We describe a layered evaluation approach to assess the accuracy of traffic speed estimation based on floating phone data. We designed a simulation framework to investigate the influence of various parameters on the estimation. Multiple quality metrics and combinations of metrics have been used and intervals of metric values have been mapped to quality ratings to enable a uniform and comprehensible presentation of results. We simulated traffic on several artificial and real maps and identified influencing parameters and their optimal values to achieve higher speed accuracy estimation. The results for artificial maps can be transferred to realistic maps with restrictions.

Keywords: Floating Phone Data, Data Quality, Traffic Information

INTRODUCTION

Traffic congestion is not only an annoying fact in individual transport, but has also severe impact on a country’s economy. In Germany, road freight traffic constitutes by far the biggest fraction of means of freight transport [1]. Hence, traffic congestion imposes a big efficiency problem in business models such as Just-In-Time production. Therefore, valuable traffic information, such as the current and prospective traffic state, has to be provided to turn delays to a foreseeable and manageable risk.

The creation of traffic information is based on data gathered in a road network. There exist different means which can be divided mainly into stationary and mobile detection mechanisms. Stationary detectors are placed along a road, e.g., inductive loops, radar detectors, or imaging devices [2]. High setup and maintenance costs are significant disadvantages of these methods. For mobile detection the cars themselves are used as probes. One common method is the creation of Floating Car Data (FCD). FCD includes only simple information, such as the position, while Extended FCD provides more detailed data gathered by sensors of the car, e.g., rain sensors. Another mobile approach to gather traffic information is the Observer method [3] [3], in which cars “sense” other cars driving in the opposite direction. Mobile approaches often suffer from low equipment rates and are mostly used in combination with stationary detection.

An emerging research field is the derivation of traffic information from anonymously located handsets (i.e., mobile devices which are able to connect to a cellular network). In analogy to FCD this is also termed Floating Phone Data (FPD). Most approaches using Floating Phone Data include the following steps [4]:
1. **Localization of the handset** as exact as possible.
2. **Map matching** to determine the position in a road network.
3. **Derivation of traffic information** using the estimated positions, e.g., link speed.

Of course, the derived information is not exact. The precision of the estimated values is influenced by several factors. The objective of this work is to find out, which factors impact the quality of traffic information to what extent. Based on the works of Fontaine et al. [5] [6] we evaluated different parameters and measured the resulting estimation errors by using the traffic simulation software VISSIM. This work has been performed in the course of the Cooperative Cars (CoCar) project¹ which investigates the suitability of 3G cellular network technologies (UMTS, HSDPA/HSUPA) and its foreseeable extensions for Car-2-X communication for different traffic applications enhancing driver safety and comfort [7].

There has already been a lot of related work done investigating the feasibility of using mobile phone data for traffic parameter estimation. These approaches can be categorized according to different dimensions. First, the works differ in the study types used; there are simulation (e.g., [5] [8]) or field studies (e.g., [9] [10]). Furthermore, different positioning techniques (if applicable) are studied, but handover techniques dominate, i.e., determining the position of a handset at the moment the handover takes place in an active call. Also the type of streets, such as highways or urban streets can be distinguished. Finally, the estimated parameters also differ: average link speed, travel time, traffic volume, and others have been investigated. All in all the reviewed works agree in the feasibility and the potential FPD can have for traffic applications. The feasibility is also shown by commercial systems, such as TomTom HD Traffic™. It determines the position of handsets by using a patented technique based on Timing Advance and fuses this data with GPS probe data, static detection, and other sources [11]. To determine which factors may impact such a monitoring system Crayford and Johnson [12] analyzed already in 2003 the influence of location accuracy, time between successive position estimates, and the sample frequency on the accuracy of determining the correct road using a microscopic traffic simulation. Fontaine et al. [13] [6] [5] studied accuracy issues of “Wireless Location technology-based Traffic Monitoring” very detailed by implementing a simulation framework which enables them to degrade data produced by the simulation according to the factors they studied. In [13] [6] they analyzed the influence of seven factors and their combinations on the accuracy and number of speed estimates per time period in artificial road networks. They extended their work in [5] by reevaluating two case studies in simulation and analyzing the results in comparison to their results in [6].

The contribution of our work is a layered evaluation approach based on the simulation framework of Fontaine et al. [5] [6]. The approach abstracts the layer of quality metrics which are used to determine the error of the estimated traffic parameters. To achieve this abstraction we adopt quality ratings to assess the accuracy of traffic estimation and use a mapping as an adaption layer which assigns an interval of metric values to a quality rating. This makes our evaluation approach very flexible and maintainable as the quality metric as well as the mapping are easily exchangeable. We employed cumulative distribution functions of quality ratings to compare and visualize the effect of different parameter settings. Utilizing this evaluation framework we examined different factors influencing the accuracy of traffic estimation with different artificial and realistic roadway networks. Besides the factors evaluated in [5]

¹ [http://www.aktiv-online.org/english/aktiv-cocar.html](http://www.aktiv-online.org/english/aktiv-cocar.html)
[6] we further investigated the influence of potentially problematic sources, i.e., pedestrians and busses, the distance between adjacent links and the measurement time interval size. For some of them we found a turning point, where the accuracy is no longer improving or even deteriorating when the corresponding factor changes. We also found that the adaptation of a metric for different speeds represents real traffic situations more accurately and results in a better approximation in real maps.

The remainder of this paper is organized as follows: In the next section, we introduce the setup of our test bed and explain its components. The section Evaluation and Results presents the metrics, evaluation methods and test scenarios used to assess the accuracy of the derived traffic information. Furthermore, the section details the results and findings of the experiments. The last section discusses the results and presents ideas for future work.

ARCHITECTURE

An overview of the basic components of our implemented architecture, which will be detailed in the following, is depicted in Figure 1.

DATA ACQUISITION

We obtain a ground truth for our tests by creating different roadway networks and simulating varying traffic situations using the traffic simulation software VISSIM. This data is later on compared with the estimated traffic information and the introduced error is determined. Firstly, we export a representation of the roadway network at hand. A network consists of links, connectors, and detectors, where each link is composed of lanes. Each pair of lanes of two links is connected by a connector. To determine the average speeds and travel times on the links, travel time detectors are placed on the links in the network and are also exported. Secondly, we record the status and position of every vehicle in the network every one tenth second in a simulation. The status and position of pedestrians on sidewalks are also recorded. After simulation the raw traffic data is stored in a database.

Figure 1: Architecture Overview

The next step is to create the baseline data. To investigate the estimation of traffic parameters from handset positions we have to create the corresponding instances of mobile devices in the data. Since every person in a car or a bus potentially can carry a handset, we created an instance for each passenger. Finally, we calculated the true average speed for each link.
DATA DEGRADATION
In a real life scenario handset positioning introduces an error. Depending on the positioning technique the precision varies extremely, from a few meters for GPS-based systems to several kilometers in UMTS. For handsets a positioning error of 150m (handset-based systems) has been demanded by the Federal Communications Commission to satisfy requirements of emergency calls [14]. To analyze how the positioning precision of handsets influences the traffic parameter accuracy we degraded the accuracy of the handset positions. Furthermore, the sampling rate, i.e., the frequency of handset positioning, and the number of handsets tracked simultaneously, are limited by the positioning technique. Hence, these are parameters whose impact has to be analyzed, too. Another problem is that a handset may be turned off or a call is dropped, that is, the position of the handset cannot be tracked anymore. A test framework (the application in Figure 1) has been implemented which enables the configuration of these factors and the modification of the original data in a degradation step.

MAP MATCHING AND TRAFFIC DATA ESTIMATION
The traffic simulation and the degradation of the data create a traffic situation which we now want to capture. Thus, the next step is to match the degraded positions to a roadway network, which consists of a set of roadway links, each roadway link is composed of a sequence of roadway sections and specified as a sequence of points. The task is now to estimate on which point of which roadway section a handset is located. This is called Map Matching.

For map matching there exist different types of algorithms which vary in matching precision. Furthermore, the accuracy of a map matching algorithm depends on the accuracy of the input positions and the accuracy of the digital map to match to. We implemented two map matching algorithms: a simple geometric map matching and a geometric map matching with topology (evaluating other more sophisticated map matching algorithms is easily possible in our framework, but this was not the focus of our work). The simple method matches the handset to a roadway section in its location area with the shortest minimum distance (i.e., the minimum of the perpendicular distance and the distances from the point to both endpoints of the road section) to the measured position. More complex map matching algorithms utilize available knowledge about the roadway network topology and the trajectory of a vehicle, also called geometric map matching with topology. In our implementation of topology map matching, we first determine the location area of the handset and thereby reduce the number of links to match to. Then we consider the last matched position (if available) and reduce the potential sections to a subset of links, each of which is either the last matched link or another nearby upstream or downstream link directly connected to the last matched link. Based on the last measured position and the current measured position a trajectory is formed and only those links are considered further whose heading is similar enough to the one of the trajectory. For the final subset the simple geometric map matching is applied. If no history of preceding positions is available, the simple method is used. For a survey on different map matching algorithms see [15].

After positions have been matched to a road network, they can be used to derive the speed of handsets. For each link, the average speed is calculated from the individual speeds each time one measurement interval has been completed. We also investigated the effect of the length of this measurement interval and vary it in our tests.

DATA QUALITY ASSESSMENT
The next step is to estimate the accuracy of the derived traffic data. We are interested in the following values:
(1) What is the distance between the measured position and the actual position?
(2) What is the difference between the mapped position and the actual position?
(3) What is the error ratio of the individual handset speed compared to the actual speed?
(4) What is the error ratio of the average link speeds compared to the actual average link speeds?

To determine the difference between the actual and the estimated traffic data, metrics have been developed for each parameter. For the measured and mapped positions the Euclidean distance is used to determine the error of position estimation. The accuracy of the speeds of the individual handsets and the average link speeds are assessed by calculating the error ratio, which is the absolute value of the difference between the true and estimated speeds divided by the true speed. The error ratio is used as it reflects the weight of a deviation depending on the level of speed in a more adequate way. The absolute error would not be sufficient. If for example a speed of 40 meters per second (m/s) is estimated to 35 m/s and another speed of 5 m/s is estimated to 10 m/s, the error would be 5 m/s in both cases. Intuitively, these two speed estimations should not have the same accuracy. In first evaluations we observed very high error ratios for congested road networks and hence revised this error ratio. For example, if the speed is 1 m/s and it is estimated with 2 m/s, then this is as bad as estimating a speed of 20 m/s with 40 m/s. However, the difference between 1 m/s and 2 m/s is not significant for estimating the traffic state: it represents congestion in both situations. Thus, we used the absolute error divided by 5 for speeds below 5 m/s and the error ratio for speeds above 5 m/s.

Furthermore, to assess the accuracy of the data in a comparable and comprehensible way, we defined a rating for each parameter. The rating is a function (shown in Equation 1) that assigns an interval of metric values to a positive integer \( r \) between 0 (poor) and 100 (excellent).

The choice of the coefficient \( k \) depends on the probability distribution of metric values and the worst possible metric values. For the position estimates we used \( k = 1.5 \), because an error of 150 meters is considered by the Federal Communications Commission E-911 to be the worst estimate for handset-based solutions [14] and is therefore assessed with the rating 0. Speed estimates with an error ratio greater than 100 percent are also regarded as worst estimates and are assigned the rating 0. Hence, the coefficient \( k = 0.01 \) is selected for estimated speeds.

\[
[100k, \infty) \rightarrow r, r = 0 \text{ and } \[(100-r)k, (101-r)k) \rightarrow r, r \in \{1, \ldots, 100\} \] \tag{Equ. 1}
\]

To compare different values of influencing factors for each traffic parameter the cumulative probability of the ratings is calculated. The probability of a rating value is computed by

\[
Prob[R = r] = \frac{\text{Number of estimates with rating } = r}{\text{Total number of estimates}}, r \in \{0, \ldots, 100\} \tag{Equ. 2}
\]

The probability mass function (PMF) of the random variable \( R \) (the rating function) is a function which calculates the above probability for all \( r \in \{0, \ldots, 100\} \). The cumulative probability of a rating can then be calculated as follows:

\[
Prob[R \leq r] = \sum_{i=0}^{r} Prob[R = i] \tag{Equ. 3}
\]

The cumulative distribution function (CDF) is a function, which maps the cumulative probability to all \( r \in \{0, \ldots, 100\} \).
Finally, we designed a data quality model and transformed it into a corresponding database schema to store the configuration of an experiment and its results.

**EVALUATION AND RESULTS**

We evaluated ten different parameters which might influence the accuracy of traffic data. According to Fontaine et al. [13] [5] those parameters can be categorized into system-based and network-based parameters. *Network-based parameters* are dictated by the properties of the roadway network and the traffic situation while *system-based parameters* depend on the configuration of the system acquiring the traffic data.

**NETWORK-BASED PARAMETERS**

One interesting question in traffic estimation based on FPD is if characteristics of the underlying roadway network can influence the accuracy of the estimated parameters. To analyze each of these characteristics we firstly used five artificial roadway networks and varied the values for the characteristics one by one, i.e., only one property had been changed at a time. The basic artificial network consists of two parallel arterial roads which are crossed by several streets in a fixed distance. In each direction the major roads have two lanes, minor roads one lane. The traffic simulation has been run seventy-five minutes for each experiment. After a start-up period about 800 vehicles are simulated. In the following, the analyzed characteristics and their values are described.

1. **Length of roadway links**: The length of a monitored road section between two intersections (or nodes) in meters. It is expected that longer links are more advantageous as e.g., positions which are near intersections maybe mistakenly matched to a wrong link. We analyzed major roads of 800 meters which is the typical distance between two inductive loop detectors on highways in Germany and an arbitrary longer length of 1600 meters.

2. **Distance between adjacent links**: The distance between two roads with the same orientation, i.e., parallel or almost parallel, in meters. We assume that a greater distance will result in a higher accuracy of positions and speeds as the matching will work better. In our experiments we used distances of 300 meters and 600 meters.

3. **Traffic flow**: The density of the traffic in vehicles per hour. It is expected that low levels of traffic flow are beneficial to the accuracy of traffic parameters, since the traffic speeds on uncongested roadway networks are more stable than those on congested roadway networks. We investigated traffic flows of 250 vehicles and 500 vehicles per hour on main roads as these are common traffic flows for roads in Germany. Minor roads are always loaded with 20% of the main roads traffic.

4. **Potentially problematic sources**: In a real traffic scenario, pedestrians, busses or trains are also sources of traffic data derived from FPD. Intuitively, these sources are harmful to the measurement of average link speeds, because pedestrians can be mistakenly sampled as vehicles and busses can be sampled as multiple vehicles. First, we compared situations with and without problematic sources. On major roads 50 persons per hour and on minor roads 10 persons per hour have been simulated. Ten percent of the vehicles are busses and on each bus there are nine passengers and a driver. Furthermore, we investigated how these problematic sources influence the traffic separately and analyzed scenarios with and without pedestrians and with and without busses.
The default values comprise a link length of 800 meters, an adjacent link distance of 300 meters, a fluent traffic flow and no problematic sources.

After using an artificial map, we analyzed if the same observations can be made using a real map. VISSIM provides a part of the roadway network of Redmond, WA, as a sample map. The map comprises an area of 15 square kilometers with 238 links, where the average link length is 255 meters. The map contained highways, rural roads as well as city streets.

**SYSTEM-BASED PARAMETERS**

System-based parameters are mainly dictated by the properties of the cellular network, the positioning method and the map matching algorithm.

1. **Number of simultaneously tracked handsets:** Positioning systems in cellular networks can only track a certain number of handsets at the same time. When a vehicle leaves the network or its position is lost, another vehicle is randomly picked up somewhere in the network so that the same number of vehicles will continue to be tracked. It is anticipated that tracking more vehicles is advantageous, because from a statistical point of view, a large quantity of vehicle samples always better reflect the generality of the traffic condition than a small quantity. In our experiments we tracked 25, 50, 100, 200 and 300 handsets simultaneously.

2. **Probability of instantaneously losing a handset:** In cellular networks a handset may be lost while being tracked due to dropping a call or switching it off. If the tracking of handsets is often suspended this will influence the estimates in a negative way. We experimented with probabilities of 1%, 3%, 10%, 20% and 30%.

3. **Time between two succeeding position readings:** Besides the number of simultaneously tracked handsets the positioning system also limits the frequency of position estimates. Intuitively, determining the vehicle positions frequently is favorable. More handset positions are retrieved, which results in a higher number of vehicle speed samples. The time is modeled by a normal distribution noted by $N(\mu, \sigma)$, where $\mu$ is the mean and the $\sigma$ the standard deviation. We analyzed the accuracy of traffic parameters with $\mu = 10, 20, 30, 45$ and 60 seconds. The standard deviation was always 5 seconds.

4. **Positioning error:** Depending on the positioning method used the accuracy of the estimated position varies. Obviously, the lower the positioning error is, the more accurate the positions and speeds are determined. Both, the horizontal and the vertical error, have been independently modeled by a normal distribution, noted $N(\mu, \sigma)$, where $\mu$ is the mean and the $\sigma$ the standard deviation. We analyzed errors with $\sigma = 50, 75$ and 100 meters. The mean has been always 0.

5. **Map Matching algorithm:** It is expected that a more accurate and intelligent map matching algorithm generates more accurate traffic data. We used the simple and topological matching described in Section *Map Matching*.

6. **Measurement time interval:** The measurement time interval represents the period of time for which the average link speeds are aggregated. Intuitively, prolonging the measurement time interval is beneficial to the measurement of average link speeds, because more vehicle speed samples can be generated in a long time interval than in a short one. We analyzed time intervals of 5, 10 and 15 minutes length following to [16] [17].

Default values for the parameters are: Number of simultaneously tracked handsets = 100, probability of instantaneously losing a handset = 10%, time between position readings =
N(30,5), positioning error = N(0,100), Map Matching algorithm = topological map matching and measurement time interval = 5 minutes.

RESULTS
First we will discuss the results of the experiments based on the artificial maps and afterwards the results obtained using the Redmond map. Recall that we analyzed the Cumulative Distribution Function (CDF) of ratings for measured positions, mapped positions, individual vehicle speeds and link speeds to determine the impact of the different factors on the accuracy of these traffic parameters.

Length of roadway links: The rating distributions of vehicle speeds and link speeds derived by using a network with major links of 1600 meters length show a substantially higher accuracy than the rating distributions of a network with 800 meters link length. This approves the results in [6], where this is accounted to the higher number of samples on longer links.

Distance between adjacent links: The experiments showed that the rating distributions of mapped positions, vehicle speeds and link speeds derived when using a roadway network with parallel roads with a distance of 600 meters are significantly better than the rating distributions of a network with 300 meters link distance.

Traffic flow: It turned out that the rating distributions of vehicle speeds and link speeds in a simulation without congestion are significantly better than with congestion.

Potentially problematic sources: As expected, the rating distributions of vehicle speeds and link speeds are significantly better in a network without simulation of pedestrians and busses than in a network including these sources. Furthermore, as shown in Figure 2 (the lower curve represents the better result) the pedestrians heavily influence the accuracy of the average link speeds while the passengers in busses have a smaller influence. The influence of passengers may be explained by a resulting smaller sample size of vehicles tracked.

![Figure 2: Influence of problematic sources on average link speed](image_url)

Number of vehicles tracked simultaneously: As can be seen from the results in Figure 3 increasing the percentage of the tracked vehicles in a certain value range is advantageous to get more accurate link speeds. The accuracy of the link speed estimation is improved, when the percentage of vehicle samples grows from 3% to 12% (tested with 25, 50 and 100 vehicles). However, no further enhancement can be discovered, when the percentage grows from 12% to 37% (tested with 100, 200 and 300 vehicles).
Figure 3: Influence of the number of simultaneously tracked vehicles

**Probability of instantaneously losing a vehicle:** Our experiments showed that suspending tracking vehicles with a high probability harms the accuracy of link speeds. The accuracy of the link speed estimation deteriorates slightly, when the probability of instantaneously losing a vehicle increases from 10% to 30%. However, the disruption of tracking within the reasonable scopes contributes to more accurate link speeds to some extent. The accuracy is slightly improved with an increase in the probability from 1% to 10%.

**Time between two succeeding position readings:** Our experiments showed that increasing the frequency of vehicle position measurements within reasonable bounds is favorable to get more accurate link speeds. In the tests the accuracy of the link speed estimation has improved, when the time between successive vehicle position readings decreased from 60 seconds to 30 seconds. However, an excessive increase in the frequency is harmful. The accuracy deteriorates with a further decrease from 30 seconds to 10 seconds. The bad results for a frequency of 10 seconds are probably due to the fact, that with high sampling frequencies the positioning error has a higher impact on the vehicle speed estimates than with lower sampling frequencies, because the distance between the positions is smaller. As illustrated in Figure 4, the rating distribution of link speeds obtained by reading vehicle positions every 30 seconds is just slightly better than the rating distribution for 20 seconds, which is significantly better than the rating distributions for 10 seconds.

**Figure 4: Experimental results for time between succeeding readings**

**Positioning error:** Figure 5 shows the rating distributions of measured positions (left) and average link speeds (right) with $E = N(0, 50)$, $E = N(0, 75)$, and $E = N(0, 100)$. Obviously, the measured position is better for smaller error rates, but the positioning error has only a small impact on the measured average link speed. Thus, FPD can be used for traffic state estimation although it might have a significant positioning error.
Figure 5: Experimental results for positioning error

Map Matching: As expected the rating distributions of the vehicle speeds and the link speeds obtained by applying the topological method have shown to be substantially better than the distributions using the simple method.

Measurement time interval: Surprisingly, by testing the 5-min, 10-min and 15-min time intervals, no obvious distinction can be found in the rating distributions for not only vehicle positions and speeds but also link speeds. The experiment shows that shortening the measurement time interval reasonably does not influence the accuracy of traffic estimates.

Experiments with the real map:
After first experiments with the real map, we recognized that the results from the artificial maps are not directly transferrable to the real world scenario. Similar results had been found by Fontaine et al. [5]. However, we analyzed the results for the real map in more detail and realized that the average speed of vehicles in the real map was much lower than that in the artificial maps. We concluded that for very low speeds our initial quality metric for speed estimates is not appropriate for real world scenarios. We adapted the quality metric using the absolute error divided by 5 for speeds below 5 m/s and the error ratio for speeds above 5 m/s as low speeds indicate congestion anyway. The results for the adapted metric compared to the original metric are shown in Figure 6.

Figure 6: Experimental results for the two metrics

Furthermore, we analyzed the characteristics of the roadway network. The real map contained many very short links (shorter than 600m) due to the urban roads which were also modeled in the map. As we already found out with artificial maps, that long links are beneficial for the estimation of link speeds, we separated the links of the real map in short and long links and measured the quality for the speed estimates separately. Figure 7 shows that the quality of the estimates for long links in the real map almost achieved the quality of the results for the artificial map. Thus, if FPD is used for traffic state estimation, it should be rather considered only for long links, or shorter links should be combined so that they form longer links.
**Figure 7: Experimental results for an artificial and a real map (shown on the right)**

**Discussion:** The results show that the results for artificial networks cannot directly be applied to real maps, due to the complex properties of the maps. But the experiment presented in the preceding section showed that an analysis of the characteristics of the network and an adaptation of metrics improve the situation. This leads to the general problem of executing simulation studies including ours – the complexity of real world traffic can only be simulated in a limited way. We used means, such as inclusion and evaluation of the influence of pedestrians in the traffic, to approximate the real world as far as reasonable for our purpose.

**CONCLUSION AND FUTURE WORK**

In this work we presented a layered approach to investigate the influence of several factors on the accuracy of speed estimates based on FPD. Our approach is extensible and flexible in several ways, as we can use different metrics and metric combinations, and different kinds of mappings which assign an interval of metric values to a certain quality rating. Furthermore, different map matching algorithms can be applied and compared. The described adaptations can be easily deployed without editing the underlying data structure of the database. We extended the list of investigated factors and the used test values presented in [6] [5] and found some interesting turning points in the experiments for the number of simultaneously tracked vehicles, the probability of losing a vehicle and the sample frequency.

The complexity of a real world map with inhomogeneous traffic flows are challenging for traffic monitoring based on FPD and more fine tuning of algorithms is required, as also indicated by Fontaine et al. [5]. However, our results have shown that the results of artificial maps are transferable to a subset of the links in a real map, i.e., the long links. As these longer links (as they often constitute the main roads) are also more important for traffic state estimation, the application of traffic information derived from FPD is feasible. We believe, that the better the underlying traffic model is understood, the more accurate traffic information can be derived. As we have proved, problematic sources, such as pedestrians have a heavy impact on the accuracy. If we are able to apply sophisticated filtering techniques to exclude such sources we can gain a further improvement of the estimates’ accuracy. In [18] for example the acceleration, speed and rotation rate of a road user with a handset is used to determine their motion traces and distinguish their type. The investigation and feasibility of such filtering methods will be subject of future work. We will also investigate the requirements for estimates according to their accuracy in a broader scenario. We will extend the framework and investigate the gain of the integration with other mobile sources, such as sensor and event data, for link speeds in a fusion architecture which has been presented in [19]. We also plan to investigate the derivation of other traffic information such as traffic states based on FPD. This includes the assessment of the suitability and accuracy of different data mining algorithms.
Acknowledgements: This work has been supported by the German Federal Ministry of Education and Research (BMBF) under the grant 01BU0690 (Project Cooperative Cars, www.aktiv-online.org/english/aktiv-cocar.html) and by the Research Cluster on Ultra High-Speed Mobile Information and Communication UMIC (www.umic.rwth-aachen.de). We would like to thank the PTV AG, who kindly provided us VISSIM and VISUM licenses.

REFERENCES