City-wide bottleneck and deficiency analysis on a road network generated from the Open Street Map road network using Floating Car Data (FCD)

Josephine Grau*[1], Lea Fuchs[1], Torben Lelke[2], and Peter Vortisch[3]

[1]M.Sc., Institute for Transport Studies, Karlsruhe Institute of Technology, Germany
[2]M.Sc., Institute of Transportation and Urban Engineering, Technische Universität Braunschweig, Germany
[3]Prof. Dr.-Ing., Institute for Transport Studies, Karlsruhe Institute of Technology, Germany

Short summary

The German Highway Capacity Manual (HBS) and the German Guideline for Integrated Network Planning (RIN) rely on travel time distributions to assess the Level-of-Service (LOS) of roads and road networks. Usually, these values are generated by traffic measurements or with the help of traffic flow simulations. In recent years Floating Car Data (FCD) has become an essential data source for analyzing traffic quality because of its easy accessibility and growing coverage. This paper proposes a method to perform a city-wide analysis on the Open Street Map (OSM) road network using raw FCD. Therefore, OSM road segments of a city are aggregated to longer network sections on which travel times are estimated. Performance indicators can be calculated using these travel time distributions. Examples are shown for the cities of Karlsruhe and Hannover.
Keywords: floating car data, incident analysis, open data, key performance indicators, big data analytics

1 Introduction

The German Highway Capacity Manual (HBS) is the primary tool used to evaluate the quality of traffic flow. However, the calculation methods of the HBS are not generally applicable, e.g., for roads with a speed limit of 30 km/h or if cyclists are guided in mixed traffic (HBS (2015)). Since the required conditions are often not met in reality, especially in inner-city areas, it is necessary to use different means for the evaluation. This paper will present such an alternative by evaluating FCD.

FCD offers the advantage of mapping historically long periods of data obtained from vehicles moving in traffic. It can add to classical stationary detectors by providing information about driven trajectories. Other than stationary detectors, only a subset of the total traffic volume is represented. However, preceding research shows that even low-frequency FCD (around four data transmissions per minute) has adequate statistical power to calculate travel time distributions (Schäfer et al. (2003), Jenelius & Koutsopoulos (2013)). Furthermore, FCD can be used to analyze traffic flow regarding incident recognition. Altintasi et al. (2017) showed that FCD could recognize traffic patterns defined by LOS and detect congestions and bottleneck locations.

For an intuitive analysis of a city-wide road network, travel time indices are helpful. The TomTom traffic index, for example, uses FCD to generate several indicators, e.g., an average travel time per 10 km and offers thus the possibility to compare different cities (*TomTom Traffic Index* (2023)). Further studies developed and evaluated several indices for motorways (see, for example, Peter et al. (2021), and Radde et al. (2016)) and for urban road networks (Ulm et al. (2015)).

Similar to the study conducted by Axer & Friedrich (2014), we use historical FCD, transformed to trajectories, to calculate average travel times on all road network links resulting in traffic quality indices using the traveled distance and the elapsed time between two data transmissions. In Axer & Friedrich (2014), FCD was mapped to the OSM road network via the TMC (traffic message channel) system, used as an intermediary georeferencing tool. We use a more straightforward approach and map the FCD directly to the OSM road network. This offers an analysis method using FCD for a whole city road network. The determination of decision values (travel time distributions and percentiles) within the framework of the German guidelines of the HBS and the RIN is possible. The aim is to provide practitioners with a time and cost-efficient method to evaluate the quality of every single link in an urban network and anchor this method in the mentioned guidelines.

2 Methodology

In order to perform a city-wide analysis, the following method is proposed. First, a suitable road network needs to be constructed. Our work is based on the OSM road network, as it is publicly available. Roads in the OSM network consist of so-called OSM ways, each of which is a list of nodes representing a small road segment of some meters. Inspired by the functional road classes defined in the German Guideline for Integrated Network Planning (RIN (2008)), these OSM ways are grouped into longer road segments representing, for example, main roads, which we refer to as network sections in the following. The FCD can be used to estimate a travel time or speed distribution for each network section. Finally, performance indicators are calculated to evaluate the traffic conditions on each network section. Network sections with low performance measures can be analyzed in greater detail in further steps.

The OSM import is done via the Python OSM API packages *osmapi* (2023) and *overpass* (2019). As they are decisive for travel planning, only higher road classes are considered. Therefore, e.g., residential roads are not taken into account. The road classes are defined by the OSM attribute "highway". Besides that, OSM offers many other attributes for the OSM ways. For this application, we use the id of the OSM way ("id"), the allowed maximal speed on the OSM way ("maxspeed"), whether the OSM way is a oneway road ("oneway"), and the coordinates of the nodes ("geometry.coordinates"). The information is then augmented by node ids, start and end nodes of the OSM ways, and a doubling of non-oneway OSM ways to separate the two driving directions.

After this preparation, the automatic identification of network sections is made. The general idea is to add an OSM way to a network section as long as it belongs to the same road category and does not cross an OSM way of the same or higher-order road category. In the end, each direction of an OSM way is part of precisely one network section. The algorithm works as follows: at the beginning, there are no network sections, and no links of the network are assigned to a network section. Each algorithm step randomly selects a link that must still be assigned to a network section. This link establishes a new network section from which the section is extended in and against the direction of travel until a node is reached where routes of the same or higher category cross. This determines and delimits the network section, and the procedure continues with the next unassigned link. Directions of travel are considered separately, so OSM ways representing two directions of travel must be duplicated and used as two separate links. After all links are assigned, the procedure is finished. For the choice of the next link to be attached to a network section, the start and end nodes and angles are considered, i.e., OSM ways are not attached to a network section if a steep curve occurs at the current node. This reflects typical situations in the network structure and avoids, e.g., U-turns. Particular attention should be paid to avoid round trips; thus, it has to be checked whether the following link is already part of the considered network section.

This method provides plausible and intuitive results in the OSM road networks of the considered cities. The algorithm depends on a correct assignment of coordinates and attributes in OSM, which is mostly the case. The following figures provide some insights into the algorithm's results.

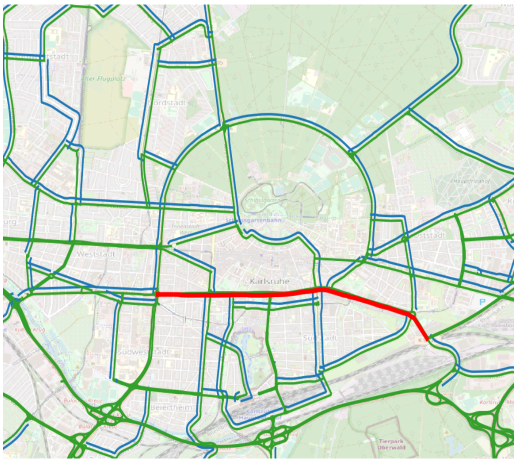

Figure 1: Recognition of the network section Kriegsstraße in the city center of Karlsruhe

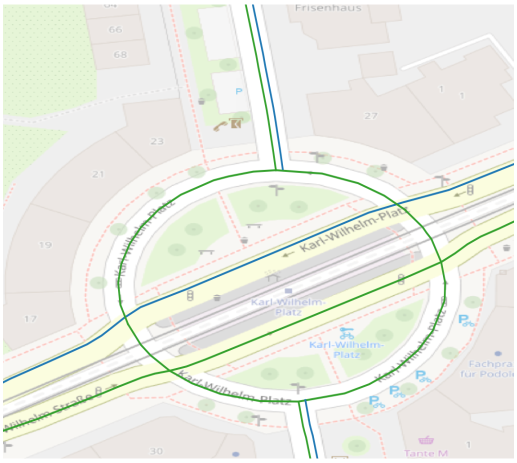

Figure 2: Recognition of a complex roundabout with separation of driving directions

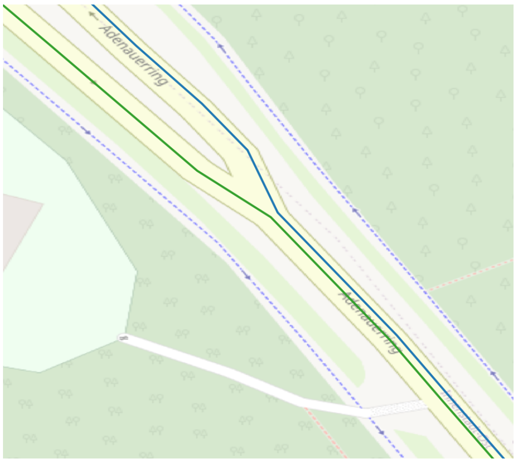

Figure 3: Continuation of the driving directions on roads with structural separation of the driving directions

Figure 4: Faulty detection of two driving directions due to a wrongly assigned oneway attribute in OSM

The following step uses FCD to calculate travel times on the network sections. To obtain a reliable evaluation, the FCD must be processed. First, we use only the weekdays Tuesday, Wednesday, and Thursday to compare similar demand situations that represent typical working days. The data quality is ensured by excluding implausible high speed values and cutting trajectories at stops longer than 60 seconds.

For each network section, each direction's start and end nodes are used to define a catch radius to filter floating cars traversing the entire section. This imitates the principle of an ANPR measurement. Travel times were calculated from the difference between the time stamps. A travel speed distribution can be determined using the haversine function to calculate road lengths. The percentiles of these distributions are used to calculate incident indices. For this step, raw FCD compared to aggregated FCD information is essential. Calculating travel time (or travel speed) distributions of a network section from shorter road segments is impossible. Mathematically, the random variables representing travel times on different road segments are not independently distributed. Therefore, a convolution of travel time distributions to obtain overall distributions is incorrect.

Next, several indicators are used to evaluate traffic flow on the network sections. The reliability index, as proposed in Peter et al. (2021), is calculated as the quotient of the 90% and 50% percentile of travel time denoted by $t90$ and $t50$ accordingly:

$$Reliability\ Index = \frac{t90}{t50}. \tag{1}$$

If the travel times vary little, there are hardly any disturbances ($t90$ and $t50$ are low) or permanent disturbances ($t90$ and $t50$ are high) and the reliability index is close to one. The higher the reliability index, the more the travel times fluctuate on the network section. In particular, the reliability index does not indicate whether there are disruptions but whether the same traffic condition occurs reliably. The reliability index has the same significance as the buffer time index (BTI) used in Radde et al. (2016) normalized with the 50% percentile. It can be interpreted as additional time to the median travel time to be considered in planning to arrive on time. The index is calculated by the difference between the 90% and 50% percentiles and a normalization for comparability:

$$BTI = \frac{t90 \ - \ t50}{t50} \ = \ \frac{t90}{t50} \ - 1 = Reliability\ Index - 1. \tag{2}$$

Another index proposed by Peter et al. (2021) is the travel time index which evaluates the ratio between the 50% percentile of the travel time and a target travel time t_{target}. We use the travel time when driving at the maximum permitted speed to approximate t_{target}. This is a plausible assumption for a city road network of higher-order road categories.

$$Travel\ time\ index = \frac{t50}{t_{target}} \tag{3}$$

The above indices refer to travel time percentiles, which are robust quantities. However, an aggregation of travel time distributions over all daytimes only gives a rough overview and does not account for peak hours. On the other hand, indices with higher temporal or spacial resolutions can only be used if a sufficiently large amount of data is available. This would apply when evaluating the indices above for daytime hours. Another index of that category is the Cumulated Differences Index (CDI) from Radde et al. (2016):

$$CDI = (n-1)^{-1} \sum_{t=1}^{n-1} |\,(v_{t+1} - v_t)|, \tag{4}$$

where n denotes a considered number of days and v_t denotes the mean speed at the considered hour on day t. This index reflects user expectations based on the fact that travel speeds do not vary from day to day. A high CDI indicates the low reliability of the route segment. The index provides good results in the situation considered by Radde et al. (2016) on highways with high traffic volumes. In contrast, it may be considered unsuitable for urban roads with few vehicles per hour at low FCD penetration rates. Indicators that require many data sets due to high temporal or spatial resolution may be of limited suitability for FCD evaluation.

3 Results and discussion

The method explained above is used for the cities of Karlsruhe and Hannover in Germany. For Karlsruhe, we used FCD from September 2021, whereas for Hannover, data were available for the entire year 2021. As mentioned above, only Tuesdays, Wednesdays, and Thursdays were considered. The transmission interval of the FCD is mostly around five seconds.

The following pictures show the results for the reliability index (formula 1) and the travel time index (formula 3) for the city of Karlsruhe. It can be seen that the two indices make different statements but lead to plausible results. They allow the identification of network sections with poor traffic conditions, which can be subjected to further, more detailed analyses. For example, one can take the Kriegsstraße, the east-west axis in the center of Karlsruhe, which performs poorly in both indices as expected. It was the subject of extensive reconstruction measures until 2022.

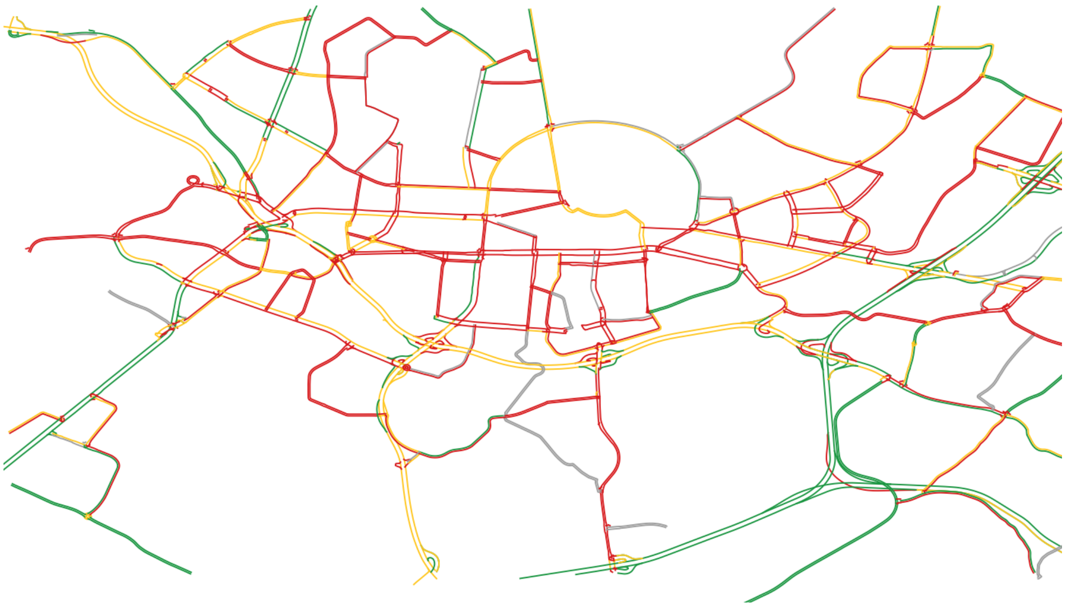

Figure 5: Reliability index Karlsruhe (green: value range 1 - 1.5, yellow: 1.5 - 2, red: > 2)

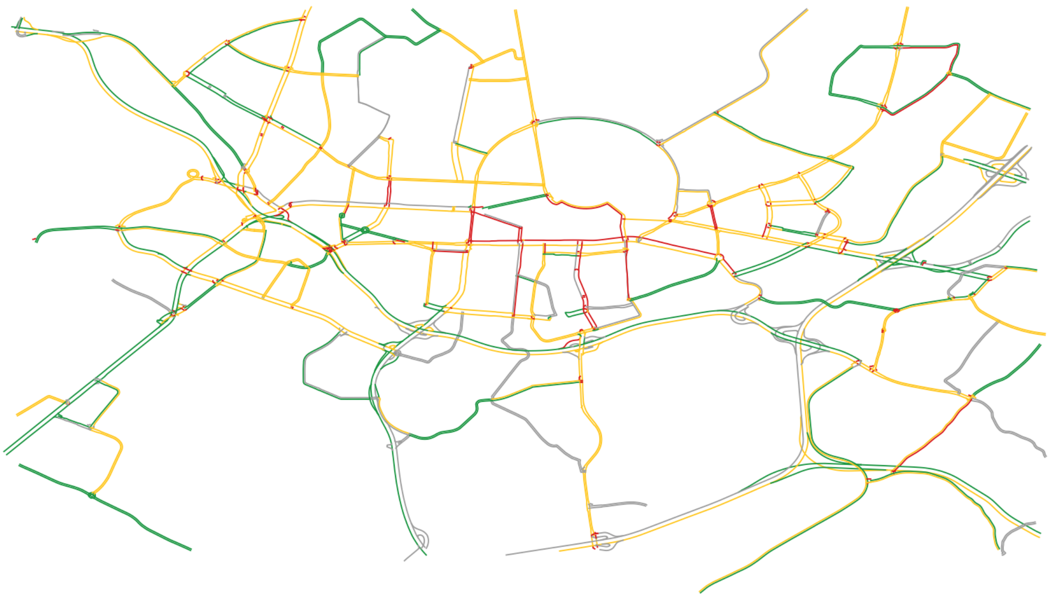

Figure 6: Travel time index Karlsruhe (green: value range 0 - 1, yellow: 1 - 2, red: > 2)

As we had enough data for Hannover, the CDI (formula 4) could be evaluated. The figures 7 and 8 show the results for 7 a.m. and 11 a.m. The results do not show a considerable variation between different evaluated hours.

Figure 7: CDI Hannover 2021, 7 a.m.

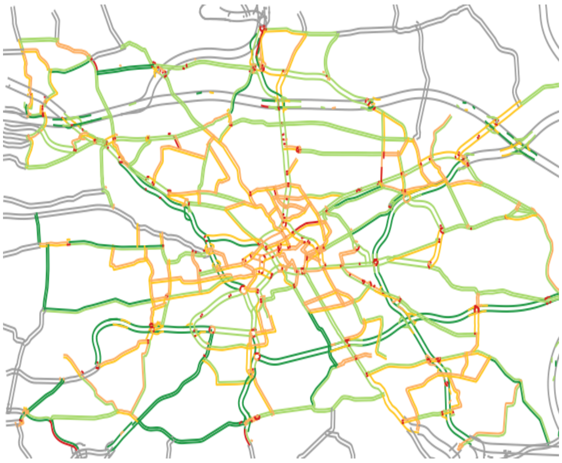

Figure 8: CDI Hannover 2021, 11 a.m.

These indices were chosen as examples from the literature. When interpreting the indices, it is important to ensure that they reflect the different calculation methods. The performance index must be chosen according to the intended application.

The indices rely on FCD's high penetration rates, transmission frequencies, and coverage of all road segments. For the reliability index and the travel time index, an average of 640 vehicles per network section were evaluated in the example of Karlsruhe. For the CDI, around 128 days with an average of 5.6 vehicles per hour were available as a data basis for each network section. Thus, the travel time distribution's representativeness for calculating the CDI is significantly less robust than the distribution used for the other indices. The evaluation also depends on the chosen color scale, which is different for every index and might be shifted to one or another direction based on the knowledge of the local situations or the intention of the statement. A particular artifact of the simple travel time calculation with catch radii is the misevaluation of short segments, roundabouts, or motorway links, where start and end nodes lie too close together. This problem could be tackled using more refined techniques to process the FCD, notably a map-matching algorithm.

4 Conclusions

The presented method offers an efficient and valuable procedure to gain a quick overview of a city's traffic conditions. It uses raw FCD and evaluates performance indices based on travel time distributions on network sections. The road network is constructed using open-source data, namely OSM ways. The presented method can provide an extension to the existing methods of the German Highway Capacity Manual (HBS) or the German Guideline for Integrated Network Planning (RIN) to evaluate traffic quality.

The method could further be used to identify problematic road segments, which can be analyzed in more detail to find the reasons for disruptions and bottlenecks. A similar analysis is also conceivable for other transport modes, e.g., using floating bike data.

Acknowledgements

The research for this paper was conducted as part of the project "Bottleneck Analysis for Urban Road Networks" funded by the German Federal Ministry for Digital and Transport, represented by the Federal Highway Research Institute (BASt). We want to thank our project partners for their excellent cooperation and for providing FCD for the city of Hannover, as well as our colleagues Marvin Baumann and Claude Weyland for their advice and ideas on this work.

References

Altintasi, O., Tuydes-Yaman, H., & Tuncay, K. (2017). Detection of urban traffic patterns from floating car data (FCD). , *22*, 382–391. (Publisher: Elsevier)

Axer, S., & Friedrich, B. (2014). Level of service estimation based on low-frequency floating car data. , *3*, 1051–1058. (Publisher: Elsevier)

HBS. (2015). *Handbuch für die Bemessung von Straßenverkehrsanlagen: Teil S: Stadtstraßen.* FGSV - Forschungsgesellschaft für Straßen- und Verkehrswesen.

Jenelius, E., & Koutsopoulos, H. N. (2013). Travel time estimation for urban road networks using low frequency probe vehicle data. , *53*, 64–81. (Publisher: Elsevier)

osmapi. (2023). Retrieved 2023-02-28, from `https://pypi.org/project/osmapi/`

overpass. (2019). Retrieved 2023-02-28, from `https://pypi.org/project/overpass/`

Peter, L., Janko, J., Schick, N., Waßmuth, V., Friedrich, M., & Bawidamann, J. (2021). Entwicklung eines aktuellen, echtzeit-verfügbaren Key Performance Indicator (KPI) Systems für das deutsche Autobahnnetz.

Radde, M., Leerkamp, B., & Klemmer, T. (2016). Messung der Zuverlaessigkeit von Strassennetzen mit Floating Car Data. , *60*(11).

RIN. (2008). *Richtlinien für die integrierte Netzgestaltung.* FGSV - Forschungsgesellschaft für Straßen- und Verkehrswesen.

Schäfer, R.-P., Gühnemann, A., & Thiessenhusen, K.-U. (2003). Neue Ansätze im Verkehrsmonitoring durch Floating Car Daten.

Tomtom traffic index. (2023). Retrieved 2023-02-28, from `https://www.tomtom.com/traffic-index/about/`

Ulm, M., Heilmann, B., Asamer, J., Graser, A., & Ponweiser, W. (2015). *Identifying congestion patterns in urban road networks using floating car data.*