Dynamic Origin-Destination-Matrix Estimation Using Smart card Data: An entropy maximization approach

Abderrahman Ait-Ali [1], Jonas Eliasson
Royal Institute of Technology, KTH

Abstract

Problems of dynamic origin-destination-matrix estimation using smart card data can be modelled as entropy maximization problems and efficiently solved using solution techniques such as Lagrangian relaxation. In this paper, we briefly review the literature of OD-matrix estimation and the use of smart card data. We show how this problem can be modelled and solved with incomplete smart card data where the trip distribution incoming to stations is known but not the outgoing distribution. A large non-linear entropy maximization problem is solved using an iterative algorithm solution based on Lagrangian relaxation. The algorithm is tested using a case study from the commuter train service in the great Stockholm region. Even though there are no theoretical guarantees for the algorithm's convergence, the results show that with the incomplete smart card data, the solution method converges and finds an estimate of the dynamic OD-matrix reflecting the reported aggregate statistics from the local operator.

Keywords: smart card; origin-destination estimation; Lagrangian relaxation; Entropy maximization.

1 Introduction

The problem of origin-destination (hereafter OD) matrix estimation appears in various domains and has been extensively studied by many researchers. In the four-step transport planning model, OD-matrix estimation forms the second component, i.e. an output of the trip generation model where the zones are determined and serves as an input to the mode choice and route assignment model [11].

The literature on this problem is rich, several researchers have studied the subject and different methods have been developed. There are generally two types of problems: static and dynamic [5]. The latter is called so because of the additional dimension of time. Thus, the static OD-matrix estimation aims at finding an estimate of the total number of trips between pairs of zones (e.g. stations) during a certain time interval (e.g. a working day). The dynamic OD-matrix estimation looks at the number of trips in every time period (e.g. every 15 minutes) over a certain time interval (e.g. a working day). Besides, The increasing availability of larger amounts of data (i.e. big data) led to many estimation models that can be used to improve the quality of the trip distributions in transportation models using data such as smart card data.

The aim of this paper is to show the use of the entropy-maximization based OD-matrix estimation approach with smart card data to determine the trip distribution in a large railway network. Starting from an entropy-maximization model and using Lagrangian relaxation to solve it, we derive an iterative solution algorithm. The latter

[1] Corresponding author: abde@kth.se

allows to find the dynamic OD-matrix estimate. The case study is the commuter train service in the greater Stockholm region and the available input data is that of the number of trips from stations at different time periods of the day [2].

The paper starts with a brief literature review of the existing OD-matrix estimation methods. The mathematical formulation of the model and solution algorithm follows before presenting the experiment and the results. The conclusion ends the paper.

2 Literature Review

There are several modeling approaches for OD-matrix estimation problems. Even though some approaches are equivalent, the choice of one technique rather than the other is often depending on the characteristics of the studied case and the available input data. This section summarizes some of the most popular state-of-the-art tools on the subject. It reviews the methodologies and techniques that have been commonly used in the literature.

There has been mostly three main methodologies for OD-matrix estimation: data survey-based methods, trip distribution models and OD-matrix updating methods [7]. In the first, the estimation is based on a data survey of the trip subjects (e.g. survey of car-owners, travellers). The second is based on estimation models that use incomplete or basic trips data. The last uses a flow information to update an old estimate of the OD-matrix. The methods and techniques that have been mostly used are a combination of one or more of these methodologies. What follows is a review of some of these methods.

2.1 Growth models

Growth factor models are OD-matrix estimation methods where an outdated matrix is updated based on the estimation and calibration of different growth rates. These methods are also simply called *growth models*. These models are very simple and are mostly based on the estimation of a *growth factor* (i.e. f) which is multiplied by the outdated matrix n^{old} as in equation 1 where n is the new OD-matrix estimate.

$$n = fn^{old} \tag{1}$$

An example of growth models is the *factors methods* where the growth rate is a constant factor that is computed uniformly from survey data (i.e. uniform factor method) or as an average factor (i.e. average factor method). These are some of the basic and first developed estimation techniques [9].
The *Fratar method* is also a popular growth-based technique. The growth parameters are computed to reflect the attractiveness of the inter-zonal movements. The method improves the accuracy of the factor technique by providing a more accurate estimate of the parameters [18]. Another variant of the growth models is the *Furness method*. It is a Fratar-like method but comes to reduce the amount of computation needed by using an iterative approach to compute balancing growth factors a_i and b_j as in equation 2 where n_{ij} are new estimates of the trips from i to j [12].

$$n_{ij} = a_i b_j n_{ij}^{outdated} \tag{2}$$

[2]This is because travellers in the commuter network in Stockholm use their ticket cards only when entering the stations and not when leaving them.

2.2 Gravity models

Gravity models attempt to estimate the OD-matrix using a gravitational attraction model, hence the name. Initially introduced as the Reilly's law of retail gravitation where the main idea was that the attractiveness of retails depends on the physical distance and the size of the retail centers [14]. This idea was used later to develop a mathematical model to predict traffic patterns based on land-use. In this study, the traffic patterns identify the number of trips from each origin to each destination which makes the model one of the earliest use of gravity models for OD-matrix estimation [17]. In addition to discrete choice models and urban economics, gravity models were among the most influential approaches to the analysis of land-use and transportation interaction [8].

One of the formulations of the gravity function is given in equation 3 where f is a deterrence function of the generalized travel cost c_{ij} and o_i is a parameter that relates to the trip origin and d_j to the destination. a_i and b_j are balancing factors as in the Furness method in equation 2.

$$n_{ij} = a_i b_j o_i d_j f(c_{ij}) \tag{3}$$

Entropy maximization approach is an equivalent approach to gravity models with a perspective that is based on the field of statistical theory of probability [19]. Entropy techniques were afterwards widely used to solve the OD-matrix estimation problem [3]. Another method that was based on the statistical theory of probability is the opportunity model. The development in its theory led to subsequent studies using one of the two variants: intervening and competing opportunity variant [15]. The main idea of these methods is that trip distributions are not explicitly related to distances but rather to the relative accessibility of opportunities to satisfy the purpose of the trip.

2.3 Other models

Many other studies used various additional techniques to solve the estimation problem. For instance, logit and discrete choice models were used to estimation trip distribution by combining the destination choice and mode choice models [2]. Other statistical techniques such as generalized least square [4], Bayesian inference [10] and principal component analysis [6] were also applied.

3 Methodology & Mathematical Model

This section describes the methodology that is adopted as well as the mathematical model for OD-matrix estimation in this paper. It also presents the solution algorithm to solve this model. First, an overview of the entropy maximization model is given. Second, we introduce some basic notations that are used to describe the mathematical model. Based on Lagrangian relaxation, we finally derive the solution algorithm to solve the estimation problem.

3.1 Entropy maximization and smart card data

Motivation

Entropy maximization models were shown to be equivalent to gravity models [19]. It has also been shown that information minimizing approach (i.e. entropy maximization) are strongly related to the utility maximizing models (e.g. discrete choice models) in spatial interaction [1]. The OD-matrix estimation model that is developed in this paper is based on an entropy maximization approach. Hence, gravity approach and other equivalent approaches. This is motivated by the nature of the studied problem and its assumptions and the available input data. Indeed, the problem can be accurately modelled as an optimization problem with functional constraints.

Public transport networks all over the world are increasingly using smart card automated fare collection systems. This led to the availability of data that can be useful for transportation planner. These data can been used for strategic, tactical, and operational planning [13]. The latter has not been widely studied by researcher and this paper shows an example of the use of smart card data in operational planning of commuter train services. The available smart data is that of the number of passengers entering each station during each 15 minutes of a working weekday. The dynamic OD-matrix, i.e. the temporal variation of passenger flow can be estimated by adding data related constraints to the optimization problem.

Methodology

Entropy maximization models can be formulated as convex nonlinear optimization problems where the objective function to maximize is the total entropy of the system under certain functional constraints. The solution to this optimization problem gives an estimate of the most probable OD-matrix in the system. The unknown variables in the optimization problem are the number of trips between the pairs of stations at each period of time (i.e. the three-dimensional OD-matrix). The objective function to be maximized is the entropy of the system as formulated in the original paper [19]. Furthermore, it is assumed that there are no trips from and to the same station and that the final estimate is non-negative.

Once the optimization problem is formulated, the constraints are relaxed and Lagrangian multipliers are associated to each of these. The optimality conditions give the relation between the unknown variables and the multipliers. Based on the inter-dependency between the multipliers, an iterative solution method is developed to estimate the Lagrangian multipliers associated with the relaxed constraints until the dual gap is small enough (i.e. up to a certain error tolerance). Hence, the value of the unknown variables and therefore the OD-matrix estimate. These are compared with the aggregated travel statistics from the local operator.

3.2 Basic notations

We consider the following useful notations

- T is the set of all time intervals (typically 15min intervals over a day).
- S is the set of all the network stations.
- n_{ij}^t is the number of trips from station i to j at time interval t.
- $n = (n_{ij}^t)$ is the dynamic OD-matrix to be estimated.
- $O_k = \sum_{t \in T} O_k^t$ is the total number of trips from station k (typically in one day).

One can already notice that the main unknown variables here are $n = (n_{ij}^t)$. The input data provides the values for O_k^t and hence the sets T and S. We note the sums with i, j, k instead of $i, j, k \in S$ and t instead of $t \in T$.

Furthermore, We assume that there is no trip from and to the same station, i.e. $n_{kk}^t = 0$. Therefore and for lighter formulations, we omit these trips in the summations, i.e we use $\sum_{ijt}$ instead of $\sum_{ijt, i \neq j}$.

3.3 Model

One can derive the expression of the total entropy of the system at hand as in the original paper that introduced the use of entropy [19]. The main idea is that the most probable trip distribution in the system is the one that maximizes the entropy of the system, i.e. chaos. This is inspired by a similar idea from the field of statistical mechanics. Using the previous notations, the number of possible states W in the system is given by equation 4 where $N = \sum_{it} n_{ij}^t = \sum_{jt} n_{ij}^t$ is the total number of trips in the system.

$$W = \sum_{ijt} \frac{N!}{\prod_{ijt} n_{ij}^t} \tag{4}$$

The total system entropy [3] is $E = ln(W)$ with the assumption that ln is the natural logarithm. Since N can be large for most real-world cases as with the case study in this paper (e.g. $\frac{1}{4}$ million trips), one can use the *Stirling* formula to simplify the log-factorial term which yields the simplified expression of the total entropy E in equation 5 where K is a constant term that does not depend on the variables n_{ij}^t.

$$E(n) = \sum_{i,j,t} n_{ij}^t - n_{ij}^t ln(n_{ij}^t) + K \tag{5}$$

The constant term K is omitted in the upcoming formulations. The maximization of the total entropy is equivalent to the minimization of $-E(n)$ which is adopted in the formulations that follow.

Using the previous notations, the mathematical formulation of the entropy maximization model is given in equation 6 with the functional constraints of the problem that is considered in this paper.

$$\begin{aligned} \underset{n}{\text{minimize}} \quad & \sum_{i,j,t} n_{ij}^t ln(n_{ij}^t) - n_{ij}^t \\ \text{subject to} \quad & \sum_{j} n_{kj}^t = O_k^t, \ \forall k \in S, \forall t \in T \\ & \sum_{i,t} n_{ik}^t = \sum_{j,t} n_{kj}^t, \ \forall k \in S \\ & n_{ij}^t \geq 0, \ \forall i, j \in S, \forall t \in T \end{aligned} \tag{6}$$

The model in equation 6 presents several constraints that each OD-matrix estimate n should satisfy. Some are generally found in most OD-matrix estimation models whereas others are specific to the problem in this paper and depend on the input data.

The first set of constraints enforces that the sum of trips from a station k during a time interval t should be equal to the input data count O_k^t of trips from the same station during the same time interval, i.e. $\sum_j n_{kj}^t = O_k^t$.

The second set formulates that the sum of trips from a certain station k (typically over one day) should be equal to [4] the sum of trips back to the same station k, i.e. $\sum_{j,t} n_{kj}^t = \sum_{i,t} n_{ik}^t$.

The last set of constraints makes sure that the estimated number of trips are non-negative for all the pairs of stations and for all time periods.

[3] Entropy is typically noted as S but we chose E instead to avoid confusion with the set of stations S

[4] This builds on the assumption that all people leaving from a station for work come back home to the same station

3.4 Lagrangian relaxation

The mathematical model is a very large nonlinear convex minimization problem (i.e. magnitude order of $|S|^2 * |T|$ variables and constraints); it is hard to solve for real world scenarios using existing solvers such as *CPLEX* or *Gurobi*. Hence, the model needs further simplification using methods such as Lagrangian relaxation which is adopted here.

One first step is to get rid of the third set of constraints by omitting the variables n_{kk}^t for all k, t. Thus, the formulation of the sum will only take into account non-identical pairs of stations, i.e. (i, j) such that $i \neq j$.

We now use Lagrangian relaxation to get rid of the first two sets of constraints. This leads to the Lagrangian function D, i.e. relaxed dual objective function, in equation 7.

$$\begin{aligned} D(n, \lambda, \mu) &= -E(n) - \sum_{k,t} \lambda_{kt} (\sum_{j \neq k} n_{kj}^t - O_k^t) - \sum_{k} \mu_k (\sum_{i \neq k, t} n_{ik}^t - \sum_{j \neq k, t} n_{kj}^t) \\ &= \sum_{i, j \neq i, t} (ln(n_{ij}^t) - 1 - \lambda_{it} + \mu_i - \mu_j) n_{ij}^t + \sum_{i,t} \lambda_{it} O_i^t \end{aligned} \tag{7}$$

Note that the primal objective function is $-E$. The multipliers λ are associated with the first set of constraints in equation 6 whereas μ are associated with the second set.

For a fixed value of λ and μ, the dual objective function can be minimized with respect to the variables n_{ij}^t which leads to the optimality conditions in equation 8 for $\forall i, j$ where $i \neq j$ and $\forall t$.

$$\frac{\partial D}{\partial n_{ij}^t} = ln(n_{ij}^t) - (\lambda_{it} - \mu_i) - \mu_j = 0 \implies n_{ij}^t = e^{\lambda_{it} - \mu_i + \mu_j} \tag{8}$$

It is now possible to write the dual objective function at the optimal value of $n_{ij}^t = e^{\lambda_{it} - \mu_i + \mu_j}$ as in equation 9.

$$D(\lambda, \mu) = \sum_{i, j \neq i, t} n_{ij}^t (ln(n_{ij}^t) - 1 - \lambda_{it} + \mu_i - \mu_j) + \sum_{k,t} \lambda_{kt} O_k^t = - \sum_{i, j \neq i, t} n_{ij}^t + \sum_{k,t} \lambda_{kt} O_k^t \tag{9}$$

The dual optimization problem is therefore the maximization of the dual with respect to the Lagrangian multipliers. The dual formulation (after simplification) is given in equation 10.

$$\underset{\lambda, \mu}{\text{maximize}} \quad D(\lambda, \mu) = \sum_{i,t} \lambda_{it} O_i^t - \sum_{i, j \neq i, t} e^{\lambda_{it} - \mu_i + \mu_j} \tag{10}$$

3.5 Solution algorithm

The optimality conditions on λ of the dual maximization problem are given by 11 whereas those on μ are given by 12.

$$\frac{\partial D}{\partial \lambda_{it}} = 0 \implies e^{\lambda_{it}} = O_i^t \frac{e^{\mu_i}}{\sum_{j \neq i} e^{\mu_j}} \tag{11}$$

$$\frac{\partial D}{\partial \mu_j} = 0 \implies e^{2\mu_j} = \frac{\sum_{i \neq j, t} e^{\lambda_{jt} + \mu_i}}{\sum_{i \neq j, t} e^{\lambda_{it} - \mu_i}} \tag{12}$$

The optimality conditions show that the dual variables are interdependent, i.e. one set of variables depend on another set. This leads to an iterative solution algorithm where one set of variables is computed based on the values of the other set. The algorithm is presented in Figure 1.

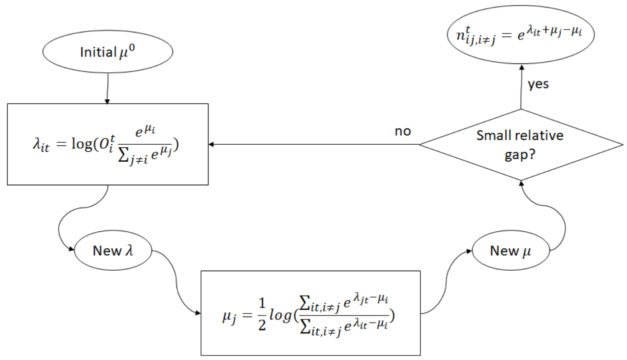

Figure 1. Overview of the iterative solution algorithm

The algorithms starts from an initial value of the multipliers μ (e.g. vector of zeros) which is used to compute the values of the multipliers λ using the optimality condition in equation 11. This leads to new updated values for λ which are in turn used to compute the new values for μ using equation 12. The algorithm iterates until the relative gap G_{rel} is small enough (e.g. smaller than a tolerance $\epsilon = 10^{-10}$) as in equation 13.

$$G_{rel} = \frac{|P - D|}{P} \leq \epsilon \tag{13}$$

Once the relative gap is small enough, the algorithm stops with the optimal values of the multipliers λ and μ maximizing the dual function D. These values allow to compute the OD-estimate using equation 8 which maximizes the total entropy of the system.

Since this original program is a non-linear convex optimization problem, one way to prove that an optimal solution exist is to use one of the constraint qualification conditions such as Slater condition for regularity conditions. It is possible to show that the problem has a feasible solution, it has therefore an optimal solution. It is however important to mention one major limitation of the solution algorithm which is that there is no theoretical guarantee that the algorithm in Figure 1 will always converge. However, if it converges then it does to an optimal solution as shown in the case study.

3.6 Model extension with additional data

It is possible to extend the model and improve its accuracy by using additional data. This data can come from different sources such as smart card or other digital systems used by the local operator.

The additional data can include information about the trips such as travel distances or travel times. In this paper, we assume that the additional data is as follows:

- $d_{ij} = d(i, j)$ is the trip distance between pairs of stations i, j.
- d^t is the average travelled distance during time period t per passenger.

It is possible to include these data by adding, to the optimization problem in equation 6, the constraint in equation 15, where $O^t = \sum_k O_k^t$ is the total number of passengers during time period t.

$$\forall t \quad \sum_{ijt} d_{ij} n_{ij}^t = d^t O^t \tag{14}$$

In the same way as before, the constraints are relaxed and we note $\theta = (\theta_t)$ as the corresponding Lagrangian multipliers. The Lagrangian function is formulated in equation 15.

$$D(\lambda, \mu, \theta) = \sum_{i,t} O_i^t \lambda_{it} + \sum_t d^t O^t \theta_t - \sum_{i,j \neq i,t} e^{\lambda_{it} - \mu_i + \mu_j + \theta d_{ij}} \tag{15}$$

Following the same steps, the optimality conditions on the Lagrangian function will lead to the new formulation of the number trip estimate in equation 16

$$\frac{\partial D}{\partial n_{ij}^t} = 0 \implies n_{ij}^t = e^{\lambda_{it} - \mu_i + \mu_j + \theta_t d_{ij}} \tag{16}$$

The iterative algorithm will be based on the same idea. The different is the additional cost parameter $\theta_t d_{ij}$ and that the Lagrangian multipliers θ_t are estimated by enforcing the satisfaction of the additional optimality constraints in equation 17. For fixed values of λ and μ, there is no general closed for for the estimate of θ_t in equation 17. One possible alternative is to estimate it numerically in the form of a unconstrained nonlinear minimization problem.

$$\frac{\partial D}{\partial \theta_t} = 0 \implies \sum_{i,j \neq i} d_{ij} e^{\lambda_{it} - \mu_i + \mu_j + \theta_t d_{ij}} = d^t O^t \tag{17}$$

Even though it is possible to successively estimate the Lagrangian multipliers up to a certain tolerance as previously in Figure 1, there is no guaranty that the algorithm will converge and this depends on the values of the additional data. However, if it converges then the OD-matrix estimate is optimal.

4 Experiment

This section describes how the model was implemented and tested using realistic input data. The case study is presented before briefly describing the implementation. Finally the results are presented and discussed.

4.1 Case study

In this experiment, we consider the dynamic OD-matrix estimation of the commuter train services in the greater Stockholm region. The network that is considered here (see. Figure 2) is from 2015, e.g. before the introduction of new stations in the end of 2017. The network has a total of 48 stations.

The main input data is the number of passengers arriving at each station during time periods of 15 minutes over an entire working day in 2015. It can be illustrated using a 3D plot with stations and time periods as in Figure 3.

Figure 2. Stockholm commuter network in 2015 (Patrik Ekengren, CC BY-SA 3.0)

4.2 Implementation

The iterative solution model was developed and tested using *MATLAB* R2017b. The input data was imported as a *CSV* file. The experiment was performed using a laptop with *8 GB* RAM and *CORE i5* processor.

The iterative algorithm has several parameters. The stopping condition was implemented using an error tolerance of $\epsilon = 10^{-10}$. The algorithm stops after a maximal number of iteration $K_{max} = 100$ if the stopping condition is not satisfied.

The input data includes some zero values. In order to avoid getting undefined number due to the logarithm function, these values were replaced with the epsilon value in *MATLAB* which has a value of 10^{-16}.

4.3 Results and discussions

The experiment allowed to compute the dynamic OD-matrix estimate that maximizes the total entropy of the system under the constraints in the model 6. The solution time is only few seconds for the whole network in the case study over one day. Besides, the experiment allowed to capture the evolution of the relative gap as well as the dual objective function.

The dual objective function for the different iterations is presented in Figure 4. The figure also shows the optimal primal value which corresponds to the maximal entropy of the system. The dual objective values are always below the optimal primal due to the weak duality. We notice that the dual function reaches the primal optimal value after 6 iterations. However, it does reach the requested error tolerance for the stopping condition after 45 iterations.

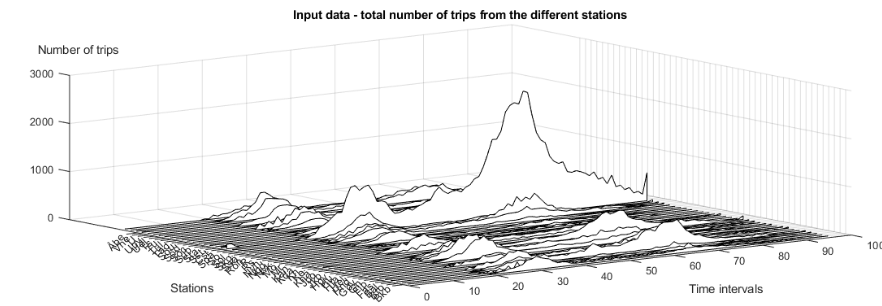

Figure 3. Input data - total number of trips from each station for every time period

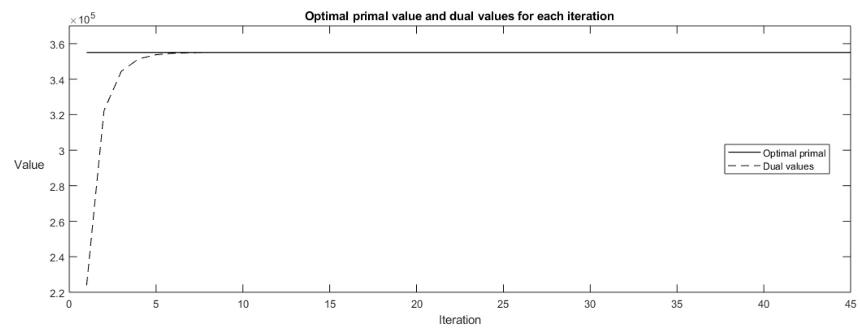

Figure 4. Dual objective values for each iteration (with optimal primal objective value)

The stopping condition depends on the relative gap which is defined in equation 13. The evolution of the relative gap is presented in Figure 5. This gap decreases and is practically zero after 10 iterations. However, it does reach the tolerated gap of $\epsilon = 10^{-10}$ after 45 iterations before stopping the algorithm with an estimate of the dynamic OD-matrix.

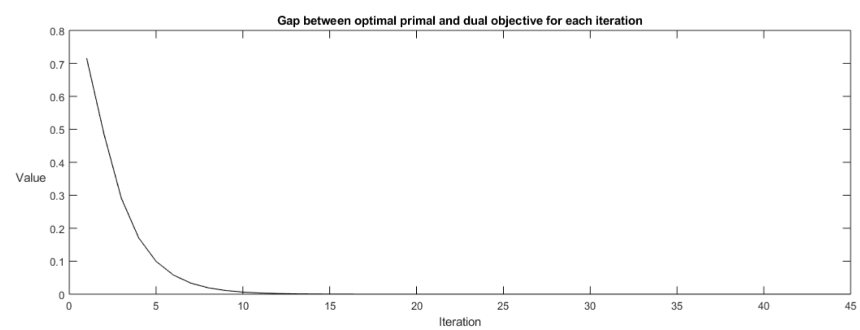

Figure 5. Relative gap for each iteration

The estimate of the dynamic OD-matrix that results from the solution algorithm is illustrated in Figure 6. This is estimate corresponds to the number of trips $n = (n_{ij}^t)$ that maximizes the entropy of the system, i.e. most probable situation.

There are several ways to validate the results of the model. One way to do that is the use of the estimated total average travel distance per person. *SLL*, responsible for the commuter train services in in Stockholm, provides each year these statistics [16]. For instance in 2015, the average travel distance per person was $18.7km$ per person. Using data on the distances between the pairs of stations, we find that the average total travel distance corresponding to the OD-matrix estimate is $16.9km$ per person which is

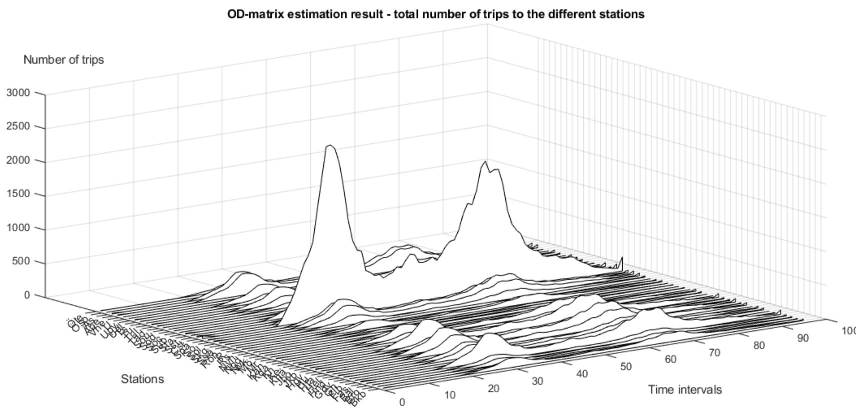

Figure 6. Result - dynamic OD-matrix estimate maximizing the entropy of the system

comparable with the reported statistics from *SLL*. The latter validation discussion shows that it is possible to have additional functional constraints in the estimation model in order to get statistical results that are similar to those in the report of *SLL*. Such constraints can relate to the average travel distance, average travel time or total kilometer person to name few of them.

5 Conclusions and future work

We have shown that there are several methods for OD-matrix estimation through the historical background in the literature review. We also developed an optimization model for the dynamic OD-matrix estimation problem before using the Lagrangian relaxation method to derive the iterative solution algorithm. The latter was tested on a case study from Stockholm's commuter train services. The iterative solution method reduced the dual gap in few seconds with the whole network and the resulting OD-matrix estimate gives comparable travel statistics to the ones reported by the local operator.

This developed model shows the use of smart card data for operational planning in commuter train services. It can be used to estimate microscopic trip distributions in large transport networks given partial information on the real trip distributions. The accuracy of the model can be simply improved by adding constraints based on the available data.

Since there is no guarantee that the iterative algorithm will always converge, an improvement can be to establish the theoretical conditions under which the algorithm converges. It can also be possible to compare the performances of this method in terms of accuracy and computational time with other solution methods.

References

1. A. Anas. Discrete choice theory, information theory and the multinomial logit and gravity models. *Transportation Research Part B: Methodological*, 17(1):13–23, 1983.

2. M. Ben-Akiva and S. R. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand*, volume 32 of *Transportation Studies*. Cambridge, MA: MIT Press, 9 edition, 1988.

3. M. Brenninger-Göthe, K. O. Jörnsten, and J. T. Lundgren. Estimation of origin-destination matrices from traffic counts using multiobjective programming formulations. *Transportation Research Part B*, 23(4):257–269, 1989.

4. E. Cascetta and S. Nguyen. A unified framework for estimating or updating origin/destination matrices from traffic counts. *Transportation Research Part B*, 22(6):437–455, 1988.

5. Q. Deng and L. Cheng. Research review of origin-destination trip demand estimation for subnetwork analysis. *Procedia - Social and Behavioral Sciences*, 96:1485–1493, 2013.

6. T. Djukic, G. Flötteröd, H. v. Lint, and S. Hoogendoorn. Efficient real time od matrix estimation based on principal component analysis. In *2012 15th International IEEE Conference on Intelligent Transportation Systems*, pages 115–121. IEEE.

7. J. Doblas and F. G. Benitez. An approach to estimating and updating origin–destination matrices based upon traffic counts preserving the prior structure of a survey matrix. *Transportation Research Part B: Methodological*, 39(7):565–591, 2005.

8. J. Eliasson and L.-G. Mattsson. A model for integrated analysis of household location and travel choices. *Transportation Research Part A: Policy and Practice*, 34(5):375–394, 2000.

9. A. W. Evans. Some properties of trip distribution methods. *Transportation Research*, 4(1):19–36, 1970.

10. M. J. Maher. Inferences on trip matrices from observations on link volumes: A bayesian statistical approach. *Transportation Research Part B*, 17(6):435–447, 1983.

11. M. G. McNally. *The Four-Step Model*, volume 1. Emerald Group Publishing Limited, 2007.

12. R. Morphet. A note on the calculation and calibration of doubly constrained trip distribution models. *Planning - Policy - Research - Practice*, 4(1):43–53, 1975.

13. M.-P. Pelletier, M. Trépanier, and C. Morency. Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, 19(4):557–568, 2011.

14. W. J. Reilly. *The law of retail gravitation*. W.J. Reilly, New York, 1931.

15. M. Schneider, J. T. Lynch, G. E. Brokke, and A. M. Voorhees. Panel discussion on inter-area travel formulas. *Highway Research Board Bulletin*, (253), 1960.

16. S. L. SL. Fakta om sl och länet 2015. Report, Storstockholms Lokaltrafik SL, 2016-10-25 2015.

17. A. M. Voorhees. A general theory of traffic movement. *Transportation*, 40(6):1105–1116, 2013.

18. F. Willekens. Log-linear modelling of spatial interaction. *Papers of the Regional Science Association*, 52(1):187–205, 1983.

19. A. G. Wilson. A statistical theory of spatial distribution models. *Transport Research*, 1(1967):253–269, 1967.