Visual Analysis of Floating Taxi Data Based on Interconnected and Timestamped Area Selections

Andreas Keler and Jukka M. Krisp

Abstract Floating Car Data (FCD) is GNSS-tracked vehicle movement, includes often large data size and is difficult to handle, especially in terms of visualization. Recently, FCD is often the base for interactive traffic maps for navigation and traffic forecasting. Handling FCD includes problems of large computational efforts, especially in case of connecting tracked vehicle positions to digitized road networks and subsequent traffic state derivations. Established interactive traffic maps show one possible visual representation for FCD. We propose a user-adapted map for the visual analysis of massive vehicle movement data. In our visual analysis approach we distinguish between a global and a local view on the data. Global views show the distribution of user-defined selection areas, in the way of focus maps. Local views show user-defined polygons with 2-D and 3-D traffic parameter visualizations and additional diagrams. Each area selection is timestamped with the time of its creation by the user. After defining a number of area selections it is possible to calculate weekday-dependent travel times based on historical taxi FCD. There are 3 different types of defined connections in global views. This has the aim to provide personalization for specific commuters by delivering only traffic and travel time information for and between user-selected areas. In a case study we inspect traffic parameters based on taxi FCD from Shanghai observed within 15 days in 2007. We introduce test selection areas, calculate their average traffic parameters and compare them with recent (2015) and typical traffic states coming from the Google traffic layer.

Keywords Floating car data (FCD) · Traffic map · Focus map · Travel time estimation · Velocity estimation · Movement data · Visual analysis · Linked views

A. Keler (✉) · J.M. Krisp
Institute of Geography, Applied Geoinformatics, University of Augsburg,
Alter Postweg 118, 86159 Augsburg, Germany
e-mail: andreas.keler@geo.uni-augsburg.de

G. Gartner et al. (eds.), *Progress in Cartography*, Publications of the International Cartographic Association (ICA), DOI 10.1007/978-3-319-19602-2_8

1 Introduction

Floating Car Data (FCD) is a relatively new technology, which appeared as a useful source of information about the actual traffic situation (Liu et al. 2012). It is acquired by a tracking device inside a vehicle, which is in this case the sensor itself (Cohn and Bischoff 2012). FCD is often provided by whole taxi fleets for long time periods. The resulting massive data sets are difficult to handle with common analysis and visualization methods and need advanced methods. In the case of FCD of a taxi fleet, information on the traffic situation within selected investigation areas is detectable due to thousands of simultaneously tracked vehicles. Besides detection of traffic patterns, the traffic information visualization is challenging. One common possibility for describing the traffic situation of urban areas is the use of interactive traffic maps. These products allow users to detect traffic movement patterns, visualized by only differently colored road segments in a scalable 2-D map.

These maps are not only used as traffic prediction services but as well as tools for transportation planning (Sohr et al. 2010). The classification of two-dimensional road segments in traffic maps can be provided qualitatively by the values for traffic congestion high, medium, low (Liu et al. 2008; Goldsberry 2005) or quantitatively by classes of average velocity (Goldsberry 2008). The connection between traffic data coming from FCD and digitized road segments is realized by Map Matching (MM) algorithms. For the case of pneumatic counters (Dirks et al. 2003) and inductive loops (Leduc 2008), the connection to the road element is already set previously. Since there are still positioning errors in the acquired positions by using GNSS devices the connection with road segments is more difficult, especially in areas near intersections. That's why Map Matching (MM) has become a frequently used group of methods with over 36 different algorithms in the year 2006 (Zhao et al. 2012; Quddus et al. 2007). We have to respect the fact that MM achieves differing results and quality dependent on the different inspected road networks. In case of complex transportation infrastructure elements, matching is difficult. Another fact is the high computational effort of MM (Zhao et al. 2012), especially in dense situated road segments in large surface areas.

Keeping this in mind, we want to design a simpler way to connect vehicle trajectories to road segments. One basic aspect in doing this is to localize personalized areas of interest. The idea behind this is to derive detailed vehicle traffic information in short spatial distance to personal points of interest (POI). The arcs of the inspected road network are then enriched with information on relationships between the defined selection areas. One simple example for relationships is the connectivity of road segments between two area selections and the calculated traffic-aware travel times.

The base for this kind of personalization is the recording of time stamps, representing time of creation and inspection of each user-defined selection area.

In terms of visualization, these user-defined area selections are not following the already mentioned interactive traffic maps but other approaches, where visual analysis tools with linked views are used. In these examples large FCD sets were visualized by different interconnected views on the data (Tominski et al. 2012; Guo et al. 2011; Ferreira et al. 2013; Wang et al. 2013). All these examples make, similar to our idea, use of area selections for recorded tracks of moving objects with the aim of representing more detailed information at certain locations. Our idea consists of designing a geovisualization tool for interactive inspection of historical traffic information derived from taxi FCD. We follow the idea of a focus map (Zipf 2002; MacEachren 1995; Freksa 1999), where important areas are emphasized in a specific way. Based on our idea the emphasized areas are user-defined and represent average values of travel times, vehicle speeds and vehicle densities in a so called local view on the data. The global view corresponds to the focus map itself, which is simply a usual overview map containing the emphasized selection areas that have varying size, shape and a time component. The last mentioned is generated user-dependently and may give some support for inspecting time-dependent traffic pattern.

Within user-generated selection areas (in "local view") other points of interest (POI) may be included, such as "my house", "the place I work" or "where I want to go". In the end we can define different types of relations between the user-generated selection areas.

2 Analysis and Visualization Methods for Large FCD Sets

The analysis of massive FCD includes often data visualization, which might help to discover new insights into various traffic patterns. The visual analysis process itself is dynamic and implies testing of different visualization techniques. This follows rudimentary the idea of an exploratory data analysis (EDA). In addition to this fact we have the established visual analytics methods for movement data. Starting from both ideas, we want to discover if personalized analysis on FCD benefits from personal inspection of individual mobility and travel times.

2.1 Visual Analysis of FCD

The possibly simplest method for visualizing massive positions of vehicle movement, especially for the case without pre-processing, includes the generation of Dot maps (Stanica et al. 2013). The dots are often represented with differing color based on the classification of instantaneous attribute values, such as velocity or driving direction. Dependent on the selected data partition more or less point overlapping appears. The reason for this appearance results often from the high number of tracked vehicles. Due to overlapping of the dot symbols it is difficult to detect stop-and-go traffic patterns (Liu and Ban 2013). This might be solved by aggregation methods (Andrienko and Andrienko 2013). Andrienko and Andrienko (2007) for example use grid cells to summarize the values describing traffic (velocity, vehicle density). Additionally it is possible to represent individual movement of an object as a sequence of grid identification numbers (Moosavi and Hovestadt 2013). The visual representation of the latter might imply coloration or other symbolization of grid cells. Influenced by point aggregation, Sun and Li (2012) use of a pyramid-based approach for the visual exploration of large FCD sets.

Other visual analysis approaches use point clustering techniques on FCD for the detection of traffic patterns. Krisp et al. (2012) extract pick-up and drop-off points of taxi passengers for applying k-means point clustering with the aim to detect the busiest places in Shanghai. Others studies use density-based point clustering methods like DBSCAN (Tang et al. 2015) or OPTICS (Rinzivillo et al. 2008) for detecting vehicle movement patterns. In case of dense point distributions it is possible to use kernel density estimation (KDE). Krisp et al. (2011) extend this idea and use adaptive and directed KDE for the visual traffic analysis, which helps to detect movement trends within dynamic point data.

2.2 Exploratory Data Analysis (EDA)

In general it is to say that some of the mentioned visual analysis examples make use of procedure flow, which is described by the term exploratory data analysis (EDA). This term was summarized by Keim et al. (2004) as a sequence of the three steps "Overview", "Zoom and Filter" and "Details on demand". The overview represents the inspected data in a summarized view, which will be called in our case the global view. By using the "Zoom and Filter" functions, which are the used data analysis methods, it is possible to detect movement patterns. After some patterns are detected "Details on demand" refers to the inspection of certain details in the data, which is dependent on the field of interest for analysis. This inspection in a detailed view on the data is called in our case local view. The aim of "Details on demand" or the local view is to propose a hypothesis in the end (Keim et al. 2004).

2.3 Geovisual Analytics of Movement Data

Exploratory Data Analysis and the mentioned examples for visual analysis of FCD will be the theoretical base for creating selection areas for our test data set.

Additionally it is to mention that the term selection area refers to area wise selection of spatio-temporal data sets for further visualization.

In a similar way selection areas are termed differently in numerous Visual Analytics approaches. They are named time lenses in Tominski et al. (2012), trip views in Liu et al. (2011) and spatial traffic views in Guo et al. (2011). The general connection of these examples to our term selection area is the use of a dynamic spatial query (Tominski et al. 2012). By using dynamic spatial queries data partitions of a vehicle movement data set are selected and inspected further by other linked views on the data. With the use of a time lense it is possible to inspect certain partitions of the data set in time steps. Tominski et al. (2012) show that numerous trajectories of multiple moving objects can be inspected visually for long time periods. In Guo et al. (2011) the TripVista application has an interlinkage between global and local views with selection possibilities for certain trajectories. Examples for windows or views despite the selection area are histograms, parallel coordinate plots, scatterplots and heat maps.

Keeping this in mind, an additional aim for using selection areas for FCD is interactive linking to other reasonable views on the data.

3 Test Traffic Data for the Vehicle Transportation Network of Shanghai

For testing our approach we use 3 different data sources. Besides FCD from taxis, we use obtained traffic states from the Google traffic layer. The base for comparing the different traffic states is an extracted partition of the road network in Shanghai, coming from the OpenStreetMap (OSM) project. In the following we will discuss the properties of our used test data sets.

3.1 Floating Taxi Data from Shanghai

The inspected FCD set is the result of a survey on a taxi fleet in Shanghai with an average of 7120 frequently observed vehicles. This number represents the average for each hour. In total there are around 10,000 different taxi identifications. We can detect this important pattern by simple inspection of vehicle identifications for certain hours of the day. Depending on the time of day some of the taxi drivers turn their tracking device off and some new appearing turn it on. The data structure of the inspected data set is shown in the following Table 1.

Most of the inspected original 10 attributes are not used in our study. A pre-processing step was provided by means of selecting certain attributes, where only the car ID, longitude, latitude, time and instantaneous velocity are kept for

Table 1 Data structure of the inspected taxi FCD set of Shanghai

Fieldname	Details
Car ID	The unique ID of the car, in 5 digits
Longitude	In degree [°]; accurate to the 6th decimal place
Latitude	In degree [°]; accurate to the 6th decimal place
Instantaneous velocity (km/h)	accurate to 0.1
Record date	In form of yyyy-mm-dd
Record time	In form of hh:mm:ss

further analysis. We inspect in our study FCD partitions of 15 selected days between the 1st of February and 1st of March 2007.

By previous inspection of the data it should be noted that the sampling interval in time is differing. This means the jumps in time between consecutive points of one and the same vehicle are not constant. These time jumps vary between 1 s and 30 s and have an average of around 12 s for each inspected hour of the data set.

3.2 Street Network from OpenStreetMap (OSM)

We extracted the street network of the entire city extent of Shanghai from the OpenStreetMap (OSM) project. Based on Stanica et al. (2013) the digitized road network of this source has one of the best quality of accessible street geodata. Nevertheless, we have to inspect the reasonability of this information, mainly in road types, driving directions and restrictions. All these attributes of the OSM road segments are crucial for achieving reasonable routing results.

In terms of connectivity of different road segments, we tested Shanghai's road network with computed test routes. The results indicate a more or less realistic connection between the roads, with some small mistakes in driving directions and in complicated crossing. For comparing the reliability of the computed OSM network routes the routing service of Google Maps was used.

In a case study the quality of the OSM road network was already evaluated for its suitability for vehicle routing (Graser et al. 2014). For the city of Shanghai, we extract all available road networks that are accessible by car. For testing reasons we will mainly focus on the highway network in Shanghai.

3.3 Web-crawled Traffic States from Google Traffic Layer

The Google traffic layer is an additional layer, which can be optionally switched on, in the Google Maps service. Within this service it is possible to inspect segments of the road networks by their recent traffic situation. There are 4 quality values ranging from slow moving traffic to fast moving. As an extension users may use the Google route service within the Google Maps GUI and predict the certain travel times for the routes between selected start and end points. The simplest routing strategy is shortest path based on Dijkstra (1959), which is not always the best solution for users, as for example traffic congestion can influence the travel time. Therefore, a traffic-aware route is often proposed as the best route, respectively the fastest route within specific time windows. Traffic-aware routing means that the information that is partially visualized in the Google traffic layer is respected. Besides the information that is usually transmitted via TMC, such as accidents or closed roads, the travel time for each segment is calculated based on tracked data. The traffic layer itself is updated in irregular time intervals always dependent on available input data.

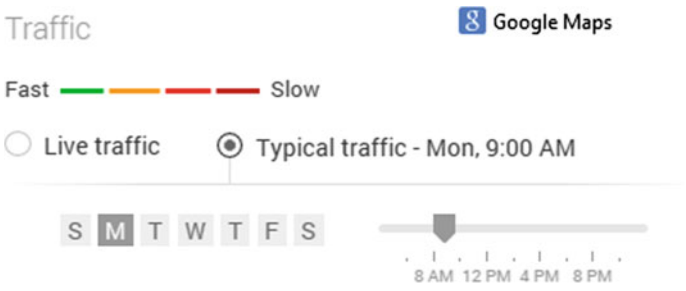


Fig. 1 Interactive legend of the Google traffic layer in Google Maps for the visualization of different traffic states

The data source for the Google traffic layer or any kind of metadata is not given by the service provider. Nevertheless direct information from Google state that the data source is a merged solution from several sources.[1] This includes data from government departments of transportation, private data providers and mobile phone users. The last mentioned provide anonymous speed information via Google's traffic crowdsourcing feature. In reality users of Google of Google Maps for mobile phones that enable the "My Location" function are helping Google to provide traffic information more accurately.[2] This should be seen critically, since no information is available on positioning quality of the data source, quality of the MM results and reliability of the provided traffic information.

In our approach we inspect the live and the typical traffic information from Google's traffic layer, which is an available option within an interactive legend in Google Maps. Figure 1 pictures one example of such an interactive legend. We use a web-crawling technique that was already tested for the traffic layer in Bing Maps by Tostes et al. (2013) for extracting time dependent traffic states (4 different) for each weekday in Shanghai. The weekdays are important for providing comparable traffic situations. Unfortunately all the Google traffic information is from 2015 and is difficult to compare with taxi FCD from 2007. This is problematic due to the fact that numerous changes in the urban transportation infrastructure appeared in the last 7 years in Shanghai.

We use historical traffic information of the road segments in Shanghai for comparing it with the taxi FCD-based traffic situation weekday-wise. This has the idea that same weekdays appear more similar than differing ones. After the first inspections of the traffic layer from Google Maps for Shanghai, we detected that there is only typical or live traffic information available for highways. Resulting from this we will only inspect and compare input data associated to the highway segments in Shanghai.

[1]http://geeknizer.com/how-google-maps-traffic-works/.

[2]https://googleblog.blogspot.in/2009/08/bright-side-of-sitting-in-traffic.html.

4 Introducing Selection Areas for Visual Analysis of Taxi FCD

The first part of our approach consists of introducing selection areas for the inspection of certain FCD partitions in time and space. The inspection is mainly a visualization of averaged traffic parameters such as velocity and vehicle density. Since each selection area has a timestamp, we connect several defined areas by spatiotemporal and semantic relations. The spatiotemporal relations, mainly on the connected road segments, are then enriched with traffic information such as congestion level or travel times for computed traffic-aware routes based on various data sources. The aim of the approach is to provide exploratory data analysis on taxi FCD based on selected areas of interest with recent and historical travel time information.

4.1 User-defined Spatial Area Selection

The selection areas appear in two forms: as a selection circle without association with a given road network and as road selectors with dependency on certain road segments. Figure 2 shows the usual appearance of the selection areas used in our approach.

The introduction of these elements is user-dependent and can indicate a personal context like "at home" or "place of employment". In our approach the shape of selection areas has the two options as pictured in Fig. 2, mainly dependent on the selected point or line of interest. This means that selection areas are similar to spatial point and line buffers. The size of these buffers is selected by the user and gives an indication how much area is included into specific interests. Selection areas represent areas of interest and define the localization of subsequent calculation and visualization of spatially intersecting taxi FCD.

Fig. 2 Selection areas as **a** selection circle and as **b** road selector

4.2 Different Views on Taxi FCD: Global View and Local View

With our visual representation we want to give insight into parts of the inspected FCD sets. Therefore, we focus on associating selection areas with the deduced FCD information on average velocity and vehicle density.

For reasons of achieving overview there is a differentiation between a global view and a local view on inspected taxi FCD sets. Figure 3 pictures this differentiation by the workflow of data selection and visual inspection. Additionally, Fig. 3 shows how we link supplementary views, which may be charts on temporal distributions, on the data spatially included within selection areas.

As can be seen in Fig. 3 two differently calculated attributes are presented for the local view, which are always based on the selected area of interest. In our case, we make use of two types of simple calculations for two different attributes. For calculating the vehicle density we simply use the sum of counted vehicle positions within the selection areas. The average velocity is calculated by the average of n points within the selection area.

These are the 2 basic attributes, which are fixed in our approach, since this method relies on the existence of at least one small FCD set.

In Fig. 3 there is a global and a local view on the data. The global view has the intention to give some orientation on the investigation area and to show interesting appearance for placing a selection area. The base or base map of a global view for example can imply a usual traffic map with 3 qualitative values describing congestion (low, medium, high) or classified by ranges of average velocity. Other examples could be thematic and topographic maps, satellite imagery, the raw FCD points or simply a digitized road network. Here it should be mentioned that this view on the data may only be a visualized layer and does not represent the original form of the data. Additionally we can show the original positioning points by dots colored by velocity ranges (in this case: instantaneous velocity).

With the visual detection of interesting places (points of interest) in a static traffic map, we can use a selection area for further inspection. There are many different options for the appearance of global views. Therefore Fig. 4 pictures two possible displays for a global view on FCD.

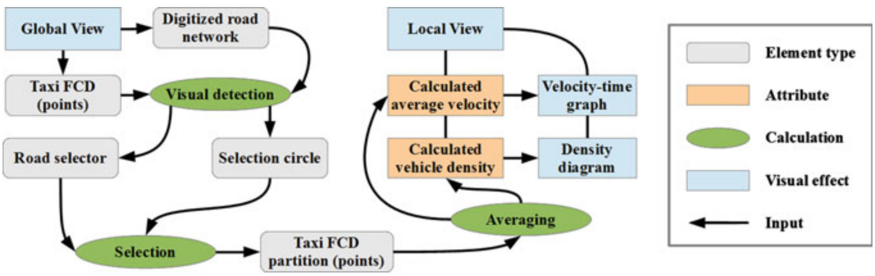


Fig. 3 Workflow of FCD inspection by selection areas and linked charts

Fig. 4 Examples for selection areas in a global view with **a** FCD records as points and **b** classified road partitions

Figure 4a shows the original dot representation of the used FCD and Fig. 4b a road network representation colorized by average velocity classes. The displays in Fig. 4 are showing the results for a test taxi FCD partition of 10 min in the global view. A selection of a bigger taxi FCD partition can be realized by queries on the spatial position of the introduced selection area. Afterwards it is possible to average and count FCD records for other time ranges.

4.3 The Method for Defining Individual Time-dependent Selection Area Sequences and Its Visualization

The proposed method for providing multiple successive selections structures as the following:

1. User defines polygons based on one selected point (POI) or on one line (selected road segment).
2. Sequence of user is recorded (ID, pol_ID, time, type, name).
3. Defined Polygons are enriched with average information on traffic states and travel times (with the option for different travel modes).

The deduced FCD information of the enriched polygons is visualized within a local view.

One example for a local view with linked displays is pictured in Fig. 5 with a road selector as selection area.

In Fig. 5 we test the selection area method on the Shanghai FCD set. As already mentioned, we follow with our approach the idea of a “Focus Map” (Zipf 2002), where the important areas of the map are displayed more dominantly then the others. In this example we investigate one FCD partition of one hour between 8 a. m. and 9 a.m. on a Wednesday in the center of Shanghai.

The visualized information about the average velocity and the vehicle density can be estimated with 10 min steps. Following the aggregation of movement is based on 10 min in time and the size of the selection area in space.

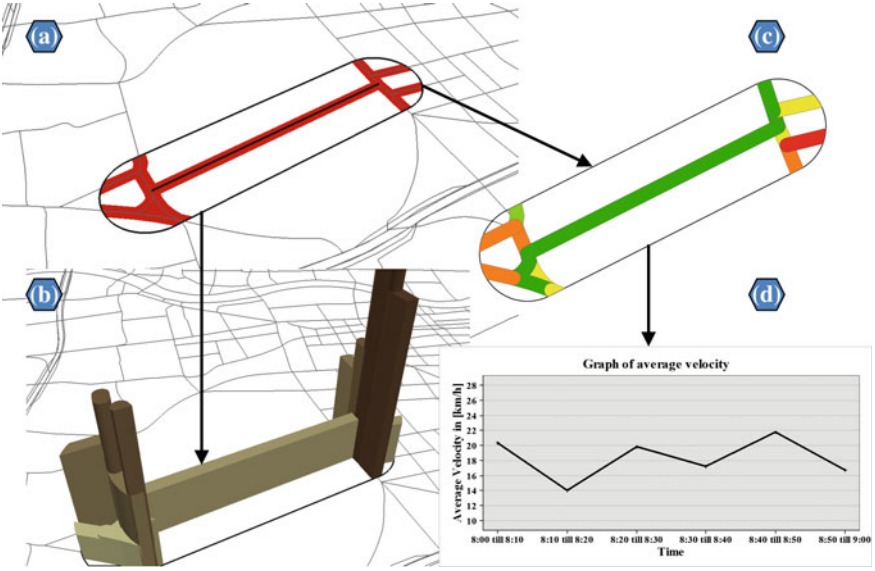


Fig. 5 Possible display of a global and local view on FCD with **a** road selector on road network; **b** extrusion of road segments based on taxi density; **c** coloration based on average velocity ranges and **d** associated graph of average velocity

The main selection steps consist of querying FCD by certain time windows and coordinate restrictions in space. This is included in the global view. The counting of the vehicle positions (for the vehicle density) and calculating average velocity values are the main calculation steps and part of the local view. Each of these steps is based on selected areas in the form of selection circles and road selectors.

Figure 5d shows the chart of the taxi average velocity by its variation in time, which is linked with the same selection area showing consistent time steps. Each time step includes a data partition based on time stamp selection. Similar to the mentioned data sets it is possible to link other spatio-temporal information and data sources as for example weather, social events or other changes in the environment. Interesting for personalized selections within location-based services (LBSs) are location-aware news feeds or user ratings of spatial items as additional data sources (Mokbel et al. 2011).

4.4 *Defining Semantic, Topological and Temporal Relations of Multiple Successive Selections*

Multiple successive selections are connected by semantic relations (between POIs in different selections), topological relations (same road network) and temporal relations (working hours).

Each selection area is product of personal interest ("personalized traffic information"). Depending on the type of averaged data within selection areas we can define semantic relations between certain selection areas. These semantic relations often result from temporal relations. One example might be the time when a user is leaving home and going to work. Resulting from the topological relations we know the usual travel time for the shortest path between the points in two different selection areas. Depending on the deduced traffic information from FCD we can estimate if a variation in travel time is possible or not.

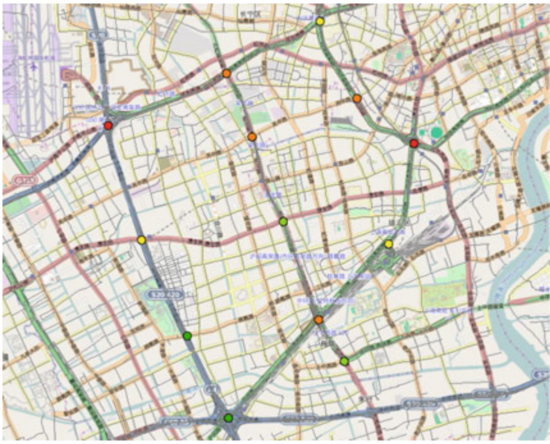
5 Results

By the use of combining a global and a local view on taxi FCD by selection areas it is possible to get insight into average velocity and absolute vehicle density values of the selected areas. The selection itself appears interactively within the global view, where raw FCD points, traffic base maps or simply the digitized road network are used for orientation.

In our case study we selected 13 prominent highway crossings in Shanghai for testing our approach on traffic congestion events, traffic states (slow or fast moving traffic) and travel times. The latter was calculated from historical taxi trajectories and recorded from the routing function of Google Maps. Traffic congestion was calculated from taxi FCD velocity records and the density of taxi positions. The traffic states were computed by average velocities and travel times from taxi FCD and recorded from the Google traffic layer.

Even if the inspected taxi FCD sets are from the year 2007, there are similar weekday-dependent traffic patterns comparing to the live and typical traffic classifications from Google traffic layer. The peak hours on weekdays, in particular from 8 to 9 a.m. and from 5 to 7 p.m. (Sun et al. 2009), that are characterized by heavy traffic congestion events are detectable in both data sets. Additionally there are only relatively small variations in comparing calculated traffic-aware travel times. For the latter case we used only the Google live traffic routing function respectively between the selected crossings in Fig. 6. We compared the traffic dependent shortest path of the life traffic of selected weekdays for the same times of the day.

Another interesting result is the traffic-aware routing for a given visiting order of the timestamped selection areas (see Fig. 6 on the right). Following the ascending order of selection area definition by a route delivers other travel times as the descending order of selection areas. This has the reason that the segments in between are associated with frequently changing traffic state information or travel time. Therefore differing starting and end points deliver different travel times for the same and in different directions driven routes. This factum shows the dynamically changing traffic situation on highways on workdays in Shanghai.



ID	Sel_area	Sel_lon	Sel_lat	Sel_ID
388	Wai Huan Hu Qing Ping Li Jiao Qiao	121.358094	31.183053	301
389	849 Hong Xu Lu	121.388660	31.193598	302
390	827 Zhong Shan Xi Lu	121.412212	31.205220	303
391	1885 Cao Bao Lu	121.367646	31.158105	304
392	1619 Gu Dai Lu	121.379107	31.136764	305
393	Xin Zhuang Li Jiao	121.389511	31.118375	306
394	708 Wu Zhong Lu	121.394401	31.180858	307
395	2 Wu Zhong Lu	121.421279	31.189230	308
396	71 Cao Xi Lu	121.436061	31.179283	309
397	688 Cao Bao Lu	121.403223	31.162119	310
398	Hu Min Lu Fu Lu	121.419189	31.148112	311
399	76 Liu Zhou Lu	121.429660	31.156339	312
400	710 Shang Zhong Xi Lu	121.418159	31.131703	313

Fig. 6 Selected crossings classified by quality of traffic congestion (5 classes) based on taxi FCD from Shanghai

6 Conclusion

The idea of introducing area polygons and circles as selection areas is helpful, as it was already used for visual analysis on different data sets (Tominski et al. 2012; Guo et al. 2011; Ferreira et al. 2013; Wang et al. 2013). Nevertheless, the extension to a user-driven selection process with establishing relations between individual selection areas is relatively new.

The proposed specified visualization methods are only examples of designing user-friendly visual analysis methods. One main issue of this testing may be the evaluation of the practical use of selection areas for getting insight into microscopic traffic patterns.

The 13 selected crossings in Fig. 6 are good indicators for detecting traffic variations on the highway network in Shanghai. We can connect averaged taxi FCD from various time windows with these selected areas. Especially in comparison with the Google traffic layer solution, colorized selection areas show a suitable and informative extension for commonly known interactive traffic maps.

There are several questions that result from the first tests of our analysis method. One of them implies the questioned usefulness of extending traffic maps with more interactivity. We cannot estimate the usefulness of our tool for daily commuters. One clear benefit of our user-based area selection method is the extension of the visual analysis process for FCD. Still, we need to detect the ease of use for a potential user by providing evaluations with individual selection areas. Defined preferences of selection areas are user-specific. Therefore, we need to respect the frequency of selection area inspections, which will be an addition to their order (temporal) and definition (local knowledge). This might be an initial point for the conception of a GUI for the visual analysis of FCD.

7 Outlook

All the presented visualization possibilities in Fig. 5 appear often trivial, since only highly generalized and averaged traffic information is visualized.

Future work may include the extension of 2D road networks into 3D representations of the entire vehicle transportation network of dense populated cities with the aim of expanding the possibilities for modelling vehicle traffic. Showing two variables of one feature is a good base for starting geovisual analyses. The linkage with the global view and estimated travel times and traffic states between user-defined areas can supplement this analysis, since we can detect dependencies in the typical or "usual" dynamics of vehicle traffic.

Another idea, which may include the use of the third dimension, is about representing different traffic situations on one and the same crossing visually. This may include the use of extrusion in 3D views with the vehicle density as the extruded value on the z-axis and additional coloration by average velocity for selected segments of the transportation infrastructure. Figure 7a pictures one imagined example for one hour of taxi FCD in one selected crossing in Shanghai. There are six partitions that appear different depending on the ratio of different vehicle driving directions. It should be mentioned that this appearance is highly dependent on the working times of traffic lights. This brings us to the favored linking of this kind of information, which is hard to detect but relatively simple to simulate.

In all of these cases the differentiation between times of the day, the certain weekday, and typical and averaged time windows is crucial for comparing results.

Another idea for further analysis and visualization of selection areas is the inclusion of spatial point interpolation and density estimation methods. This might be useful for cases of multiple or overlapping POI in the same selected area. One possibility is to introduce weightings for different POI as pictured in Fig. 7b. In this figure the points are weighted based on the Euclidian distances to one selected POI. Insight into weightings of POIs is important for applying and interpreting spatial interpolation methods and their results. In case of using kernel density estimation

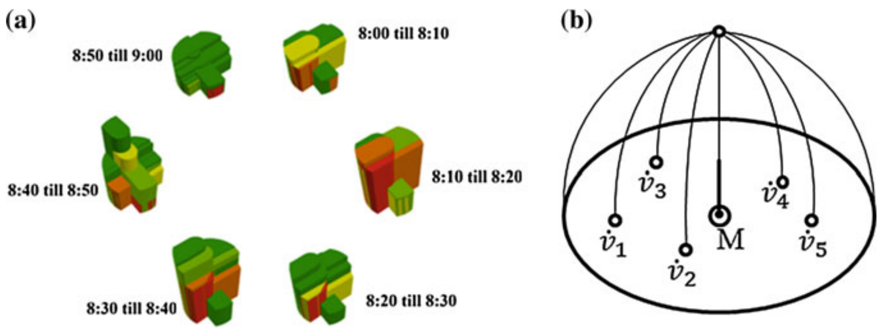


Fig. 7 Possible visual representations of **a** different traffic situations on the same crossing and **b** weighting of different POIs

(KDE) the views as in Fig. 7b are useful for selecting the kernel bandwidth. Another example for creating surface information out of spatial points is presented in Keler and Krisp (2015) by the inclusion of spatially interpolated PM2.5 values based on measurement from static monitoring stations. The connection between vehicle traffic and particulate matter is an important topic nowadays, especially since exceeded PM concentrations are a big problem in today's Chinese cities like Beijing and Jinan.

Acknowledgements The described taxi Floating Car Data set of Shanghai ('SUVnet-Trace Data'[3]) was obtained from the Wireless and Sensor networks Lab (WnSN) at Shanghai Jiao Tong University. We would like to thank the Laboratory for Wireless and Sensor Networks at Shanghai Jiao Tong University, especially Prof. Min-You Wu and Jia Peng, for providing access to this data.

References

Andrienko N, Andrienko G (2007) Designing visual analytics methods for massive collections of movement data. Cartographica 42(2):117–138

Andrienko N, Andrienko G (2013) Visual analytics of movement: an overview of methods, tools, and procedures. Inf Vis 12(1):3–24

Cohn N, Bischoff H (2012) Floating car data for transportation planning. Explorative study to technique and applications and sample properties of GPS data. NATMEC 2012 Dallas; Presentation slides

Dijkstra EW (1959) A note on two problems in connexion with graphs. Numer Math 1(1):267–271

Dirks KN, Johns MD, Hay JE, Sturman AP (2003) A semi-empirical model for predicting the effect of changes in traffic flow patterns on carbon monoxide concentrations. Atmos Environ 37 (19):2719–2724

Ferreira N, Poco J, Vo HT, Freire J, Silva CT (2013) Visual exploration of big spatio-temporal urban data: a study of New York City Taxi Trips. IEEE Trans Visual Comput Graphics 19 (12):2149–2158

Freksa C (1999) Spatial aspects of task-specific wayfinding maps—a representation-theoretic perspective. In: Gero JS, Tversky B (eds) Visual and Spatial Reasoning in Design, p 15–32

Goldsberry K (2005) Limitations and potentials of real-time traffic visualization for wayfinding. Proceedings of the 22nd ICA/ACI International Cartographic Conference; A Coruna, Spain, 9–16 July 2005

Goldsberry K (2008) GeoVisualization of automobile congestion. Proceedings of the AGILE 2008 Conference: GeoVisualization of Dynamics, Movement and Change; Girona, Spain, 5 May 2008

Graser A, Straub M, Dragaschnig M (2014) Is OSM good enough for vehicle routing? A study comparing street networks in Vienna. In: Gartner G, Huang H (eds) Progress in Location-Based Services, Lecture Notes in Geoinformation and Cartography, Springer, Berlin, p 3–18

Guo H, Wang Z, Yu B, Zhao H, Yuan X (2011) TripVista: triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. In: Di Battista G, Fekete J-D, Qu H (eds) Proceedings of the Pacific Visualization Symposium 2011 (PacificVis 2011), IEEE, Session 5: Space and Time, p 163–170

[3]http://wirelesslab.sjtu.edu.cn/taxi_trace_data.html.

Keim DA, Panse C, Sips M, North SC (2004) Visual data mining in large geospatial point sets. IEEE Comput Graphics Appl 24(5):36–44

Keler A, Krisp JM (2015) Visual analysis of floating taxi data based on selection areas. In: Gartner G, Huang H (eds) Proceedings of the 1st ICA European Symposium on Cartography (EuroCarto 2015), p 58–59

Krisp JM, Peters S, Mustafa M (2011) Application of an adaptive and directed kernel density estimation (AD-KDE) for the visual analysis of traffic data. Online Proceedings GeoViz2011, Hamburg, Germany, 9–11 Mar 2011

Krisp JM, Polous K, Peters S, Fan H, Meng L (2012) Getting in and out of a taxi: spatio-temporal hotspot analysis for floating taxi data in Shanghai. Conceptual paper—6th International Symposium "Networks for Mobility", 27–28 Sept 2012

Leduc G (2008) Road traffic data: collection methods and applications. Working Papers on Energy, Transport and Climate Change N.1, JRC Technical Notes, European Commission, Joint Research Centre (JRC), Institute for Prospective Technological Studies (IPTS), EUR Number: Technical Note: JRC 47967

Liu X, Ban Y (2013) Uncovering spatio-temporal cluster patterns using massive floating car data. ISPRS Int J Geo-Inf 2:371–384

Liu C, Meng X, Fan Y (2008) Determination of routing velocity with GPS floating car data and WebGIS-based instantaneous traffic information dissemination. J Navig 61(02):337–353

Liu H, Gao Y, Lu L, Liu S, Qu H, Ni LM (2011) Visual analysis of route diversity. In: Miksch S, Ward M (eds) Proceedings of IEEE Conference on Visual Analytics Science and Technology 2011 (VAST 2011), Session 5: Space and Time, p 171–180

Liu C, Li N, Huang M, Wu H (2012) City routing velocity estimation model under the environment of lack of floating car data. J Geo Inf Syst 4(1):55–61

MacEachren AM (1995) How Maps Work. The Guilford Press, New York

Mokbel MF, Bao J, Eldawy A, Levandoski JJ, Sarwat M (2011) Personalization, socialization, and recommendations in location-based services 2.0. In: Proceedings of the International Workshop on Personalized Access, Profile Management and Context Awareness in Databases, (PersDB 2011), co-located with VLDB 2011, Seattle, WA, Sept 2011

Moosavi V, Hovestadt L (2013) Modeling urban traffic dynamics in coexistence with urban data streams. In: UrbComp' 13: Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing; Article no.: 10

Quddus MA, Ochieng WY, Noland RB (2007) Current map-matching algorithms for transport applications: state-of-the art and future research directions. Transp Res Part C Emerg Technol 15(5):312–328

Rinzivillo S, Pedreschi D, Nanni M, Giannotti F, Andrienko N, Andrienko G (2008) Visually driven analysis of movement data by progressive clustering. Inf Vis 7(3–4):225–239

Sohr A, Brockfeld E, Krieg S (2010) Quality of floating car data. In: Conference Proceedings, paper nr 02392, www.digitalpapers.org. 12th World Conference on Transport Research (WCTR), Lisbon, Portugal, 11–15 July 2010

Stanica R, Fiore M, Malandrino F (2013) Offloading floating car data. In: Proceedings of the 2013 IEEE 14th International Symposium and Workshops on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), p 1–9

Sun J, Li X (2012) A pyramid-based approach to visual exploration of a large volume of vehicle trajectory data. Frontiers Earth Sci 6(4):345–353

Sun J, Wen H, Gao Y, Hu Z (2009) Metropolitan congestion performance measures based on mass floating car data. In: Yu L, Lai KK, Mishra SK (eds) Proceedings of the Second International Joint Conference on Computational Sciences and Optimization (CSO 2009), Computational Transportation Science, p 109–113

Tang J, Liu F, Wang Y, Wang H (2015) Uncovering urban human mobility from large scale taxi GPS data. Physica A 438:140–153

Tominski C, Schumann H, Andrienko G, Andrienko N (2012) Stacking-based visualization of trajectory attribute data. IEEE Transactions on Visualization and Computer Graphics (Proceedings IEEE Information Visualization 2012) 18(12)

Tostes AIJ, de L. P. Duarte-Figueiredo F, Assunção R, Salles J, Loureiro AAF (2013) From data to knowledge: city-wide traffic flows analysis and prediction using bing maps. In: Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing (UrbComp '13), 12

Wang Z, Lu M, Yuan X, Zhang J, van de Wetering H (2013) Visual traffic jam analysis based on trajectory data. IEEE Trans Visual Comput Graphics 19(12):2159–2168

Zhao Y, Qin Q, Li J, Xie C, Chen R (2012) Highway map matching algorithm based on floating car data. IGARSS 2012—IEEE International Geoscience and Remote Sensing Symposium, THP-P: Image Information Extraction: Detection of Man-made Features. p 5982–5985

Zipf A (2002) User-adaptive maps for location-based services (LBS) for tourism. In: Woeber K, Frew A, Hitz M (eds) 9th International Conference for Information and Communication Technologies in Tourism (ENTER 2002)