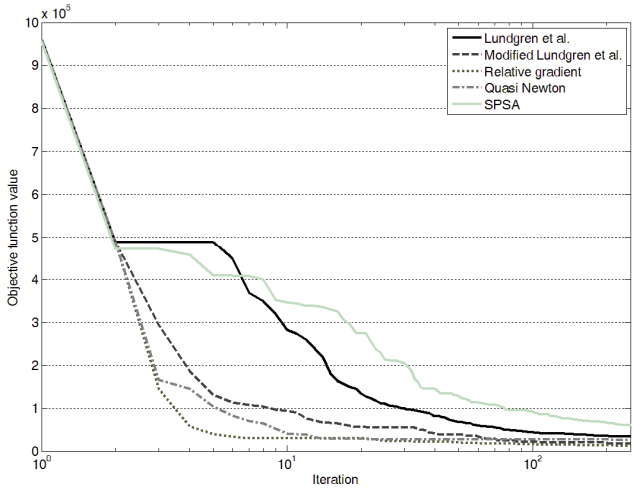

Fig. 6. Performance of various algorithms for $w = 0.1$

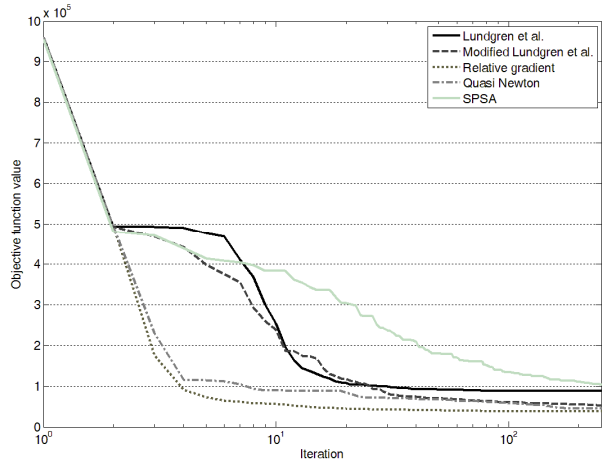

Fig. 7. Performance of various algorithms for $w = 1$

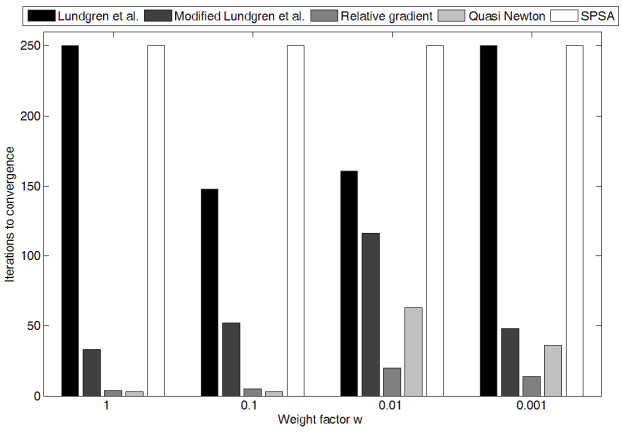

Fig. 8. The number of iterations to convergence for the various algorithms and values of the weight factor w

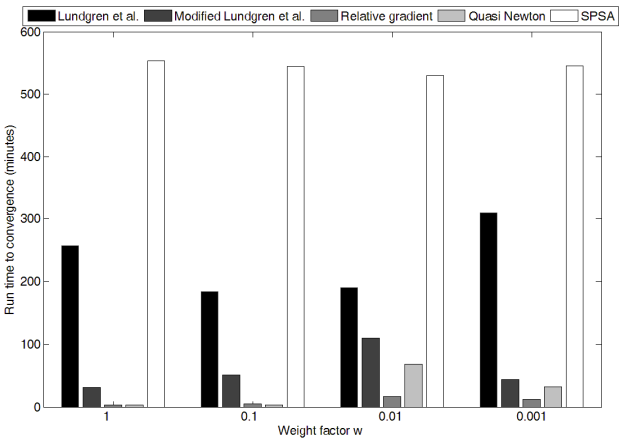

Fig. 9. The run time to convergence for the various algorithms and values of the weight factor w

VIII. Conclusions

This paper presented a general solution scheme and specific algorithms for the DODE problem. A critical construct in DODE is an assignment matrix, which maps OD demand flows to link sensor counts. In congested networks, the assignment proportions depend on the unknown time-dependent OD demands. The methods presented in this paper are based on use of linear approximations of the assignment matrix in the optimization iterations. Several specific solution algorithms that differ in the search direction they use were proposed. A case study demonstrated the applicability of the developed algorithms to large-scale complex networks and their computational efficiency compared to current state-of-the-art approaches.

The work presented in this paper may be extended and strengthened in several ways. First, additional case studies, with seed OD matrices that were generated in different ways (e.g. with random rather than uniform perturbation of the "true" demand), different network layouts, congestion levels and DTA or traffic simulation models are needed in order to strengthen the findings regarding the performance of these methods. Furthermore, it may be possible to improve the assignment matrix approximation in various ways, such as using higher order polynomials or by appropriate selection of setup parameters (e.g. the number and selection method of previous solutions) used in the approximation. Finally, in this work historic OD demands and link traffic counts were used exclusively in the estimation process. The possibility to accommodate other types of measurements (e.g. speeds, travel times, GPS and cellular tracking information) within the proposed approach may be investigated.

Acknowledgment

The authors thank Dr. Wilco Burghout from KTH in Stockholm for providing the Mezzo model and for his valuable assistance and the anonymous referees for many useful comments and suggestions.

References

[1] H. Yang, Y. Iida, and T. Sasaki, "The equilibrium-based origin-destination matrix estimation problem", *Transportation Research* part B, vol. 28, pp. 23-33, 1994.

[2] H. Yang, "Heuristic algorithms for the bi-level origin-destination matrix estimation algorithm", *Transportation Research* part B, vol. 29, pp. 231-242, 1995.

[3] M.L. Hazelton, "Estimation of origin-destination matrices from link flows on uncongested networks", *Transportation Research* part B, vol. 34, pp. 549-566, 2000.

[4] E. Cascetta and M.N. Postorino, "Fixed point approaches to the estimation of O-D matrices using traffic counts on congested networks", *Transportation Science*, vol. 35, pp. 134–147, 2001.

[5] M.J. Maher, and X. Zhang, "Algorithms for the solution of the congested trip matrix estimation problem", Proceedings of the 14th

International Symposium on Transportation and Traffic Theory, pp. 445-469, 1999.
[6] E. Cascetta, D. Inaudi, and G. Marquis, "Dynamic estimators of origin-destination matrices using traffic counts". *Transportation Science, vol.* 27, pp. 363-373, 1993.
[7] H. Tavana and H.S. Mahmassani, "Estimation of dynamic origin-destination flows from sensor data using bi-level optimization method", Preprints of the 80th Transportation Research Board Annual Meeting, National Research Council, Washington DC, 2001.
[8] K. Lindveld, "Dynamic O-D matrix estimation: A behavioural approach", Ph.D. dissertation, Faculty of Civil Engineering and Geosceinces, Technical University of Delft, The Netherlands, 2003.
[9] K. Ashok and M. Ben-Akiva, "Estimation and prediction of time-dependent origin-destination flows with a stochastic mapping to path flows and link flows", *Transportation Science, vol.* 36, pp. 184-198, 2002.
[10] E. Cipriani, M. Florian, M. Mahut and M. Nigro, "A gradient approximation approach for adjusting temporal origin–destination matrices", *Transportation Research part C*, vol. 19, pp. 270-282, 2011.
[11] W. Burghout, "Hybrid microscopic-mesoscopic traffic simulation", PhD Thesis, Royal Institute of Technology, Stockholm, Sweden, 2004.
[12] Cascetta E, "Transportation systems engineering theory and methods", Kluwer Academic Publishers, The Netherlands, 2001.
[13] H.D. Sherali and T. Park, "Estimation of dynamic origin-destination trip tables for a general network", *Transportation Research* part B, vol. 35, pp. 217-235, 2001.
[14] H. Tavana, "Internally-consistent estimation of dynamic network origin-destination flows from intelligent transportation systems data using bi-level optimization", PhD Thesis, Department of Civil Engineering, The University of Texas, Austin, 2001.
[15] N.J. Van der Zijpp and C.D.R. Lindveld, "Estimation of origin-destination demand for dynamic assignment with simultaneous route and departure time choice", *Transportation Research Record, vol.* 1771, pp. 75-82, 2001.
[16] X. Zhou, X. Qin, and H. S. Mahmassani, "Dynamic origin-destination demand estimation using multi-day link traffic counts for planning applications", *Transportation Research Record,* vol. 1831, pp. 30-38, 2003
[17] T. Toledo, M. Ben-Akiva, D. Darda, M. Jha, and H.N. Koutsopoulos, "Calibration of microscopic traffic simulation models with aggregate data", *Transportation Research Record*, vol. 1876, pp. 10-19, 2004.
[18] J.T. Lundgren, A. Peterson and C. Rydergren, "A heuristic for the estimation of time-dependent origin-destination matrices from traffic counts", in Advanced OR and AI Methods in Transportation, Proceedings of the 10th Jubilee Meeting of the EURO Working Group on Transportation, Jaszkiewicz, A., Kaczmarek, M., Żak, J., Kubiak, M. (eds.), Poznan University of Technology Publishers, pp. 242–246, 2005.
[19] M. Maher, X. Zhang and D. Van Vliet, "A bi-level programming approach for trip matrix estimation and traffic control problems with stochastic user equilibrium link flows". *Transportation Research* part B, vol. 35, pp. 23-40, 2001
[20] H. Kim, S. Baek, and Y. Lim, "O-D matrices estimation using genetic algorithms from traffic counts," *Transportation Research Recor.*, vol. 1771, pp. 156-163, 2001.
[21] B. Park and K. Zhu, "Time-dependent origin-destination estimation: Genetic algorithm-based optimization with updated assignment matrix," *KSCE Journal of Civil Engineering*, vol. 11, no. 4, pp. 199–207, 2007.
[22] J. Appiah and L. Rilett, "Joint estimation of dynamic origin-destination matrices and calibration of microsimulation models using aggregate intersection turn count data", *Preprints of the 89th Annual Meeting of the Transportation Research Board*, Washington DC, 2010.
[23] L. Kattan and B. Abdulhai, "Comparative analysis of evolutionary, local search, and hybrid approaches to O/D traffic estimation," *ASCE Journal of Transportation Engineering*, vol. 137, no. 1, pp. 46–56, 2011.
[24] L. Kattan and B. Abdulhai, "Sensitivity Analysis of an Evolutionary-Based Time-Dependent Origin/Destination Estimation Framework", *IEEE Transactions on Intelligent Transportation systems*, Vol. 13, no. 2, pp. 1442-1453, 2012
[25] A. Stathopoulos and Th. Tsekeris, "Hybrid meta-heuristic algorithm for the simultaneous optimization of the O-D trip matrix estimation", *Computer-Aided Civil and Infrastructure Engineering*, vol. 19, no. 6, pp. 421–435, 2004.
[26] J.C. Spall, "Implementation of the simultaneous perturbation algorithm for stochastic optimization", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, pp. 817–823, 1998.
[27] R. Balakrishna, "Off-Line calibration of dynamic traffic assignment models", PhD thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 2006.
[28] R. Balakrishna, M. Ben-Akiva and H.N. Koutsopoulos, "Offline calibration of dynamic traffic assignment: simultaneous demand-and-supply estimation", *Transportation Research Record*, vol. 2003, pp. 50-58, 2007.
[29] R. Balakrishna and H.N. Koutsopoulos, "Incorporating within-day transitions in simultaneous offline estimation of dynamic origin-destination flows without assignment matrices", *Transportation Research Record*, vol. 2085, pp. 31-38, 2008.
[30] V. Vaze, C. Antoniou, Y. Wen, and M. Ben-Akiva, "Calibration of dynamic traffic assignment models with point-to-point traffic surveillance", *Transportation Research Record*, vol. 2090, pp. 1–9, 2009.
[31] R. Balakrishna,M. Ben-Akiva, and H. Koutsopoulos, "Off-line calibration of dynamic traffic assignment: Simultaneous demand-supply estimation", *Transportation Research Record*, vol. 2041, pp. 80–88, 2008.
[32] H. Spiess, "A gradient approach for the O-D matrix adjustment problem", Publication No. 693, Transportation Research Centre, Montreal University, 1990.
[33] Y. Noriega, and M. Florian, "Some enhancements of the gradient method for O-D matrix adjustment", Research Report CIRRELT-2009-04, CIRRELT, Quebec, 2009.

Tomer Toledo received B.Sc. (1995) and M.Sc. (1998) in Civil Engineering from the Technion – Israel institute of Technology Haifa, Israel and a PhD (2002) in Transportation Systems from MIT, Cambridge MA USA.

He is currently an Associate Professor at the Faculty of Civil and Environmental Engineering, Technion – Israel Institute of Technology. His research interests are in large-scale traffic simulation models, intelligent transportation systems and driving and travel behavior.

Tanya Kolechkina received a B.E. (2005) in Multimodal Transportation Organization and Management from St. Petersburg state University of Civil Aviation, Russia and an M.Sc. (2010) in Transportation engineering from the Technion – Israel institute of Technology Haifa, Israel.

She is currently a PhD candidate at the Department of Civil and Environmental Engineering, Carlton University, Ottawa ON Canada. Her research interests are in intelligent transportation systems and connected vehicles.