EFFICIENT QUEUE LENGTH DETECTION AT TRAFFIC SIGNALS USING PROBE VEHICLE DATA AND DATA FUSION

Thorsten Neumann[1*]
1. German Aerospace Center (DLR), Institute of Transportation Systems
Rutherfordstr. 2, 12489 Berlin, Germany
+49 (0)30 67055-208, thorsten.neumann@dlr.de

ABSTRACT

In this paper, a new method for the detection of queue lengths at traffic signals is described. Based on conventional probe vehicle data and implementing an extremely flexible data fusion approach for the integration of nearly arbitrary additional traffic information, it provides an efficient way to get high-quality estimates for the traffic states at traffic signals.

A systematic evaluation based on extensive simulations addresses several issues concerning quality and demonstrates both the great potentials as well as the limitations of the new approach. In conclusion, it becomes apparent that high accuracies can be achieved even with today data.

KEYWORDS Traffic monitoring, probe vehicle data, data fusion

INTRODUCTION

Nowadays, probe vehicle systems based on GPS are an important candidate for an intelligent area-wide and cost-effective traffic monitoring. In this context, the positions of several vehicles from the overall traffic flow (transmitted every few seconds, i.e. in the range from 5 seconds up to 1 or 2 minutes) are used to derive actual travel times for the roads of the observed area [1]. Probe vehicle systems of this type are running at some German cities like Berlin, Hamburg, Stuttgart or Nuremberg, for example. Finally, intelligent navigation and route guidance systems can be implemented based on this actual traffic information (e.g. [2]).

However, the solitary use of probe vehicle data typically does not meet the requirements of an efficient traffic management including traffic signal control, for example. On the one hand, this is because of the difficult interpretation of travel times for traffic signal optimization purposes. On the other hand, quality of traffic information obtained from probe vehicle data solely often is not good enough because of very low penetration rates (< 1% probe vehicles).

A new approach overcomes both problems by yielding estimates of the queue length at traffic signals which can directly be used for traffic signal optimization, and by providing very flexible data fusion options to significantly enhance the quality of these estimates. Especially, this flexibility turns out to be a big advantage compared to other recent approaches for queue length estimation [3, 4]. Besides the needs for the probably simplest form of currently available probe vehicle data, the new method makes very low demands on the process of data collection. It just takes additional traffic data as they are available by integrating them as more or less strong evidences for particular traffic states. Moreover, it uses information from probe vehicle data sets which is typically neglected by conventional probe vehicle systems, i.e. it uses the correlations between all data points instead of trajectory information of single cars. Hence, the new approach tries to extract as much information as possible from currently (or in the future) available traffic data. In doing so, it minimizes the costs for construction and operation of infrastructure needed for urban traffic monitoring.

THE NEW APPROACH

With regard to inner-city traffic, it is an important fact that vehicles usually cluster at traffic signals. Assuming that the probe vehicles are distributed homogeneously amongst all cars, the same clustering effect is observable in real probe vehicle data, i.e. there are more probe vehicles sending GPS coordinates located in front of traffic signals than from other regions of the road network. Thereby, it turns out that the relative frequency of a specific position is directly correlated with the (time-dependent) local traffic density at that location (see [5] for a more detailed description of the new approach).

Hence, if the possible density profiles with the corresponding queue lengths (related via a virtual traffic demand q in the model below) are known, the queue lengths can be estimated easily by selecting that density profile which makes the observed set $x_1,\ldots,x_n$ of probe vehicle positions to be as likely as possible. For the concrete implementation of this approach, assuming stationary traffic conditions, the possible density profiles $K(q)$ can be derived analytically (cf. [6]) based on a common traffic flow model [7] which can also be used to describe urban traffic [8]. Moreover, depending on the virtual traffic demand q, the density profiles $K(q)$ are uniquely associated with the corresponding queue lengths $L(q)$. Then, based on the normalized versions K_q of $K(q)$, the probability for registering a given single probe vehicle position x is defined by $K_q(x)$, and the queue length L^* for the observed data set $x_1,\ldots,x_n$ can be estimated via

$$L^* := L(q^*) \quad \text{where} \quad q^* := \underset{q\in[0,1]}{\arg\max}\, f(q \mid x_1,\ldots,x_n) := \underset{q\in[0,1]}{\arg\max}\left\{ w_q \cdot \prod_{i=1}^{n} K_q(x_i) \right\}. \qquad (1)$$

Mathematically spoken, this is a kind of generalized maximum-likelihood estimation where w_q are suitable a-priori weights. At this point, by modifying these weights, the search space for the correct virtual traffic demand q^*, i.e. the correct queue length $L(q^*)$ can be adjusted in a very flexible way. For example, if the correct queue length $L(q^*)$ or the correct virtual traffic demand q^* were already known from some additional traffic information, it would be even possible to reduce the set of potential estimates so that it would comprise the value $L(q^*)$ only. Just let $w_q = 0$ for all $q \neq q^*$. Hence, these a-priori weights obviously allow for the integration of nearly arbitrary additional traffic information, i.e. they provide an extremely flexible way for real data fusion with other than probe vehicle data.

SYSTEMATIC EVALUATION BASED ON SIMULATION

To evaluate the new approach, extensive simulations with about 5 million scenarios (combinations of different penetration rates, transmission frequencies, etc.) covering the whole range of possible traffic demand values q were performed. To this end, the general setup was given by an isolated road section comprising a distance of about 750m (≈ 100 veh. lengths[1]). The inflow is driven stochastically following a Poisson distribution approximately, and the outflow is controlled by traffic signals with fixed cycle times during the whole simulation campaign. Then, virtual probe vehicle data were extracted from the simulations to be processed using the above algorithm. Finally, the obtained estimates were compared to the 'real' queue lengths observed during the simulation. The result is a systematic evaluation of the new approach. The comprehensive findings are described on the following pages.

Influence of different lead times

To get reliable and comparable results, each simulation was initialized with free flow traffic for the complete road section under consideration, i.e. without any queue. Hence, some time T_0 is needed until the initial build-up of the queue at the traffic signal is finished. At that, T_0 depends on the average queue length (associated with traffic demand q) and on q itself. In this context, Figure 1 shows two typical plots of the average estimation error for different values T_0 with and without data fusion[2].

a) without data fusion b) with data fusion

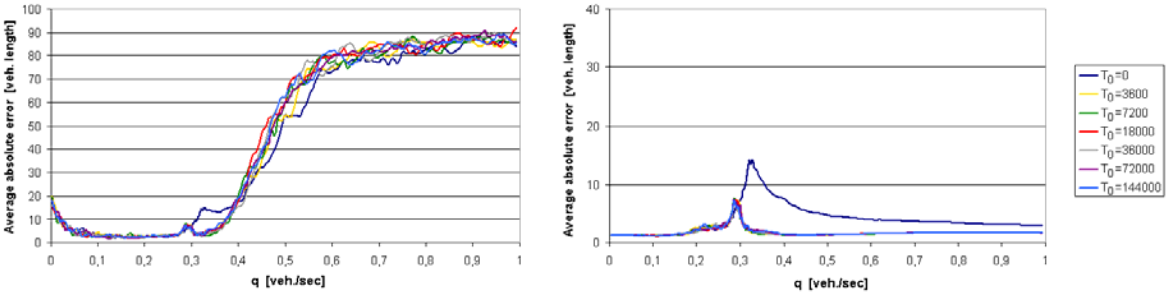

Figure 1 – Average estimation error (depending on lead time T_0)

As can be seen, lead times $T_0 \geq 3600$ sec are sufficient for stable estimation results in general. Accordingly, all further analyses are based on lead times between 3600 and 7200 seconds. But even without lead time ($T_0 = 0$ sec), the same stable results can be observed at least for appreciably under-saturated traffic conditions ($q < 0.28$ veh./sec). For, in this case, the average queue is comparatively short and thus is build up rapidly. The most crucial differences between $T_0 = 0$ sec and lead times $T_0 \geq 3600$ sec can be found near to the critical traffic demand $q_{crit} \approx 0.289$ veh./sec at slightly over-saturated traffic because of long but slow developing queues at the traffic signal. With increasing q however, the deviations become smaller again due to a faster build-up of the queue (see Fig. 1b).

Influence of penetration rate and transmission frequency

In general, it can be stated that the quality of estimated queue lengths becomes better with increasing penetration rates ρ and increasing transmission frequencies $1/\Delta t$ as expected. At

[1] The effective average vehicle length in congestion is 7.5m approximately (cf. [7]).
[2] The fatal estimation errors in the case without data fusion at over-saturated traffic conditions with large q (see Fig. 1a) will be clarified below.

penetration rates above 5%, average estimation errors less than 2 vehicle lengths are theoretically possible even without any data fusion (see Fig. 2a). By aggregation over longer time intervals T and by using the options of data fusion (assuming that the real traffic demand is at least roughly known[3]), such a high quality can be achieved even by penetration rates of about 1% and transmission frequencies with about Δt = 60 sec between two subsequent position messages of the same probe vehicle (see Fig. 2b). Obviously, this is very interesting for reality since many current probe vehicle systems are working with similar parameters.

a) without data fusion ($T = \Delta T$) b) with extended aggregation and data fusion ($T = 15 \cdot \Delta T$)

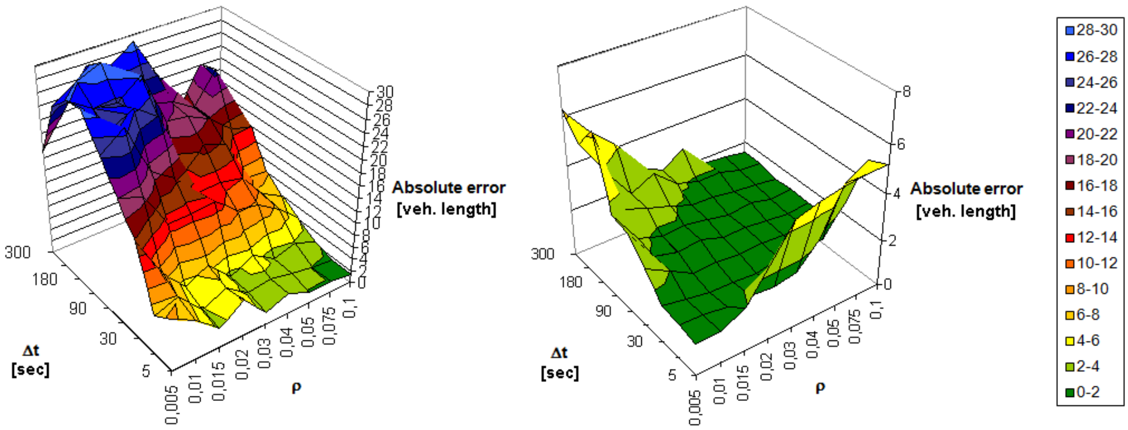

Figure 2 – Average estimation error (depending on penetration rate ρ and transmission frequency $1/\Delta t$) where $0 \leq q \leq q_{crit}$ and ΔT = 3600 sec

Nevertheless, Figure 2b also shows some problems at large penetration rates and high transmission frequencies in case of too much aggregation. For seemingly mysterious reasons, the estimation error begins to increase at some point although the amount of data increases, too, and thus better results should be expected due to simple statistics. A reasonable explanation of this effect will be given in context of the detailed analysis of the influence of aggregation in the corresponding section below.

At this point, however, Figure 3 demonstrates that a significant improvement of the results from Figure 2a can be achieved even with aggregation or data fusion alone. In case of a still larger degree of aggregation (see Fig. 3a), penetration rates of about 1% and transmission frequencies with only 1 position message every 60 seconds from each probe vehicle are sufficient for an average estimation error less than 2 vehicle lengths again. With data fusion alone, such high-quality estimates are obtained with less than twice as many probe vehicles (i.e. ρ = 1.5%) even at comparatively short aggregation intervals $T = \Delta T$ (see Fig. 3b).

Finally, Figures 2 and 3 show that the same high quality is achieved for relatively large areas of the parameter space (ρ, Δt). Hence, the described estimation approach is very robust within wide limits concerning penetration rate and transmission frequency. Especially with regard to local fluctuations of the penetration rate or varying transmission frequencies in real probe vehicle systems, this is an important property of the new method.

[3] More precisely, it is assumed that the real traffic demand is known up to ± 360 vehicles per hour approximately (e.g., because of loop detector data or others).

a) with extended aggregation only ($T = 30 \cdot \Delta T$) b) with data fusion only ($T = \Delta T$)

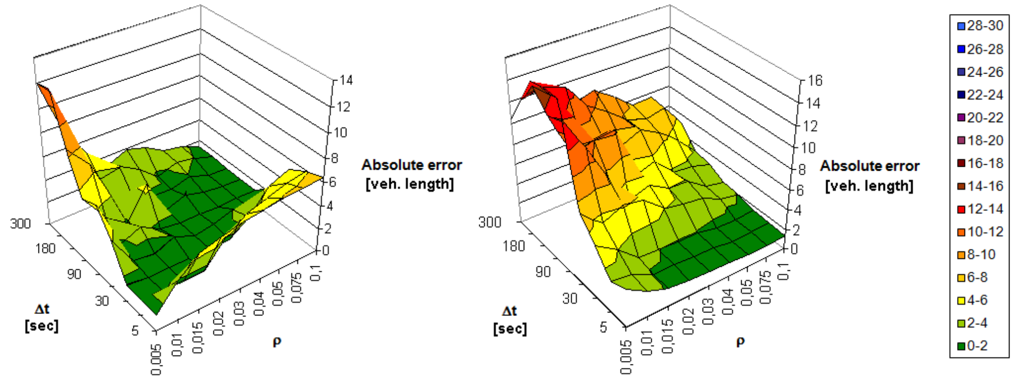

Figure 3 – Average estimation error (depending on penetration rate ρ and transmission frequency $1/\Delta t$) where $0 \le q \le q_{\text{crit}}$ and $\Delta T = 3600$ sec

Furthermore, even the collective processing of data from more than one probe vehicle system with different penetration rates and/or transmission frequencies is possible without any problems. At that, erroneous single data points affect the overall quality only marginally if the sample of probe vehicle positions is large enough.

Influence of data fusion

In order to identify sources of error in the estimation process as well as to point out solutions, it is useful to have a closer view on the estimates itself. In contrast to the above consideration of the average estimation error, the detailed analysis of distributions of the estimates yields important insights into the problems of queue length estimation. Since data fusion will eliminate or at least reduce the effects of some sources of error, the analysis starts with the case where all a-priori weights w_q are equal to 1, i.e. without data fusion.

a) $\Delta t = 5$ sec b) $\Delta t = 60$ sec

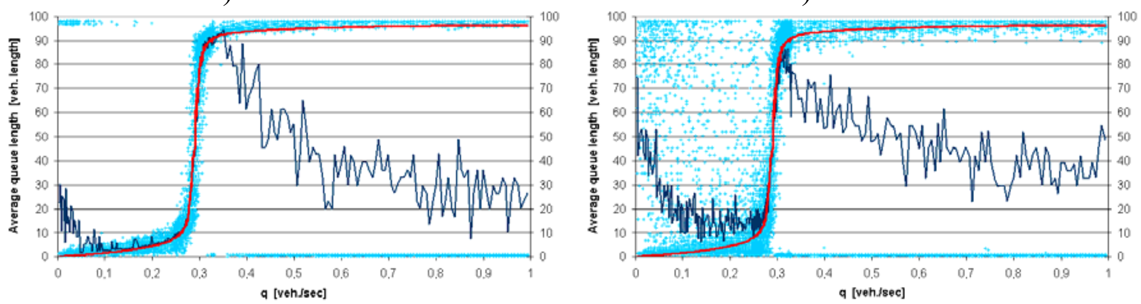

Figure 4 – Distribution of estimated queue lengths with mean value (blue line) in comparison to 'real' queue lengths (red line) from simulation without data fusion ($\rho = 2\%$, $T = 3600$ sec)

In doing so, at least three different sources of error can be identified. First of all, Figure 4a shows that there sometimes (when $q \ll q_{\text{crit}}$ or $q \gg q_{\text{crit}}$) is a strong tendency to a bimodal distribution. The new method seems to be unable to distinguish between very short and very long queues correctly. The reason is that in both cases the traffic density is constant nearly on the whole road section. Either it is equal to the jam density nearly everywhere or free flow density can be found on the complete road section except for the very few meters in front of the traffic signal. Thus, the final decision between very short and very long queues mostly depends on very few data points located near to the upstream and downstream exit of the considered road section. So, if there are just a few data points at the 'wrong' position, it is easy possible that the described estimation method yields the corresponding wrong result as can be seen in Figure 4a. As a consequence, of course, the average estimation error increases significantly when $q << q_{\mathrm{crit}}$ or $q >> q_{\mathrm{crit}}$. Especially, the fatal errors in Figure 1a at over-saturated traffic can be explained in this way.

Now, if the transmission frequency decreases (see Fig. 4b), a second source of error can be identified. Obviously, the reduction concerning the amount of data points has got the effect that in many cases the new method is not able any more to achieve sensible estimates at least at under-saturated traffic. There are just not enough data points to get constantly good estimation results.

Nonetheless, it looks somewhat curious (see Fig. 4b) that the effect of noisy estimates stops as soon as traffic demand reaches the region of over-saturation ($q > q_{\mathrm{crit}}$). However, there is a simple explanation. For, (probe) vehicles remain much longer on the considered road section in case of long queues as they are typical for over-saturated traffic conditions. Hence, assuming a constant Δt, there are much more data points as soon as the traffic demand q exceeds q_{crit}. Consequently, following general statistical arguments, the estimates are significantly less noisy than in case of $q < q_{\mathrm{crit}}$. If, however, the transmission frequency is reduced still more ($\Delta t > 180$ sec), the same noise arises at over-saturation, too.

Finally, the third source of error is a little bit more subtle and has (small) effects mostly when $q \approx q_{\mathrm{crit}}$, i.e. at the transition from under- to over-saturation. Although Figure 4 suggests quite good results at this point, it has been observable before (see Fig. 1) that the estimation error at $q_{\mathrm{crit}} \approx 0.289$ veh./sec nevertheless is slightly increased in comparison to several not as critical traffic demand values q.

The explanation is mainly a statistical one and is strongly related to the traffic dynamics near to the critical point between under- and over-saturation. It has been shown previously, e.g. by Brilon and Wu (see [9]), that the distribution of queue length in terms of vehicle counts at a fixed time within the traffic signal cycle becomes significantly broader when q approaches q_{crit}. Mathematically spoken, the standard deviation of this queue length distribution near to the critical traffic demand q_{crit} is much larger than for other q (i.e. $q << q_{\mathrm{crit}}$ or $q >> q_{\mathrm{crit}}$). Of course, the same is true if the distributions are considered as functions of average queue length (during a complete traffic signal cycle) and in terms of distances instead of vehicle counts (see Fig. 5).

Hence, if $q \approx q_{\mathrm{crit}}$, the range of possible traffic states (i.e. queue lengths) is much larger than in case of $q << q_{\mathrm{crit}}$ or $q >> q_{\mathrm{crit}}$. Schematically, the corresponding profiles of local traffic densities at the critical point between under- and over-saturation can be drawn as in Figure 6a where the average queue length L_0 ranges within wide limits, i.e. from very short queues up to a completely jammed road section nearly everything is possible. Indeed, these profiles are stable over more or less short time intervals and thus were observable during the simulations.

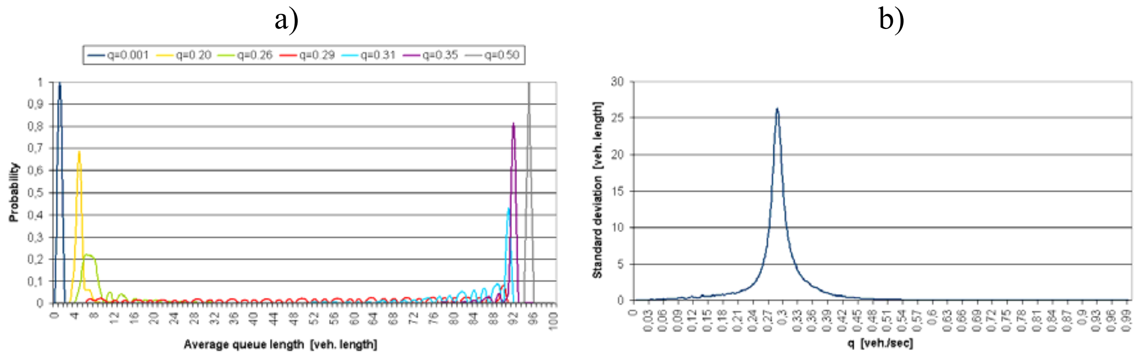

Figure 5 – a) (Stationary) probability distributions of the average queue length – b) Standard deviation of the average queue length

However, they are not stationary[4] in sense of a time-independent statistical equilibrium (cf. [9]) which principally can be obtained by summation of all non-stationary profiles each weighted by its (stationary) probability from Figure 5a. The resulting stationary profile in case of $q \approx q_{\text{crit}}$ is schematically shown in Figure 6b.

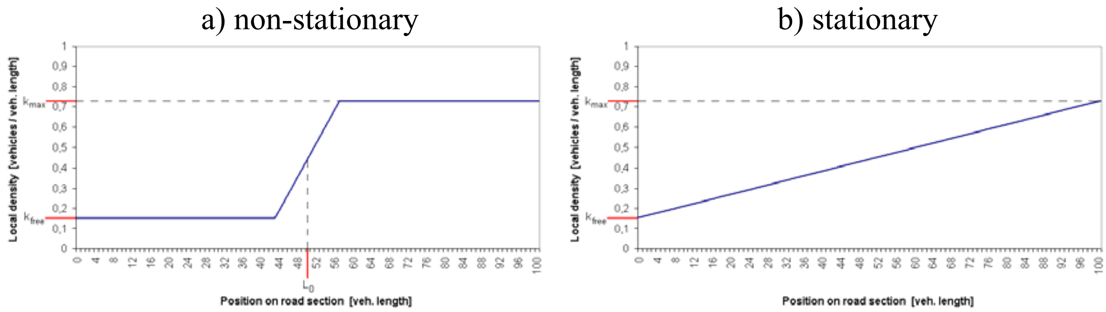

Figure 6 – Sketch of a) non-stationary and b) stationary profiles of local traffic densities ($k_{\max}$ = average jam density, k_{free} = free-flow density) in case of $q \approx q_{\text{crit}}$

Now, for the reason of simplicity, only stationary profiles of local traffic densities were taken into account during the estimation process for this paper. Hence, especially near to the critical traffic demand, appropriate profiles for the description of the occurring traffic states are missing sometimes. Indeed, the estimation method nevertheless tries to find that profile which fits in with the given probe vehicle positions best. But, due to the problem of non-stationary traffic (especially if $q \approx q_{\text{crit}}$), the estimation results cannot be expected to be just as well as in case of less critical traffic demand values q.

Obviously, this third type of error is immanent in the used algorithms and thus can only be eliminated by improvements of the estimation method itself[5]. The other described errors, however, can be reduced also by using data fusion, i.e. by adjusting the a-priori weights w_q. A first approach is to assume that because of additional information it is always known if traffic is under- or over-saturated. Adding a certain fuzziness (tolerance of 15% beyond the critical

[4] Due to the underlying mathematics, stationarity has to be interpreted as in the context of Markov chain theory.
[5] A possible solution could be the additional consideration of non-stationary profiles of local traffic densities (see Fig. 6a) during the estimation process. Further analyses in this context are part of future research.

traffic demand), the corresponding a-priori weights are defined as follows where q_{real} always is the correct traffic demand:

$$\text{If } q_{\mathrm{real}} \le q_{\mathrm{crit}} \text{ (under-saturation), } w_q := \begin{cases} 1 & \text{if } q \le 1.15 \cdot q_{\mathrm{crit}}, \\ 0 & \text{else.} \end{cases} \tag{2}$$

$$\text{If } q_{\mathrm{real}} > q_{\mathrm{crit}} \text{ (over-saturation), } w_q := \begin{cases} 0 & \text{if } q < 0.85 \cdot q_{\mathrm{crit}}, \\ 1 & \text{else.} \end{cases} \tag{3}$$

Figure 7 shows the resulting distributions of estimated queue lengths when this kind of data fusion is used. At that, with regard to a decided comparison, all other parameters are exactly the same as before (cf. Fig. 4).

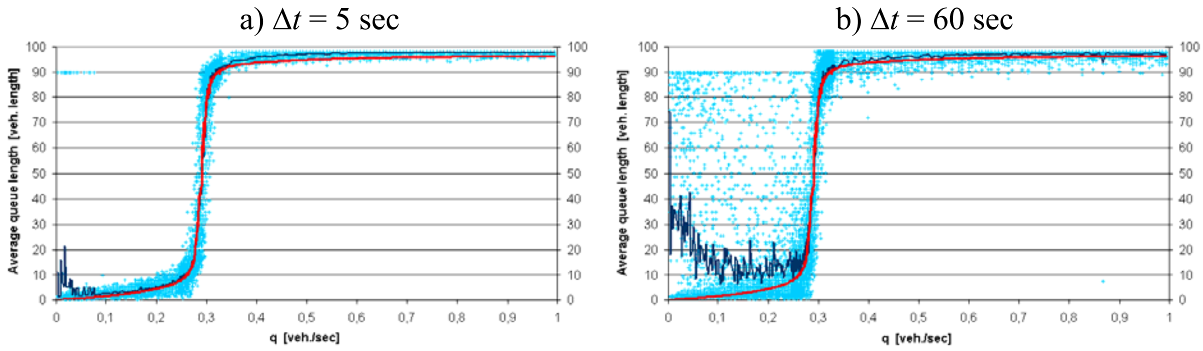

Figure 7 – Distribution of estimated queue lengths with mean value (blue line) in comparison to 'real' queue lengths (red line) from simulation with data fusion (differentiation between under- and over-saturation, ρ = 2%, T = 3600 sec)

As can be seen, the results are much better than in the situation without data fusion (cf. Fig. 4). Especially, the ambiguity error (first source of error) in case of $q \ll q_{\mathrm{crit}}$ and even more at over-saturation ($q \gg q_{\mathrm{crit}}$) is significantly lower or even nearly completely eliminated. Furthermore, the maximum deviation at under-saturation is reduced to about 90 vehicle lengths instead of almost 100 vehicle lengths before. Hence, an additional (but small) reduction of the average estimation error is achieved in general. Nevertheless, the problem of noisy estimates (second source of error) because of very few data points has not been solved (see Fig. 7b). Of course, the above way of defining a-priori weights w_q just excludes the most fatal errors, but still allows for wrong estimates within the whole range of under-saturation, respectively over-saturation (even with a fuzziness beyond the critical traffic demand). Thus, the results are not surprising.

However, selecting the a-priori weights in a more restrictive (but still realistic) way also reduces the effects of this second source of error to a certain degree. Thus, let w_q := 1 for all $q \in [q_{\mathrm{real}} - q_{\mathrm{tol}}; q_{\mathrm{real}} + q_{\mathrm{tol}}]$ and w_q := 0 else where q_{tol} is a suitable tolerance value. With q_{tol} := 0.1 veh./sec, for example, that describes the assumption that because of additional traffic information (e.g., from loop detectors, etc.) the real traffic demand q_{real} is approximately known up to ± 360 vehicles per hour. Now, of course, the effect of noisy estimates is limited to a band having the (horizontal) width $2 \cdot q_{\mathrm{tol}}$ (see Fig. 8b). Furthermore, another (small) reduction of the average estimation error can be achieved due the more restrictive way of defining the a-priori weights. Finally, even the last effects of the ambiguity problem (first source of error) are eliminated (see Fig. 8).

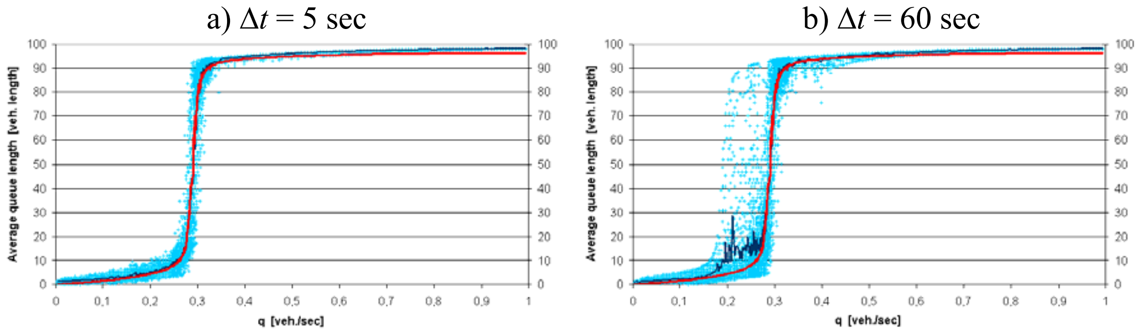

Figure 8 – Distribution of estimated queue lengths with mean value (blue line) in comparison to 'real' queue lengths (red line) from simulation with data fusion (traffic demand roughly known, q_{tol} = 0.1, ρ = 2%, T = 3600 sec)

In conclusion, Figure 8 shows the high quality of the described estimation approach. Realizing the potentials of data fusion consequently, the new method yields extremely good results even at comparatively short time intervals of data collection (T = 3600 sec) and at relatively small penetration rates (ρ = 2%). At that, the correlation between real (i.e. simulated) and estimated queue lengths can be expressed in terms of the correlation coefficient R. In case of Δt = 60 sec (cf. Fig. 8b), one gets R = 0.9776, and even R = 0.9985 when Δt = 5 sec (cf. Fig. 8a). Needless to say, this underlines the extraordinary quality of the obtained results once again.

Influence of the degree of aggregation

It has been mentioned before that the quality of estimated queue lengths essentially depends on the amount of available data points, i.e. on penetration rate ρ and transmission frequency $1/\Delta t$. Thus, raising these parameters is the generic way to get better results. For practical applications, however, the question has to be answered what to do if penetration rate and/or transmission frequency cannot be enlarged due to technical or financial limits. A simple but effective approach is the integration of historical data (observed at similar traffic conditions[6]) into current data sets.

With regard to the simulation-based evaluation presented in this paper, that means to aggregate virtual probe vehicle data from two or more simulation runs where traffic demand q as well as all other parameters are fixed. And indeed, in doing so, the average estimation error can be reduced significantly in case of sparse data densities (see Fig. 9a).

However, Figure 9b shows that aggregation also has got negative effects on the quality of the estimates if the general situation concerning data densities is quite well (cf. Figs. 2b and 3a). Especially near to the critical traffic demand q_{crit}, the estimation error becomes larger with increasing degree of aggregation as can be seen from Figure 9b, i.e. the larger the portion of traffic demand values $q \approx q_{crit}$, the larger the average estimation error.

The reason is that aggregation of data basically means to mix up different traffic states. Due to the stochastic nature of the queue at traffic signals (see Fig. 5), this however implies that

[6] Due to periodically repeating traffic patterns, a common approach is to integrate data observed on past days (typically the same or at least a similar day of week) at the same time of day.

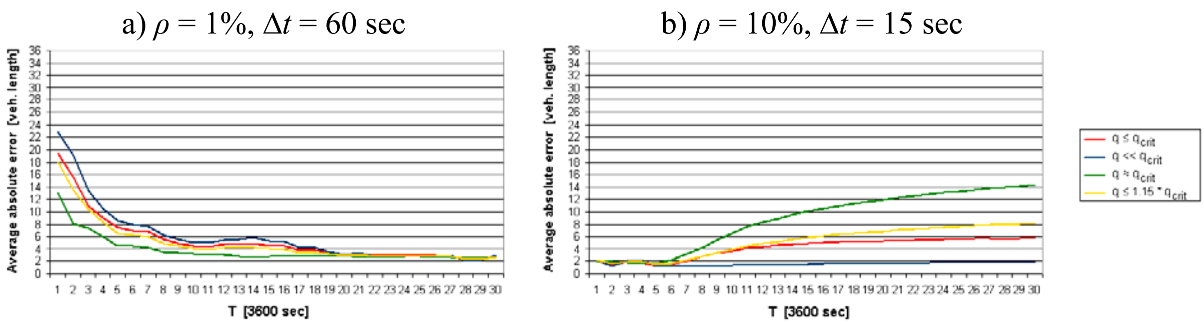

Figure 9 – Average estimation error (depending on degree of aggregation)

(especially near to q_{crit}) the resulting data sets contain data which are inconsistent in a certain way. Now, if penetration rate and/or transmission frequency become larger, the original data give an increasingly detailed picture of the traffic states being mixed up. Hence, the described inconsistency becomes more obvious for the estimation method. Thus, the corresponding error increases (cf. Fig. 9b) and finally overcompensates the gains of having more data points. At that, the largest acceptable degree of aggregation is correlated with the global data density.

Failure rates

Thus far, the analysis mostly focused on estimation errors. However, especially in case of online applications of the described queue length estimator, it is important to know about the failure rates β, too. Accordingly, since the method is based on probe vehicle data conceptually, the probability has to be estimated that during the whole time interval T of data collection not even one data point is registered. Figure 10 shows the empirical findings as average failure rates for different regions of traffic demand q.

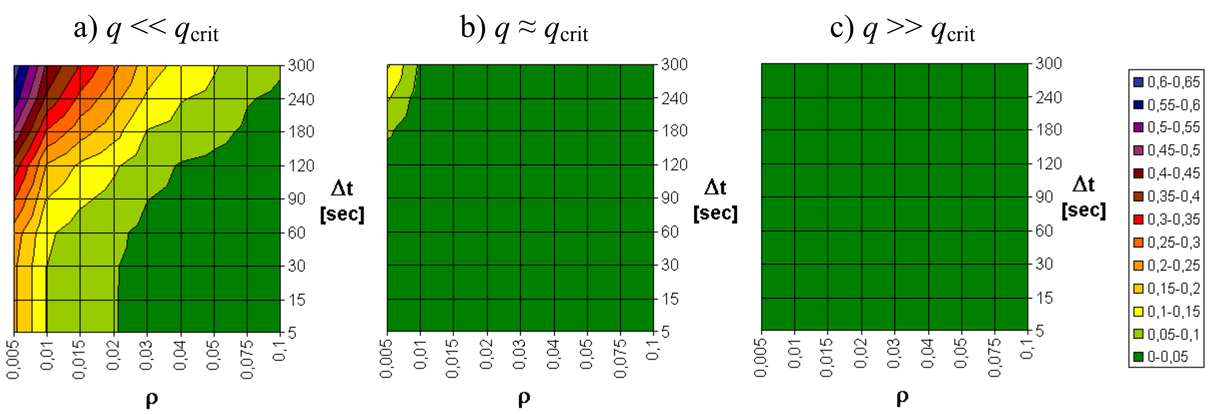

Figure 10 – Simulated average failure rates (depending on penetration rate ρ and transmission frequency $1/\Delta t$) where $T = 3600$ sec

As expected, the failure rates decrease with increasing penetration rates and transmission frequencies. Thereby, quite low penetration rates ($\rho > 2\%$) allow for failure rates less than 5% given sufficiently small Δt. Furthermore, a higher traffic demand reduces the failure rates, too.

Nevertheless, it looks somewhat curious that the failure rates β seem to be independent of the transmission frequency $1/\Delta t$ if Δt falls below a certain threshold (see Fig. 10a). However, a simple heuristic computation yields a reasonable explanation which also clarifies the general relation between failure rate β and the relevant parameters, i.e. q, ρ, Δt and T.

Obviously, the inflow rate of probe vehicles is defined by $Q \cdot \rho$ where Q is the general inflow rate[7] for given traffic demand q. Due the stochastic inflow following a Poisson distribution, then the probability $p(k)$ of observing exactly k probe vehicles during the time interval T of data collection is given by $p(k) = (Q\rho \cdot T)^k/k! \cdot \exp(-Q\rho \cdot T)$.

Now, neglecting all kinds of delay, each probe vehicle needs $N/v_{\max}$ seconds to pass the considered road section where N is the road length and $v_{\max}$ [m/sec] is the free-flow speed. Hence, due to the regular transmission frequency $1/\Delta t$, the probability p_0 of registering at least one position of a given single probe vehicle is defined by $p_0 = \min\{N/(v_{\max}\Delta t);1\}$. Thus, the failure rate β finally can be approximated by

$$\beta \approx \sum_{k=0}^{\infty} (1-p_0)^k \, p(k) = \exp\left(-Q\rho \cdot T \cdot \min\{N/(v_{\max}\Delta t)\,;1\}\right). \tag{4}$$

As can be seen, β decreases with increasing inflow Q (i.e. increasing traffic demand q) as well as with increasing ρ (cf. Fig. 10) and T. Furthermore, the failure rate β depends on the transmission frequency $1/\Delta t$ as long as Δt does not fall below $N/v_{\max}$. Now, inserting the same parameters as used above, this threshold becomes $\Delta t = 50$ sec which agrees with Figure 10a very well. At all, there is a close resemblance between Eq. (4) and the simulated results (see Fig. 11) which is not only qualitatively but also quantitatively.

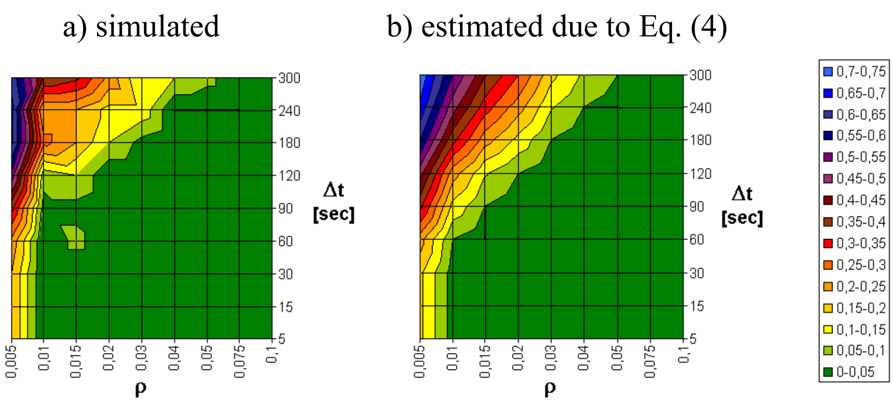

Figure 11 – Failure rates (depending on penetration rate ρ and transmission frequency $1/\Delta t$) where $T = 3600$ sec and $q = 0.1$ veh./sec

CONCLUSION

The evaluation results show that very good estimates for the traffic states at traffic signals can be found even at very sparse data densities in the context of current probe vehicle systems. Using an extremely flexible data fusion methodology, the quality can be enhanced significantly in contrast to the solitary probe vehicle approach. Even high-quality online applications with low failure rates seem to be possible given sensible penetration rates

[7] In [10], the relation between traffic demand q and inflow rate Q has been analyzed in detail.

($\rho > 2\%$) and transmission frequencies. Thus, the described approach (possibly extended for junctions without traffic signals) has got the potential to form the backbone of a highly reliable but cost-effective traffic monitoring system for urban road networks.

Nevertheless, the approach has to be tested more extensively with real data. For this purpose, also the problem of model calibration has to be solved. Because of that, just a few prototypical analyses (see [5]) with real data were possible as yet. Furthermore, the described data fusion options should be considered in more detail, i.e. the question should be answered how to choose the a-priori weights at best in respect of different types of additional traffic information.

ACKNOWLEDGEMENTS

The author wants to thank Dr. Peter Wagner for fruitful discussions about traffic data collection and traffic modelling. By interesting ideas, he and Dr. Rüdiger Ebendt initiated the work on the method described above.

REFERENCES

[1] Schäfer, R.P., Thiessenhusen, K.U., Brockfeld, E., Wagner, P. (2002). A traffic information system by means of real-time floating-car data, In Proceedings of *9th ITS World Congress* (Chicago, USA).

[2] German Aerospace Center, Institute of Transportation Systems (2003-2009). DLR-Cityrouter, URL: http://www.cityrouter.net/ (last access 2009-01-13).

[3] Mück, J. (2002). Schätzverfahren für den Verkehrszustand an Lichtsignalanlagen unter Verwendung halteliniennaher Detektoren, In Proceedings of *Heureka '02* (Karlsruhe, Germany).

[4] Priemer, C., Friedrich, B. (2008). A method for tailback approximation via C2I-data based on partial penetration, In Proceedings of *15th ITS World Congress* (New York, USA).

[5] Neumann, T. (2008). A cost-effective Method for the Detection of Queue Lengths at Traffic Lights, In Proceedings of *International Workshop on Traffic data Collection & its Standardization* (Barcelona, Spain).

[6] Neumann, T. (2009). TASEP related models with traffic light boundary, Eur. Phys. J. B **67** (1), pp. 133-138.

[7] Nagel, K., Schreckenberg, M. (1992). A cellular automaton model for freeway traffic, J. Phys. I France **2**, pp. 2221-2229.

[8] Brilon, W., Wu, N. (1999). Evaluation of Cellular Automata for Traffic Flow Simulation of Freeway and Urban Streets, Traffic and mobility: simulation – economics – environment, pp. 163-180.

[9] Brilon, W., Wu, N. (1990). Delays at fixed-time traffic signals under time-dependent traffic conditions, Traffic engineering & control **12/1990**, pp. 623-631.

[10] Neumann, T., Wagner, P. (2008). Delay times in a cellular traffic flow model for road sections with periodic outflow, Eur. Phys. J. B **63** (2), pp. 255-264.