Gunnar Flötteröd, Kai Nagel

Modeling and Estimation of Combined Route and Activity Location Choice

Flötteröd, G.; Nagel, K. (2006). Modeling and Estimation of Combined Route and Activity Location Choice. 2006 IEEE Intelligent Transportation Systems Conference. https://doi.org/10.1109/ITSC.2006.1706773

Modeling and estimation of combined route and activity location choice

Gunnar Flötteröd* and Kai Nagel**

***Abstract*— This article describes a behavioral model of combined route and activity location choice. The model can be simulated by a combination of a time variant best path algorithm and dynamic programming, yielding a behavioral pattern that minimizes a traveler's perceived cost. Furthermore, the model is extended in a Bayesian manner, providing behavioral probabilities not only based on subjective costs, but also allowing for the incorporation of anonymous traffic measurements and the formulation of a traffic state estimation problem, which can efficiently be solved by an available algorithm.**

I. Introduction

The problem of traffic monitoring and prediction has been considered by many researchers. Various approaches are data-driven [1], [2], [3], while others adjust structural models to real world measurements. The latter group can further be classified with respect to what quantities are estimated: Some consider the problem of estimating physical traffic flow properties such as densities, velocities, or flow parameters [4], [5], while others (including this work) concentrate on the underlying demand itself and consider the physics of traffic flow as a strict constraint [6]. The second point of view goes structurally deeper, since traffic demand is the cause of road usage. Still, estimation of traffic demand and network link related quantities are two aspects of the same problem and ultimately should not be separated [7].

This article describes a novel methodology of traffic state estimation based on multi-agent simulations. We link flexible but little formalized representations of individual mobility behavior such as agent-based demand generation and microsimulation [8], [9] with mathematically well understood methodologies borrowed from control engineering. More precisely, we consider the problem of estimating agents' route and activity location choice in a Bayesian setting, combining for every agent an *a priori* activity plan for a given day with anonymous traffic measurements such as flows or densities into a most likely *a posteriori* plan. Since traffic demand results from individual mobility needs, no validated individualized knowledge should be "aggregated away" during the formalizing steps of setting up a mathematical estimation problem: Our approach uses fully individualized behavioral information in terms of a daily activity plan for every agent.

Due to space restrictions, not all aspects of the problem can be analyzed in depth: This article provides detailed description of the behavioral model, but only shortly visits the equally important aspects of micro/macro traffic flow simulation and numerical optimization. Further reading is provided in [10], [11].

The remainder of this article is organized as follows. Firstly, the deterministic modeling and simulation problem is discussed: In section II-A, combined route and activity location choice is modeled in terms of an optimization problem, for which a solution algorithm is given in section II-B. Extensions and limitations of this model are discussed in II-C and II-D. Secondly, the model is extended by aspects of randomness: In sections III-A and III-B a link between an agent's individual behavior and general observations of the traffic system in terms of anonymous measurements is provided. Section III-C then combines the behavioral model with that of anonymous measurements, yielding a formal description of the combined route and activity location choice estimation problem, for which a solution algorithm is suggested in III-D. Current and future evaluation of the algorithm within a real-world setting are discussed in section IV. The methodology is summarized in section V and an outlook on further work is given.

II. Deterministic modeling and simulation

A. A model of daily plans

Every agent μ has an individual plan for a given day, which is comprised as follows: The complete day is segmented into $n^\mu + 1$ temporal stages. Every such stage $0 \leq a \leq n^\mu$ is provided with a set $\mathcal{L}_a^\mu$ of one or more locations (network links) and a discrete start time step k_a^μ with $0 = k_0^\mu < k_1^\mu < \ldots < k_{n^\mu}^\mu$. Formally, stage a is nothing but a fixed temporal interval $[k_a^\mu, k_{a+1}^\mu)$ during which μ wants to be at one of the locations in $\mathcal{L}_a^\mu$. It can be interpreted as an *activity* such as "work", "leisure" or "shopping", while its location set can be understood as the *activity locations* where the individual expects facilities for execution of the according activity, e.g. different malls for a shopping activity. An example of such an activity plan is given in Figure 1. Note that the underlying network in which the example locations are situated is not not drawn, but only the logical multi-stage structure.

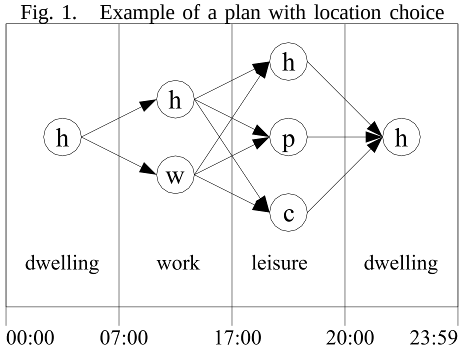
In this article, we do not consider departure time choice, although it is an important direction of research [12]. This decision is due to computational considerations given later on.

Every plan is anchored at its individual's unique home location $l_0^\mu = l_{home}^\mu$, where it starts and ends: $\mathcal{L}_0^\mu = \mathcal{L}_{n^\mu}^\mu = \{l_{home}^\mu\}$. Individual μ values the choice of location $l \in \mathcal{L}_a^\mu$ for activity a by $R_a^\mu(l)$; the cost of choosing this location is $C_a^\mu(l) = -R_a^\mu(l)$.

*Gunnar Flötteröd is with the Group for Transport Systems Planning and Transport Telematics, Technische Universität Berlin, D-10587 Berlin, Germany. Tel: +49-30-314-29520, fax: +49-30-314-26269, mail: fbetteroed@vsp.tu-berlin.de.

**Kai Nagel is head of the Group for Transport Systems Planning and Transport Telematics, mail: nagel@vsp.tu-berlin.de.

Fig. 1. Example of a plan with location choice



A four stage plan starting and ending at the agent's home location h. Stage 1 of the plan ("work" activity) can be conducted either at home (home working at location h) or at the office (work place location w). For stage 2 ("leisure" activity) there are are three possible locations, a cinema (c), a pub (p) and, again the home location. Note that the individual can choose to stay home the entire day.

A route starting at link i and time step k_0 to link j is denoted by $\boldsymbol{\mathcal{U}}(i,j,k_0)$. It will be convenient to represent it by

$$\boldsymbol{\mathcal{U}}(i,j,k_0) = \{\mathbf{u}(k)\}_{k \geq k_0} = \{(u_{rs}(k))\}_{k \geq k_0} \tag{1}$$

where $u_{rs}(k)$ is 1 if this route implies a turning move from link r to link s in time step k and zero otherwise. Here and in the following we only consider *feasible* routes in the sense that turning decisions are only made if the previous route led to a location where this turning move is physically possible. This property will in the following only be stated verbally ("$\boldsymbol{\mathcal{U}}$ is feasible"), since a formalization would just increase notational overhead.

For individual μ, the cost of traversing $\boldsymbol{\mathcal{U}}(i,j,k_0)$ is

$$C^\mu[\boldsymbol{\mathcal{U}}(i,j,k_0)] = \sum_{k \geq k_0} \sum_{rs} u_{rs}(k) c^\mu_{rs}(k), \tag{2}$$

which is additive in the nonnegative turning movement costs $c^\mu_{rs}(k)$ as perceived by μ. Link traversal costs can easily be incorporated by adding them to the turning move cost of entering the according link. The minimal cost path for μ between i and j when starting at k_0 is denoted by $\boldsymbol{\mathcal{U}}^\mu_{opt}(i,j,k_0)$ and its cost by $C^\mu_{opt}(i,j,k_0) = C^\mu[\boldsymbol{\mathcal{U}}^\mu_{opt}(i,j,k_0)]$.

During execution of their daily plans, individuals are aware of future effects their current activity location choice might have: Not the most attractive (least cost) activity location is chosen, but rather that location, which minimizes the expected cost for the entire remainder of the day. Since any individual's sequence of possible activity locations is known and finite, dynamic programming can be employed to solve this decision problem, as it will be shown in the next section.

B. Simulation of daily round trips

In order to describe the combined route and activity location choice problem as a multi-stage decision process, a residual cost $V^\mu_a(j)$ is introduced. It is defined as the minimal cost to be experienced when starting activity a at location $j \in \mathcal{L}^\mu_a$ and continuing in an optimal manner:

$$V^\mu_a(j) = -R^\mu_a(j) + \min_{l \in \mathcal{L}^\mu_{a+1}} \{C^\mu_{opt}(j,l,k^\mu_{a+1}) + V^\mu_{a+1}(l)\} \tag{3}$$

for $a < n^\mu$, while $R^\mu_0(l^\mu_{home})$ and $V^\mu_{n^\mu}(l^\mu_{home})$ can be arbitrarily set to 0 since they have no influence on the final result.

For μ being located on *any* link i at time step k and heading for activity a, the task of optimally completing its round trip can now be stated as the problem of finding a next activity location $l^\mu_a \in \mathcal{L}^\mu_a$ with minimal cost $C^\mu_{opt}(i, l^\mu_a, k) + V^\mu_a(l^\mu_a)$, being given by

$$l^\mu_a = \arg \min_{j \in \mathcal{L}^\mu_a} \left\{C^\mu_{opt}(i,j;k) + V^\mu_a(j)\right\}. \tag{4}$$

This can be achieved by calculation of a *single* best path from i to an imaginary destination d which directly succeeds all locations $j \in \mathcal{L}^\mu_a$ by means of likewise imaginary connecting links of cost $V^\mu_a(j)$. This yields the best next activity location (which is the last real link on the obtained path) as well as the best path itself.[1]

In the same manner, an optimal round trip can be obtained by one sweep through all activity stages: $l^\mu_{n^\mu} = l^\mu_{home}$ is fix. Running backwards through stages $a = n^\mu - 1, \ldots, 0$ allows to calculate for every activity location j of current stage a the optimal next activity location (4) and its residual cost (3). Having reached $a = 0$, the optimal round trip can then be obtained by moving forwards through all stages and choosing the optimal next location as annotated during the previous backwards sweep. This procedure is standard dynamic programming.

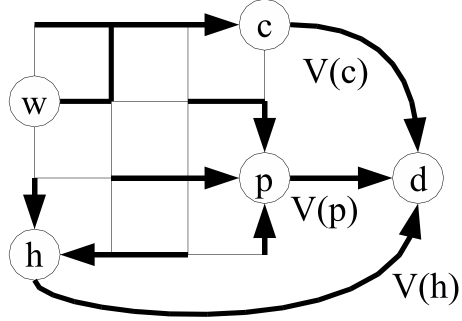
The calculation of residual costs for all activity locations requires n^μ best path tree calculations, each one connecting all locations of a given stage to the single extra node behind all locations of the next stage as it is shown in figure 2.

C. Within-day replanning

This calculation scheme can efficiently be applied for simulation of within-day replanning: Consider an individual μ, which so far followed a pre-calculated route towards its next activity location l^μ_a. Assume that μ now faces a significant deviation between the observed traffic situation and its historically learned one (on which its precomputed route is based). It appears reasonable that μ spontaneously replans its current decision stage, while keeping its evaluation of subsequent activity locations fixed. This is equivalent to direct application of (4) in order to obtain a new route (and maybe a new activity location) reflecting the current situation. The only required computation for such a single-stage decision is the calculation of *one* best path through one of the next temporal stages' locations towards the imaginary destination node behind it, as previously explained.

[1]Note that the optimal path does not change if a positive cost is equally added to all imaginary links. Raising these links' costs to a nonnegative level allows us to meet all requirements for application of a dynamic version of Dijkstra's best path algorithm [13].

Fig. 2. Calculation of a single decision stage



This figure shows a best path tree representing the optimal transition from figure 1's "work" activity stage to its "leisure" stage. The tree's root is an imaginary destination node d, which directly follows all possible activity locations h, c, and p of the "leisure" stage. Bold lines on the underlying grid network represent best paths towards d. The figure allows to identify the optimal (route and) next activity location choice for every node of the network: If the agent is currently at the office w, going to the cinema is the most attractive next step, while a home-worker at h would effectively stay home (note the shortcut from h to d). The pub p is only attractive if the agent already is in its very proximity.

Since we have shown that activity location choice can be subsumed in a slightly modified route choice problem, the following discussion will only treat the according best path problem without explicitly mentioning location choice.

D. Discussion of model limitations

Economic theory suggests that the marginal value of conducting an activity decreases over time and ultimately approaches zero. The model described above assumes duration independent activity values implying zero marginal costs, which is realistic only for long activity durations. Currently, we impose a lower bound on stage lengths when generating activity plans.

As long as departure times are fixed at stage transitions, duration dependent activity values can be incorporated by making the costs of the aforementioned imaginary links behind activity locations time variant. Realistic modeling of departure time choice would require additional state information representing the duration an agent has already been conducting an activity [12]. Since we already have to search an entire time variant traffic network in order to model spontaneous route adjustment, we will avoid this state space increase and keep departure time fixed until we have computationally investigated our approach on larger scenarios.

III. Incorporation of uncertainty and behavioral estimation

So far, a model and a simulation algorithm for combined route and activity location choice have been provided. The remainder of this article describes how these results can be adopted to calculate not only an agent's minimum cost behavior, but rather its most likely *a posteriori* behavior, given its *a priori* activity plan and additional traffic measurements.

Ultimately, this is achieved by provision of modified turning costs to the agent that represent the linearized *a posteriori* log-likelihood of *not* making the according turning moves. An agent minimizing this cost approximately maximizes the *a posteriori* likelihood of its decisions.

A. Macroscopic traffic dynamics

In the absence of measurements, the most likely *a posteriori* trip should be the one of minimal cost. If there are measurements, they have to be related to individual turning decisions in order to allow for the aforementioned turning cost modifications. This is accomplished via a differentiable macroscopic traffic flow model. Since the results given in this section follow from extensive preparatory work, the presentation is quite tight. Sources for further reading are cited.

We use a macroscopic first order traffic flow model that runs in discrete time and space [11], [14], [15]. The model is represented in state space form. Macroscopic occupancies on every link segment constitute the state vector $\mathbf{x}(k)$. The model takes accumulated turning counts $\mathbf{u}(k) = \sum_\mu \mathbf{u}^\mu(k)$ of individual, microscopically represented vehicles as control variables and internally turns them into splitting fractions of macroscopic flows at all intersections. Vice versa, microscopic vehicles are moved through the network according to velocities provided by the macroscopic model. First order traffic flow dynamics define the macroscopic system's evolution through time, formally given by the state equation $\mathbf{x}(k+1) = \mathbf{f}[\mathbf{x}(k), \mathbf{u}(k), k]$.

Measurable quantities $\mathbf{y}(k)$ such as traffic flows or velocities are calculated from the macroscopic model's states via an output equation $\mathbf{y}(k) = \mathbf{g}[\mathbf{x}(k), k]$. Since we consider a first order model, traffic occupancy on a link segment fully defines the average speed (and thus flow) on this segment.

Altogether, a differentiable relationship between any individual μ's route choice and the output of the macroscopic traffic flow model can be stated by the following set of equations [10]:

$$\begin{aligned} \mathbf{u}(k) &= \sum_\mu \mathbf{u}^\mu(k) \\ \mathbf{x}(k+1) &= \mathbf{f}[\mathbf{x}(k), \mathbf{u}(k), k] \\ \mathbf{y}(k) &= \mathbf{g}[\mathbf{x}(k), k]. \end{aligned} \tag{5}$$

B. Modeling of anonymous traffic measurements

We now assume that the control sequence $\boldsymbol{\mathcal{U}}^\mu = \{\mathbf{u}^\mu(k)\}_k$ generated by any individual μ's route choice incorporates random effects as specified in the next section. Although the state and output equation of the macroscopic traffic flow model remain deterministic, the random control component also turns the systems state and output vectors into random variables. This allows to express all causal relations between model in- and outputs by conditional probabilities.

Some components of the traffic system's output vector $\mathbf{y}(k)$ can be observed via sensors such as inductive loops, floating cars, or traffic surveillance cameras [16]. Since these sensors are prone to various sources of error, it is assumed that a concrete measurement $\tilde{\mathbf{y}}(k)$ follows a conditional probability density function $h(\tilde{\mathbf{y}}(k) \mid \mathbf{y}(k))$, which is differentiable and models the actual measurement $\tilde{\mathbf{y}}(k)$ as a random variation of the macroscopic model's output $\mathbf{y}(k)$.[2]

In the next section, we will use the notion of a measurement $\tilde{\mathbf{y}}(k)$'s conditional probability $\mathcal{P}(\tilde{\mathbf{y}}(k) \mid \mathbf{y}(k))$, which we understand as the probability that $\tilde{\mathbf{y}}(k)$ lies within a certain region Z around $\mathbf{y}(k)$ (i.e., $\tilde{\mathbf{y}}(k), \mathbf{y}(k) \in Z$) being sufficiently small to allow for the following first order approximation:

$$\begin{aligned} \mathcal{P}(\tilde{\mathbf{y}}(k) \mid \mathbf{y}(k)) &= \int_Z h(\mathbf{z} \mid \mathbf{y}(k)) \mathrm{d}\mathbf{z} \\ &\approx h(\tilde{\mathbf{y}}(k) \mid \mathbf{y}(k)) \cdot \int_Z \mathrm{d}\mathbf{z} \end{aligned} \tag{6}$$

Now, the probability $\mathcal{P}(\boldsymbol{\mathcal{Y}} \mid \boldsymbol{\mathcal{U}}^\mu)$ of a measurement sequence $\boldsymbol{\mathcal{Y}} = \{\tilde{\mathbf{y}}(k)\}_k$ can be related to the chosen route $\boldsymbol{\mathcal{U}}^\mu = \{\mathbf{u}^\mu(k)\}_k$ of any individual μ by

$$\mathcal{P}(\boldsymbol{\mathcal{Y}} \mid \boldsymbol{\mathcal{U}}^\mu) = \prod_k \mathcal{P}(\tilde{\mathbf{y}}(k) \mid \mathbf{y}(k)) \tag{7}$$

together with equations (5) and (6). Note that the resulting relationship between $\boldsymbol{\mathcal{Y}}$ and $\boldsymbol{\mathcal{U}}^\mu$ is approximately differentiable and thus can be linearized.

C. Formulation of the estimation problem

It is assumed that individual μ is currently moving through the network towards one of the activity locations of its current plan stage. Without consideration of measurements, the individual's *a priori* path and location choice can be simulated as explained in section II-B.

This choice mechanism is now probabilistically relaxed. The *a priori* probability that the individual actually chooses a path $\boldsymbol{\mathcal{U}}^\mu$ is expressed in terms of a multinomial logit model [17]

$$\mathcal{P}(\boldsymbol{\mathcal{U}}^\mu) = \frac{e^{-\beta C(\boldsymbol{\mathcal{U}}^\mu)}}{\sum_{\boldsymbol{\mathcal{V}}} e^{-\beta C(\boldsymbol{\mathcal{V}})}} \tag{8}$$

where the normalizing denominator sums over all paths $\boldsymbol{\mathcal{V}}$ the individual can choose from. Note that this choice set will never be explicitly generated. We are aware of this simple model's drawbacks, still we consider it to be a good starting point because of its tractable analytical form [17], [18].

In the absence of further information (such as measurements) the minimum cost path would have maximal probability of being chosen. Thus, a probability maximizing estimator of the individuals *a priori* route choice would yield the same result as the cost minimization procedure given in section II-B.

Now it is assumed that some measurements $\boldsymbol{\mathcal{Y}}$ are available. The *a posteriori* probability $\mathcal{P}(\boldsymbol{\mathcal{U}}^\mu \mid \boldsymbol{\mathcal{Y}})$ that an individual chose path $\boldsymbol{\mathcal{U}}^\mu$ in consideration of $\boldsymbol{\mathcal{Y}}$ is expressed via Bayes' theorem:

$$\mathcal{P}(\boldsymbol{\mathcal{U}}^\mu \mid \boldsymbol{\mathcal{Y}}) = \frac{\mathcal{P}(\boldsymbol{\mathcal{Y}} \mid \boldsymbol{\mathcal{U}}^\mu)\mathcal{P}(\boldsymbol{\mathcal{U}}^\mu)}{\mathcal{P}(\boldsymbol{\mathcal{Y}})}. \tag{9}$$

After taking the logarithm of this function, we substitute (7) and (8):

$$\begin{aligned} \ln \mathcal{P}(\boldsymbol{\mathcal{U}}^\mu \mid \boldsymbol{\mathcal{Y}}) &= \sum_k \ln \mathcal{P}(\tilde{\mathbf{y}}(k) \mid \mathbf{y}(k)) - \beta C(\boldsymbol{\mathcal{U}}^\mu) \\ &+ \underbrace{\ln \sum_{\boldsymbol{\mathcal{V}}} e^{-\beta C(\boldsymbol{\mathcal{V}})} - \ln \mathcal{P}(\boldsymbol{\mathcal{Y}})}_{\text{independent of } \boldsymbol{\mathcal{U}}^\mu} \end{aligned}$$

Substituting (2) as well as (6) and dropping all terms independent of $\boldsymbol{\mathcal{U}}^\mu = \{\mathbf{u}^\mu(k)\}_k$, the most likely *a posteriori* route $\boldsymbol{\mathcal{U}}^\mu$ of any individual μ can be stated as the solution of the following control problem:

$$\begin{aligned} &\text{minimize} \\ &J^\mu = \sum_k \left(-\ln h(\tilde{\mathbf{y}}(k) \mid \mathbf{y}(k)) + \beta \sum_{ij} c^\mu_{ij}(k) u^\mu_{ij}(k) \right) \\ &\text{s.t.} \quad \begin{array}{rcl} \mathbf{y}(k) &=& \mathbf{g}[\mathbf{x}(k), k] \\ \mathbf{x}(k+1) &=& \mathbf{f}[\mathbf{x}(k), \mathbf{u}(k), k] \\ \mathbf{u}(k) &=& \sum_\mu \mathbf{u}^\mu(k) \\ \boldsymbol{\mathcal{U}}^\mu &\text{is}& \text{feasible} \end{array} \end{aligned} \tag{10}$$

Thus, the problem of estimating a populations most likely behavior in terms of route and activity location choice is equivalent to solving problem (10) simultaneously for every agent μ in this population.

D. Solution of the estimation problem

An algorithm for approximate solution of (10) is outlined in this section, while further details are provided in [10]. Since J^μ is linear in good approximation with respect to a single agent's turning decisions, it is replaced by

$$\bar{J}^\mu = \sum_k \sum_{ij} \left(\lambda_{ij}(k) + \beta c^\mu_{ij}(k)\right) u^\mu_{ij}(k) \tag{11}$$

with real-valued coefficients $\lambda_{ij}(k)$ being defined through linearization of (10). Minimization of $\bar{J}^\mu$ by optimal choice of a feasible path $\boldsymbol{\mathcal{U}}^\mu$ can then approximately be achieved by a time variant best path algorithm that operates on modified costs $d^\mu_{ij}(k) = \max\{0, \lambda_{ij}(k) + \beta c^\mu_{ij}(k)\}$.

The lower bound of $d^\mu_{ij}(k)$ avoids cycles of negative cost so that Dijkstra's best path algorithm can be employed. Since all $c^\mu_{ij}(k)$ are nonnegative, the error introduced by this bound depends on β, which represents the *a priori* information's reliability. The usage of scaled costs $d^\mu_{ij}(k)/\beta = \max\{0, \lambda_{ij}(k)/\beta + c^\mu_{ij}(k)\}$ yields the same $\boldsymbol{\mathcal{U}}^\mu$ and is more amenable practical implementation as well as non-mathematical interpretation, since it only implies the addition of a correction term $\lambda_{ij}(k)/\beta$.

Minimizing (10) by synchronous modification of *many* agents' trajectories is more difficult. Clearly, the increased number of degrees of freedom has the potential for a better

[2]For notational simplicity it is assumed that (a) $\tilde{\mathbf{y}}$ and $\mathbf{y}$ are of same dimensions and (b) function h is time invariant.

overall solution, still this setup results in certain problems also encountered in dynamic route guidance: If many drivers are independently of each other informed of a low travel time route, they might all switch towards this route, causing a jam and very high travel times times [19]. Similarly, the individual linearization (11) of the overall problem does not allow for a coordination of different agents' path optimizations.

The proposed algorithm resembles the fixed point solution approaches to self consistent route guidance in the sense that it iteratively updates only a subset of all agents. One iteration of the algorithm is given below:

1) Load all agents of population $\mathcal{M}$ onto the network;
2) linearize target functional J^μ and obtain $\bar{J}^\mu$;[3]
3) choose a subset $\mathcal{M}' \subset \mathcal{M}$;
4) calculate a new path $\boldsymbol{\mathcal{U}}^\mu$ for every $\mu \in \mathcal{M}'$ that approximately minimizes $\bar{J}^\mu$ by dynamic best path algorithm;
5) continue with 1) if desired.

This algorithm becomes identical to a popular traffic assignment heuristic that solves the equilibrium problem in terms of a fixed point iteration if no measurements are available [20]. Since traffic assignment based on this method has become common practice, the method can be expected to also work well for our purposes.

IV. An ongoing case study

A. Setting of the test case

We have set up an extensive test case for the proposed algorithm. The geographical zone of investigation is the city of Berlin. Its traffic network is represented by a graph of currently 6400 links. The multi-agent simulation system MATSIM [9] has been used to generate activity plans for a complete microscopic representation of the Berlin population. For real time operations, a 10% sample of this demand (approx. 170.000 agents) is used.

Real world measurements are currently obtained from two sources of information: Sparse inner-city velocity information is obtained from floating car data being provided by a taxi fleet of a few hundred vehicles. Inductive loop data is available on the inner-city highway. This data is transmitted every few minutes to a standalone desktop PC that takes care of the entire estimation procedure.

As explained before, incorporation of measurements into the online traffic simulation only requires modifications of the travel costs provided to replanning agents. One iteration of the overall system at real world time t roughly involves the following steps:

1) Adaptive simulation from $t - 30$min to t: First a fraction of all agents recalculates a new route based on modified costs calculated in the previous iteration, then the demand is loaded onto the network;
2) predictive simulation from t to $t + 30$min as a continuation of 1), this step is irrelevant to the estimation itself;
3) calculation of new cost modifications for the next iteration based on measurements obtained during the last 30min.

The computational effort of a single iteration is between 300 and 600 seconds on a Pentium 4 desktop PC with 2GB RAM. The entire software system has been implemented in the Java programming language.

B. Very preliminary discussion

Against the background of the current soccer world championship in Berlin it becomes clear that the number of effects presently influencing the traffic situation is much larger than in normal operations. Although attempts have been made to incorporate event-specific behavior in the populations activity plans, a thorough validation of the method is hard with the large number of uncertainties currently influencing the results.

Still, some statements can already be made.

- A good *a priori* demand in terms of realistic activity plans for the population is important. The combined route and and activity choice problem for an entire population is utterly underdetermined if only online measurements and no a priori plans are used. If the quality of a priori information is low, little weight is put on the agents' *a priori* cost minimization (expressed by a small β parameter in eq. (8)). This results in the simultaneous attempt of all replanning agents only to reproduce the current measurements, which makes the solution prone to oscillations.
- Despite the aforementioned problems, the method works in terms of measurement error reduction. It shall again be noted that this adjustment ability only results from behavioral adaptations. Since traffic flow dynamics are treated as strict side constraints, no adjustment of link related quantities without according underlying behavioral patterns takes place.
- The current situation in Berlin made clear that an additional incident detection module will be inevitable. Otherwise, the estimator tries to adjust reasonable behavioral patterns to events of totally different causality such as a crowd celebrating their team in the middle of a major road.
- The method works in real-time. A population of 170.000 agents is adjusted on a 6400 link network in real time. The employed rolling horizon procedure updates the estimation every 300 to 600 seconds, depending on the network load. During this time interval, one iteration of a 30min estimation problem is solved and a 30min prediction is calculated. All of this takes place on a single desktop PC.

V. Summary and outlook

We presented a novel methodology of behavioral state estimation for traffic systems modeled by multi-agent simulations. Several steps were undertaken to obtain the results presented in this article: 1. Design of a differentiable, yet fast macroscopic mobility simulator for networks of arbitrary

[3] Since this functional can be considered to be the same for broad classes of travelers (e.g. "informed", "uninformed"), an efficient numerical treatment is possible.

topology; 2. movement of individual particles through this mobility simulator without loss of its differentiability; 3. representation of the overall system in state space form; 4. construction of an algorithm that solves a general nonlinear control problem for this dynamic system in terms of agents' trajectories through the network; 5. representation of travelers' route and activity location choice as an optimization problem; 6. formulation of the behavioral estimation problem in a Bayesian setting and its formulation as a nonlinear control problem, which can be solved by the proposed algorithm.

A large-scale, real-world test case has been set up. Although the current situation in Berlin is far from optimal for careful testing or fine-tuning, all infrastructure for ongoing investigations has been implemented and will be employed to collect more practical experience with the system.

Continuous operations will have another useful side-effect: Currently, all agents start their day with an offline generated activity plan. During the day this plan is adjusted to measurements. If the plan was not discarded at the end of the day (as it is currently the case), but be used as the next day's *a priori* plan, a continuous improvement of the demand could be achieved.

REFERENCES

[1] G. Huisken, E. van Berkum, A Comparative Analysis of Short-Range Travel Time Prediction Methods, *TRB 2003 (82. Annual Meeting)*, Washington DC, USA.

[2] Y. Kamarianakis, P. Prastacos, Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches, *TRB 2003 (82. Annual Meeting)*, Washington DC, USA.

[3] C. Zhou, P. Nelson, Predicting traffic congestion using recurrent neural networks, *Proceedings of the 9th World Congress on Intelligent Transport Systems*, Chicago, 2002.

[4] S. Lorkowski, P. Wagner, Parameter calibration of traffic models in microscopic online simulations, *TRB 2005 (84. Annual Meeting)*, Washington DC, USA.

[5] Y. Wang, M. Papageorgiou, Real-time freeway traffic state estimation based on extended Kalman filter: a general approach, *Trans. Res.*, vol. 39B, 2005, pp 141-167.

[6] K. Ashok, Estimation and Prediction of Time-Dependent Origin-Destination Flows, *PhD Thesis*, Massachusetts Institute of Technology, 1996.

[7] C. Antoniou, On-line Calibration for Dynamic Traffic Assignment, *PhD Thesis*, Massachusetts Institute of Technology, 2004.

[8] M. Balmer, K. Axhausen, K. Nagel, A demand generation framework for large scale micro simulations, *TRB 2006 (85. Annual Meeting)*, Washington DC, USA.

[9] M. Balmer, K. Nagel, B. Raney; Large scale multi-agent simulations for transportation applications; *J. of Intelligent Transport Systems* 8:205, 2004.

[10] G. Flötteröd, K. Nagel, Simulation and optimization of trajectories in a congested network, *VSP working paper 06-02*, http://www.vsp.tu-berlin.de/publications/, accessed 2006.

[11] G. Flötteröd, K. Nagel, Some practical extensions to the Cell Transmission Model, *ITSC 2005 8th International IEEE Conference on Intelligent Transportation Systems*, Wien, Austria.

[12] D. Charypar, P. Graf, K. Nagel, Q-learning for flexible learning of daily activity plans, *TRB 2005 (84. Annual Meeting)*, Washington DC, USA.

[13] E. Dijkstra, A note on two problems in connexion with graphs, *Numerische Mathematik*, 1959, pp 269-271.

[14] C.F. Daganzo, The cell-transmission model: A simple dynamic representation of highway traffic, *Trans. Res.*, vol. 28B(4), 1994, pp 269-287.

[15] C.F. Daganzo, The cell-transmission model, part II: Network traffic, *Trans. Res.*, vol. 29B(2), 1995, pp 79-93.

[16] J. Kamata, T. Oda, Detectors for road traffic, *Concise encyclopedia of traffic and transportation systems*, Pergamon Press, 1991.

[17] M. Ben-Akiva, S. R. Lerman, *Discrete choice analysis*, The MIT Press, Cambridge, MA, 1985.

[18] E. Cascetta, A. Nuzzolo, F. Russo, A. Vitetta, A modified logit route choice model overcoming path overlapping problems: Specification and some calibration results for interurban networks, *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, Lyon, France, 1996, pp. 697-712.

[19] J. Bottom, Consistent Anticipatory Route Guidance, *PhD Thesis*, Massachusetts Institute of Technology, 2000.

[20] E. Cascetta, *Transportation systems engineering: theory and methods*, Kluwer Academic Publishers, Dordrecht, 2001.