Data Fusion Technique in the Context of Traffic State Estimation

Irina Matschke* Bernhard Friedrich* Karsten Heinig*

*Institute of Transport, Road Engineering and Planning
University of Hannover
Appelstr. 9A, 30167 Hannover, Germany
matschke@ivh.uni-hannover.de; friedrich@ivh.uni-hannover.de;
heinig@ivh.uni-hannover.de

1 Introduction

Real time traffic management in urban road networks requires information on the present and future traffic state. This information needs to be as complete and precise as possible. Today, traffic information can be collected from various sources such as inductive loops, video observation and floating car data (FCD). To achieve a precise intelligent transportation management it is essential to obtain real-time traffic information. However, it is not realistic to get the comprehensive information of all turning movements due to the high cost of the required measuring equipments.

Therefore, the focus is to improve traffic demand estimation by using integrated data fusion techniques-based algorithms with few detector data, including FCD, as input. With these techniques only some additional information on turning movements, flow, delay and routes is needed and then the data coverage and data quality in network can be more completed and improved respectively. The benefit of such a system will thereafter be the improved accuracy of the estimation with the use of the existing detection infrastructures.

2 State of the Art

Traffic state estimation needs the information from different sources to provide consistent flows and travel time. These sources can be divided into the measurements (data collected by detectors or coming from the signal timing plan, etc.) and the additional information, gained by using the measurements in estimation algorithms or similar methods. Those additional information can be the estimated flows within intersections [1,7,9], the estimated movements at intersections with traffic lights [8] propagated link flow counts [12] or estimated queue lengths and waiting times at intersections with traffic lights [5]. Data deriving from floating cars (floating car data, FCD) have recently been described as valuable supplementary information [6]. At present those data sources exist parallel and are barely used corporately.

3 Methodology of the Approach

The concept of a new method for online traffic state estimation presented in this paper is to split the system into a network level and an intersection level. On the intersection level data fusion techniques are applied in real time to combine the detected flow data and the information on the signal timing. The respective algorithms will generate the information on turning movements, queue lengths, delay and flow. On the network level the enhanced data then is the basis for the determination of consistent flows and travel times. Figure 1 shows an overview about the concept of the methodology, the required data elements and the program flow chart. The process includes mainly six parts:

- Based on the fusion of traffic counts and traffic light timings, i.e. data which is available to a certain degree for all networks, the volumes for all movements within an intersection can be calculated.
- The calculated volumes are then propagated to the circumjacent sectors with the consideration of the accuracy. This procedure offers the opportunity to obtain more exact data at each detected intersection and also to fill data gaps on links where no detector is available.
- Queue lengths will also be estimated with the data fusion technique by combining traffic counts and traffic light timing. Floating car data (FCD) may be used for the comparison and the calibration of the estimation.
- With the given information on link flows (and in particular on turning movements) and the guess of the route choice the first OD matrix can be estimated.
- A traffic assignment then uses the processed data on volumes, queue lengths and OD relations and results in consistent flows and travel times. Based on this data a new iteration of OD estimation and assignment is performed.
- Floating car data will also be used to compare and calibrate travel times as well as an additional information (weight) for the OD estimation.

3.1 Intersection Movements

To get more and primarily precisely data a model is applied which only uses common upstream detector loops as data source. The method combines signal phase timing information and detected flow data by assigning the probability function for each particular movement over time. Disaggregated traffic flow information collected by the upstream detectors is related to the green times of the signal timing. Using the temporal dependence of detection time and signal phasing of the different streams estimates of the proportions could be made. It is because not all vehicles from the possible streams which pass the same detector are allowed

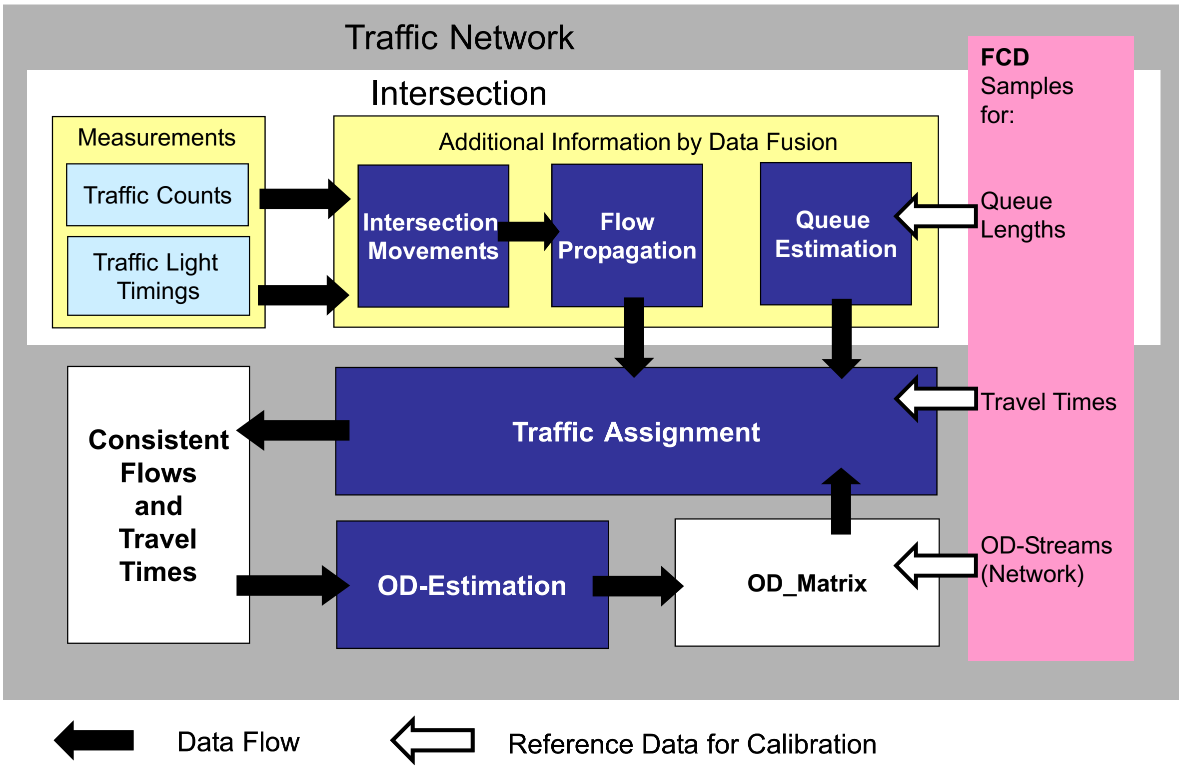


Figure 1: Data fusion technique in the context of traffic state estimation

to move at the same time. This approach was already implemented and tested by the authors [8] and resulted in a high performance of the estimation of turning flows at an intersection (determination coefficient of estimated compared to real turning flows > 0.9).

3.2 Flow Propagation

The detected or calculated volumes are now propagated to the next links upstream and downstream. The approach is here to allocate as much as possible links with flow information, because the more information the estimation algorithms can use, the better the OD estimation results can be.

In Figure 2 the concept of link flow propagation is illustrated. The information of q_{12} is calculated by propagating the detected links flows of detectors $D24$, $D26$ and $D28$ with the use of the information of intersection movements described above. Therefore the data gap can be filled out by calculating

$$q_{12} = D24 + D26 + D28$$

An upper limit for q_{12} is provided by the addition of $D13$ and $D11$:

$$q_{12} \leq D11 + D13$$

The smaller of both values is then selected.

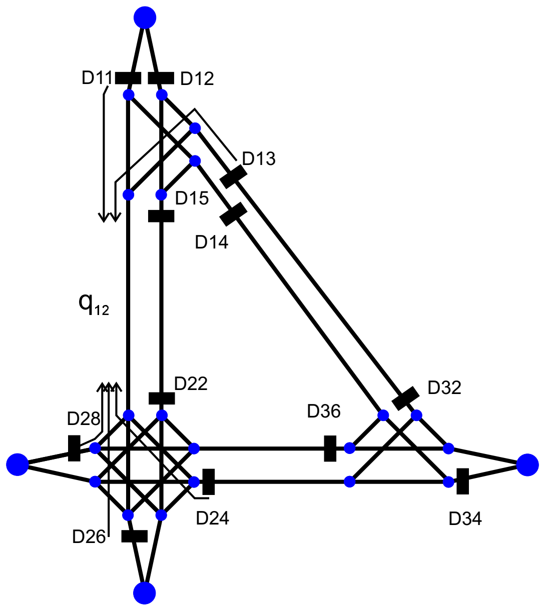


Figure 2: Propagation of link flows

3.3 Queue estimation

To involve also other additional information apart from link flows and turning movements, an approach is also implemented which determines queue lengths on the basis of vehicle counts from detectors located close to the stop line and the information of signal timings. This model was firstly introduced by MÜCK [10] and the authors [4] have proven, that a high performance gain could be achieved by employing this estimation module as a quasi measurement with Kalman filtering technique in queuing theory models (e. g. Markovian chains, Kimber-Hollis). Using this feed back procedure the problem of systematic instability of queue length and delay estimation could be improved especially in cases of the high and varying degree of saturation.

3.4 Introducing supplementary information to OD estimation

In recent years a large number of models have been developed to estimate an origin-destination matrix from link traffic count data (see [2,3,11] among the most cited papers). Typically, the entropy maximising, information minimizing, and least squares estimators have been proposed and applied. The concept of these models is to update or improve an old given OD matrix so that the estimated link volumes are consistent with the measured ones. VAN ZUYLEN, for example, has used the principle of minimum information to define the most likely OD matrix as

$$T_{ij} = t_{ij} \cdot X_0 \cdot \prod_a X_a^{\left(\frac{p_{ij}^a}{g_{ij}}\right)}$$

$$\left(g_{ij} = \sum_a p_{ij}^a\right)$$

where t_{ij} is a priori guess of the OD matrix, $X_0 = \frac{T}{\sum_i \sum_j t_{ij}}$ is a factor factor to include information on the total number of trips and $X_a^{n+1} = X_a^n \cdot \frac{q_a^{real}}{q_a^{est}}$ is a factor for the iterative solution of the equation by using observed and estimated traffic volumes q_a^{real} and q_a^{est} respectively for the adjustment.

Apart from the historic information included in t_{ij}, the assignment process which converts the knowledge of the trip matrix into link flows has a high influence on the quality of OD estimation. Besides, the accuracy of the assignment information p_{ij}^a, the fraction of trips from origin i to destination j that passes over link a, is also important in the matrix estimation process.

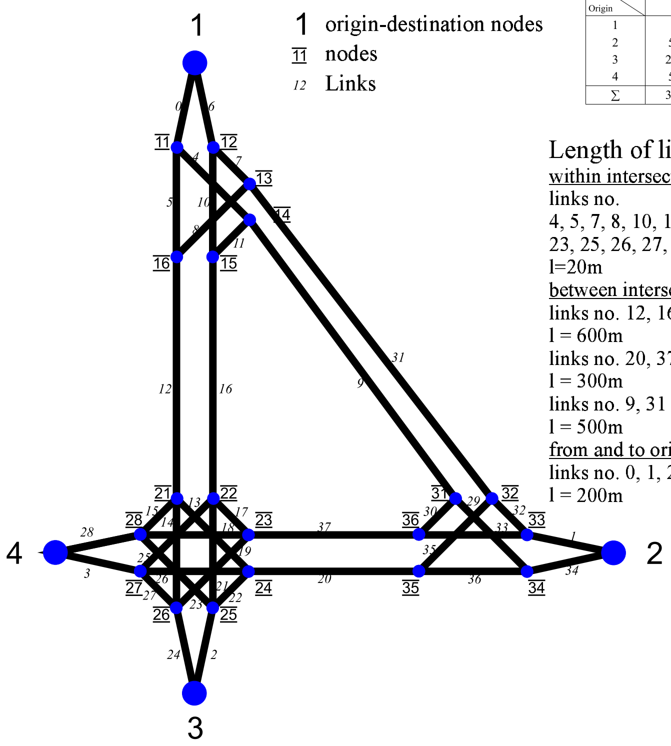
The additional information, which is generated by the afore mentioned approaches, is made available to the OD estimation algorithms described above. It is expected that the calculated OD matrix obtains therefore a higher quality. To achieve an even higher improvement the additional information is not only used in the OD estimation process but also in the assignment. This is done by using the respective data as reference values in the cost function for calculating the link travelling cost. Furthermore, the estimated delays and queue lengths on the intersection level could be used to revise the link costs which are needed in the assignment process. The OD estimation can be thereafter improved with the better assignment.

4 Computational Results

To examine the effects of the new approach on OD estimation, a simulation network, as a test bed, was constructed and displayed in Figure 3. The reference OD matrix in Figure 3, a specific OD route fraction and OD flow was assumed. Based on these assumptions the "true" link traffic counts could be determined.

The network-wide OD estimation was accomplished by the afore mentioned model of VAN ZUYLEN. In our test scenarios the priori guess of the OD matrix was the unit matrix to determine the estimation quality and only the influence of additional information was considered. Any other influences were neglected. The information on turning flows and their propagation to the circumjacent links was used as quasi traffic counts and thus the set of input data could be extended. Besides the set of counted links also the quality precision of the assignment information p_{ij}^a has an impact on the OD estimation. The impact was analyzed by comparing the qualities of the OD matrix using four different assignments, an exact assignment, a nearly exact assignment, an assignment with some different OD flow proportioning and an all-or-nothing assignment. The exact assignment was used as the reference.

Furthermore, seven different scenarios have been also developed to indicate the different influences of incomplete sets of link counts. Scenario 1 with all 38 links (14 links between intersections and 24 turning links) was used as the reference. In scenario 2 and 3 the OD matrix is estimated and based only on the links between the intersections respectively on the turning links. Scenario 4 and 5 add successively turning links to the set of links to identify which benefit could be gained by adding this additional information. In scenario 6 and 7 link information between intersections is removed and replaced by turning movements. An



Destination / Origin	1	2	3	4	Σ
1	-	100	200	150	450
2	50	-	100	150	300
3	200	150	-	150	500
4	50	100	50	-	200
Σ	300	350	350	450	1450

Length of links
within intersections
links no.
4, 5, 7, 8, 10, 11, 13, 14, 15, 17, 18, 19, 21, 22, 23, 25, 26, 27, 29, 30, 32, 33, 35, 36
l=20m
between intersections
links no. 12, 16
l = 600m
links no. 20, 37
l = 300m
links no. 9, 31
l = 500m
from and to origin-destination nodes
links no. 0, 1, 2, 3, 6, 24, 28, 34
l = 200m

Figure 3: The layout of the test network.

overview of these seven different scenarios with the corresponding sets of links is given in table 1.

The scenarios are evaluated using the average estimation error (Root Mean Square Error - RMSE). Therefore, each estimated OD matrix is compared with the reference OD matrix and the average RMSE of the OD pairs is then calculated (the diagonal elements are not considered). A comparison of the results of the different scenarios is shown in Table 3.

The results show that the inclusion of the additional turning flows improves the OD estimation. Especially the comparison between scenario 2 and 3 shows that the accuracy of OD estimation can be highly achieved when turning flows are known. Comparing scenario 6 with 7 the benefit of propagation of data could be also indicated. These two scenarios differ in the inclusion of the links 37 and 19 in the set of links. In case of estimating the OD matrix by including the turning link 19 an improvement can be made and completing the link set (plus link 37, scenario 4) via propagation the estimation error can be further reduced.

But what would happen if the determination of the turning flows is not exact and the additional information is thereafter also incorrect? This case is analyzed in scenario 5 by falsifying the turning flows of link 17, 21 and 25 and the results are shown in Table 2. These first results show that the OD estimation could be improved through additional information on the turning flows even with incorrect flows (table 4.

Table 1: Scenarios and the corresponding sets of links

Scenario	Set of links
1	all links
2	0, 6, 1, 34, 2, 24, 3, 28, 12, 16, 9, 31, 20, 37
3	4, 5, 7, 8, 10, 11, 13, 14, 15, 17, 18, 19, 21, 22, 23, 25, 26, 27, 29, 30, 32, 33, 35, 36
4	0, 6, 1, 34, 2, 24, 3, 28, 12, 16, 9, 31, 20, 37, 19
5	0, 6, 1, 34, 2, 24, 3, 28, 12, 16, 9, 31, 20, 37, 17, 21, 25
6	0, 6, 1, 34, 2, 24, 3, 28, 12, 16, 9, 31, 20
7	0, 6, 1, 34, 2, 24, 3, 28, 12, 16, 9, 31, 20, 19

Table 2: Incorrect additional information

link	correct turning flows	version 1	version 2	version 3
	veh	veh	veh	veh
17	20	15	100	15
21	200	180	180	150
25	60	50	50	50

Further on, the analysis of the impact with the consideration of the delay and the queue length in the assignment is undertaken and will be finished within few months. Based on the first results it is expected that the OD estimation can be improved even more.

References

[1] M. Cremer and H. Keller, "Dynamic Identifications of FLows from Traffic Counts of Complex Intersections", *Prof. of the 8th International Samposium on Transportation and Traffic Theory.* 121-142, University of Toronto Press, (1981).

[2] M.G.H. Bell, "The Estimation of an origin-destination-matrix form traffic counts", *Trans-*

Table 3: RMSE of the test scenarios with different assignments

scenario	exact	nearly exact	some discrepancies	all-or-nothing
	veh	veh	veh	veh
1	0.19	4.82	11.58	40.07
2	19.77	20.02	29.51	52.53
3	0.30	8.92	15.88	34.10
4	14.77	16.13	23.80	50.33
5	11.97	13.01	23.97	45.40
6	27.70	23.98	33.94	52.62
7	16.31	19.86	24.98	49.06

Table 4: RMSE of scenario 5

without additional information	correct turning flows	version 1	version 2	version 3
veh	veh	veh	veh	veh
29.51	23.97	26.80	28.30	29.47

portation Science 17 (2), 198-217 (1983).

[3] E. Cascetta, "Estimation of trip matrices from traffic counts and survey data: a generalised least squares approach estimator", *Transportation Research B* 18 (4/5), 289-299 (1984).

[4] B. Friedrich, I. Matschke, E. Almasry, J. Mück, "Data Fusion Techniques for Adaptive Traffic Signal Control". *Proceedings of the 10th IFAC Symposium, August 4-6* (2003).

[5] R.M. Kimber and E.M. Hollis, "Traffic queues and delays at road junctions", *TRRL Laboratory Report* 909 (1981).

[6] R. Kühne, R.-P. Schäfer, J. Mikat, K.-U. Thiessenhusen, U. Böttger and S. Lorkowski, "New Approaches for Traffic Management in Metropolitan Areas", *Proceedings of the 10th IFAC Symposium, August 4-6* (2003).

[7] M. Maher, "Estimating the turning flows at a junction: a comparison of three models", *Traffic Engineering and Control* 25, 19-22, (1984).

[8] I. Matschke and B. Friedrich, "Dynamic OD Estimation Using Additional Information from Traffic Signal Lights Timing", *Proceedings of the Triennial Symposium on Transportation Analysis TRISTAN IV, June 13-19* (2001).

[9] P. Mirchandani, S.A. Nobe and W.W. Wenji, "Online Turning Proportion Estimation in Real-Time Traffic-Adaptive Signal Control" *Transportation Research Record* 1784, 80-86, (2002).

[10] J. Mück, "Estimation methods for the state of traffic signals using detectors near the stop-line", *Traffic engineering and control*, (2002).

[11] H. van Zyulen and L.G. Willumsen, "The most likely trip matrix estimated from traffic counts", *Transportation Research B* 14, 291-293, (1980).

[12] P. Vortisch and S. Hengst, "Umlegungsbasierte Messwertpropagierung zur Verkehrszustandsberechnung", *Proceedings HEUREKA 2002* 209-215, (2002).