young transportation engineers conference 2020

České vysoké učení technické v Praze
Fakulta dopravní
19. listopadu 2020
Praha, Česká republika

The utilization of floating car data for traffic flow analysis

Zuzana Purkrábková, Pavel Hrubeš

Abstrakt: *Digitalizace dopravního prostředí a využití nových datových zdrojů je v dnešní době velmi diskutované téma. Systémy FCD (Floating Car Data), tzv. plovoucích vozidel, umožňují získávat data o dopravních parametrech přímo ze silniční sítě bez nutnosti instalace stacionárních senzorů. Článek se zabývá využitím* tohoto *nového datového zdroje v České republice, financovaného z veřejných zdrojů a zpřístupněného Ředitelstvím silnic a dálnic. Tato data s využitím geografických informačních programů lze zobrazit v modelu silniční sítě a dále je analyzovat.*

Klíčová slova: *plovoucí vozidlo, geografické informační systémy, data z plovoucích vozidel, analýza dat*

Abstract: *Digitization of the transport environment and the use of new data sources is a much-discussed topic today. FCD (Floating Car Data) systems, so-called floating vehicles, make it possible to obtain data of traffic parameters directly from the road network without the need to install stationary sensors. The article deals with the use of this new data source in the Czech Republic, financed from public sources, and made available by the Directorate of Roads and Motorways. This data can be displayed in the road network model using geographic information programs and further analyzed.*

Keywords: *floating car, geographic information systems, floating car data, data analysis*

1. Introduction

The article deals with the use of a new data source, financed from public sources. In the spring of 2019, the Directorate of Roads and Motorways launched a trial operation of the interface, where it is possible to request access to the free download of traffic data. These are traffic data from floating vehicles, ie time-stamped geolocation and speed data with a time stamp. Data is collected directly by vehicles in the transport network.

The first step was to obtain this data and their subsequent analysis. The author gradually obtained data from 2020, which he will subsequently process. It is obvious to acquire basic knowledge of data format and their meaning.

In the following part, the article deals with the definition of the model network. As this is data only for some vehicles in the road network, it is necessary to analyze their distribution within the road network, the penetration of this data. Due to the current situation of spread corona virus in the society, traffic data are highly affected during the year.

In the last part, the article focuses on the subsequent steps that will lead to the possible design of the model and thus support the further development of the project.

2. State-of-the-art

Applications using data from floating vehicles are not a phenomenon of recent years. Already in the last century, efforts have been made to use data on a similar principle to easily inform drivers.

The use of data from floating vehicles for traffic management at the city level is a relatively common practice today. Systems based on this data can be found, for example, in New York, Beijing, or Vienna [1] a [2].

Another example is a study [3], that uses data from floating vehicles for more precise road weather forecasts. Paper deals with increasing the spatial and temporal resolution of meteorological observations by integrating floating car data.

A closer look is made in a paper [4], where data from taxi vehicles were tested. The research examined the choice of route for taxi drivers in Berlin, and it should be noted that this group is relatively specific in its behavior. Research focuses on driver behavior rather than on specific data processing from floating vehicles.

Article [5] also deals with the behavior of taxi drivers in the city. Analysis of data from floating vehicles shows that driving time, travel distance, and road preferences are the main parameters for route selection. A similar topic is addressed in [6], which deals with vehicles in the Rome area.

The team at CTU in Prague Faculty of transportation sciences in cooperation with the Directorate of Roads and Motorways has long been dealing with the possibilities provided by data from floating vehicles within the road network in the Czech Republic. The first report of this cooperation was created in 2011 [7]. Cooperation in this field has continued and the research team is currently working on control and expert activities in the implementation and operation of the FCD. Their latest report [8] monitors data for January 2020.

Data from toll gates or cameras were already available before the FCD source. This data is generated on roads where the necessary infrastructure is installed. On the contrary, data from floating vehicles do not need to install any additional support infrastructure. Thanks to this, FCD can cover a larger range of the network, as no fixed detectors are needed. In addition, faster commissioning and lower maintenance are required.

3. Data source

Data from floating vehicles are captured by Lagrangian description of motion, in contrast to stationary detectors, which are captured by Euler's description. The vehicle behaves like a moving sensor in the network and its position is recorded using a GNNS receiver. Their behavior is compared to a fluid flow, where we do not follow every single drop, but only selected points. Thus, the model will never monitor all vehicles in the network, but only selected ones, based on which it will predict the state of traffic and create models. This principle is shown in the figure 1.

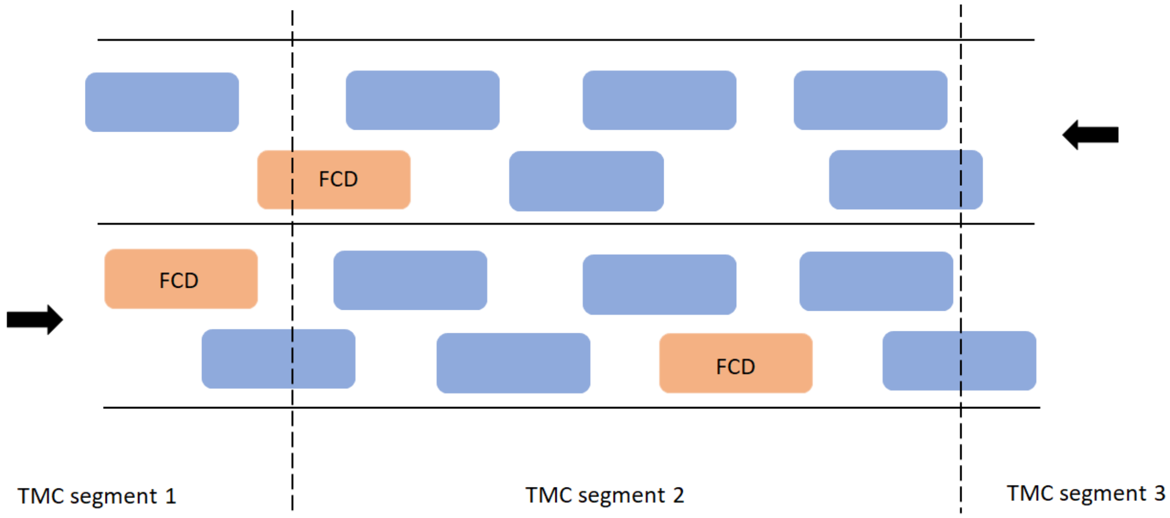

Fig. 1: *Scheme of floating car data principle*

Last year, the Roads and Motorways Directorate made available an interface for collecting data from floating vehicles. Floating vehicles are fleet vehicles equipped with a GNNS unit. The system then calculates the current speed, delay, and column length for each section above this data. In the initial mode of operation, only vehicles that transmit data are continuously being used. The fleet of floating vehicles sends data from start to engine shutdown. Therefore, it is not data from mobile devices located in vehicles. They only send information when the navigation application is switched on.
The composition of the fleet of floating vehicles in the Czech Republic is largely contained in passenger cars. This layout is logical so that trucks do not distort the data with their behavior (above all, it would artificially reduce the speed on the section).
For data collection, it is necessary to register on the interface and conclude a contract with the FCD interface operator. Based on this contract, it is then possible to collect data. The author of the article managed to overcome this procedure during the year and it is, therefore, possible to further describe the format of the sent data.

4. Data analysis

You can choose from two data formats from the Roads and Motorways Directorate distribution interface - DATEX and native XML. Data collection works based on sending data from the Roads and Motorways Directorate to the end customer according to predefined requirements. Based on the specified data subscription request, the author of the article now subscribes to data that contains the following attributes:

- defined TMC segment
- day (in format YYYY-MM-DD)
- time (in format HH:MM:SS +02:00)
- count (current number of FCD vehicles in the section that was further counted, represented by an integer)
- speed [km / h] (speed of traffic flow in current condition, represented by an integer)

- travel time [s] (travel time in current condition, represented by an integer)
- delay [s] (current delay calculated as the difference between the current travel time and the travel time in free traffic flow, represented by an integer)
- freeflow speed [km / h] (speed of traffic flow in free traffic flow, represented by an integer)
- freeflow time [s] (travel time in free traffic flow, represented by an integer)
- reliability (a measure of data reliability - characterizes the quality of data, decimal number in the range 0-1, where 1 means 100% reliable data. Quality is evaluated according to the number of data, age of input positions used and other parameters.)
- reaction time [s] (system reaction time calculated as the difference between the availability time of input data and the calculation time, represented by an integer)
- traffic level (traffic parameter marked according to convention level 1 - 5 in the Czech Republic)
- congestion (information on whether a column exists on the segment, represented by the binary notation 0/1)
- congestion from (estimation of the beginning of the congestion, defined by the string * congestion_from = "1534" *)
- congestion length (estimated congestion length in meters, defined by the string * congestion_length = "103" *)

An example of the transmitted data can be seen in figure 2. The example shows the structure of the data, a record of individual sections of the road network based on the identification of the so-called TMC segment, a selected section of road that has its unique ID. Using this identifier, the data can be assigned to the road network, which is described in the text below. Data is sent at minute intervals.


```
d="126" freeflow_time="80" reliability="0.750" reaction_time="30" traffic_level="3" congestion="0" /><tmc id="TS
mestamp="2020-07-01 01:11:01+02:00" count="2" speed="90" travel_time="108" delay="0" freeflow_speed="80" freeflc
liability="0.944" reaction_time="7" traffic_level="3" congestion="0" /><tmc id="TS01257T10378" timestamp="2020-0
:01+02:00" count="6" speed="84" travel_time="49" delay="15" freeflow_speed="128" freeflow_time="34" reliability=
eaction_time="4" traffic_level="3" congestion="0" /><tmc id="TS01271T23990" timestamp="2020-07-01 01:11:01+02:0
" speed="78" travel_time="87" delay="36" freeflow_speed="130" freeflow_time="51" reliability="0.500" reaction_ti
raffic_level="3" congestion="0" /><tmc id="TS01279T01280" timestamp="2020-07-01 01:11:01+02:00" count="2" speed=
vel_time="88" delay="33" freeflow_speed="130" freeflow_time="55" reliability="0.500" reaction_time="60" traffic_
" congestion="0" /><tmc id="TS01286T01285" timestamp="2020-07-01 01:11:01+02:00" count="3" speed="113" travel_ti
y="26" freeflow_speed="130" freeflow_time="53" reliability="0.750" reaction_time="30" traffic_level="3" congesti
><tmc id="TS01293T01294" timestamp="2020-07-01 01:07:01+02:00" count="1" speed="136" travel_time="40" delay="0"
```


Fig. 2: *Sample of subscribed data, including the highlighted ID of TMC segment*

Data from floating vehicles were then assigned to individual sections of the road network. This connection was made using the TMC segment attribute, which contains data from floating vehicles, and with the help of so-called location tables, which is a database of the road network with individual TMC sections. Thus, these two values must be the same in the data and the underlying layer. This achieves the decoding and correct assignment of each record to a segment within the spatial solution of the network. For this purpose, localization tables of *Central European Data Agency, a.s.* (CEDA) were used. The base layer used is shown in figure 3.

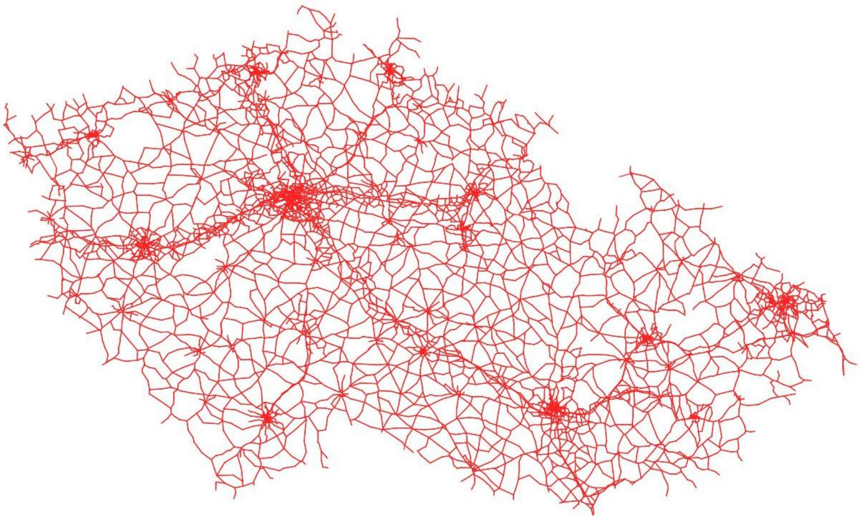

Fig. 3: *Location table visualized in QGIS*

5. Defined model network

Although the project is focused on the whole of the Czech Republic, it was necessary to define an infrastructure network on which the data can be further examined. There is a relatively large amount of data from floating vehicles throughout the country, but thanks to the parameters and origin of floating data, a large part of them are located on the main transport routes, motorways, high-speed roads, and 1st class roads.

Thanks to the cooperation with the Directorate of Roads and Motorways, sufficient penetration of data from floating vehicles in the Czech roads network was proved. These results can be seen in Figure .

As part of the calibration of the computational model for Directorate of Roads and Motorways [8] data from 2019 and 2020 were examined. The results suggest that data from floating vehicles can be considered sufficient for further investigation. From this result, the author's project is further reflected and considers the data source to be suitable for modeling alternative routes.

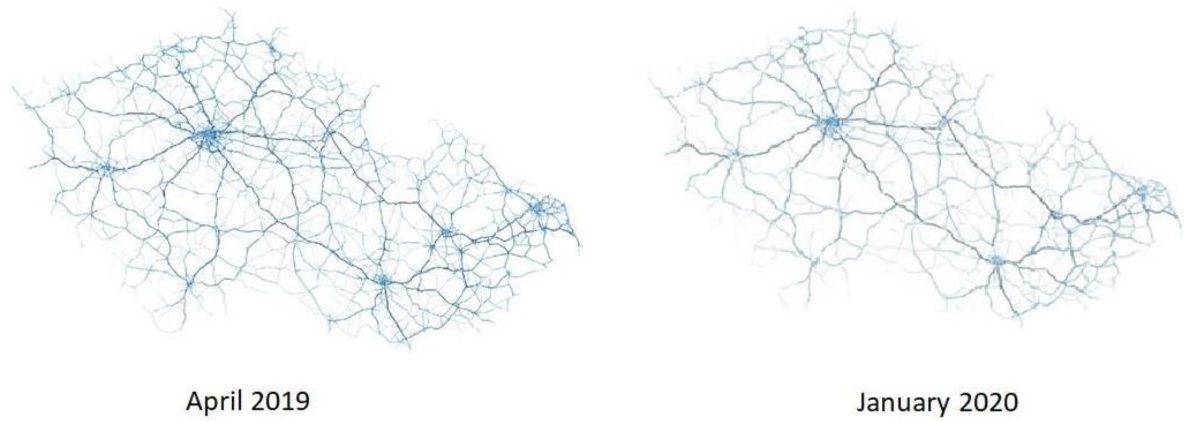

Fig. 4: *Comparison of penetration of the floating car data*

The subsequent goal of the long-term project is to model data from floating vehicles and research the creation of alternative routes in the road network. It is therefore necessary to obtain data both from the main traffic routes and from their surrounding roads. That is a road to which traffic can move in the event of the closure of major routes. The research will therefore continue to focus on the main roads, where congestion to the complete closure of the road will be examined. Subsequently, it will be examined whether it is possible to read from the data of floating vehicles the behavior of users in such situations, whether they choose the recommended detour routes, other alternative routes, or remain on the main road.

6. Subsequent steps for model creation

Based on a subjective assessment of the knowledge of drivers' behavior, it is assumed that there will appear a part of drivers who will choose other alternative routes, that is mean an alternative to a detour route. Nowadays, it is a common phenomenon to use navigation when traveling by car. Most navigations try to follow the main routes, but in the case of traffic excesses, the routes are converted to detours. Then a certain group of drivers often, based on their knowledge of the surrounding area or based on other navigation applications, choose other routes, often on lower-class roads or leading through municipalities, not through municipal bypasses.

These alternative routes may also slow down significantly at some point, for example, due to lower road capacity, as a result of increasing traffic. In the future, the author of the article questions whether these alternatives can be monitored and detected based on data from floating vehicles. Alternatively, whether these alternatives can be displayed to users on time, including the estimated travel time.

Based on traffic information, specific traffic situations will be selected, where alternative routes probably occurred when the main traffic route was closed. These situations will be monitored on the collected data. In the initial phase, it will be addressed how the closure of the main traffic route will affect the data over time. Furthermore, travel times on a closed main route and individual alternative routes will be examined and the results will be compared.

If the previous steps will be confirmed, it would be possible to examine the model, for example, using Kohonen maps, which is a type of neural network with learning without a teacher. This mathematical apparatus is suitable for the analysis of time series of traffic flows in the transport network and supports easy visualization. In the future, with the help of this method, it could be possible to analyze and predict traffic flows on multiple links simultaneously.

7. Conclusion

Sustainable transport is a big phenomenon today and this research can help identify unwanted crossings not only in municipalities and cities when the main route is closed. Thanks to data from floating vehicles, we have an up-to-date overview of the traffic situation on the road network in real-time, and it is, therefore, possible to further investigate whether the data is reflected in possible alternative routes and their load.

In further research, the authors would like to focus on the emergence of congestion and the closure of roads after specific situations and further examine how the behavior of the traffic flow changes, especially with regard to the emergence of alternative routes and possible alternatives to these detours.

Several studies have already been carried out on how to avoid the detour of paid routes, but research on the creation of alternative routes for large congestion until the closure of major roads is still a minimum.

Data from the Transport Portal of the Czech Republic were used for this project (data source: ŘSD ČR – JSDI – www.dopravniinfo.cz). This research was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS SGS20/082/OHK2/1T/16 - Research of creation of alternative routes and their impact on road infrastructure in the Czech Republic.

References

[1] PARK, H. Joon, Patrick H. KIM, Michael MARSICO and Naim RASHEED. *Data Mining Strategies for Real-time Control in New York City*. In: Procedia Computer Science [online]. Volume 32, 2014, 109-116, [cit. 2020-10-17]. ISSN 1877-0509. Available: https://doi.org/10.1016/j.procs.2014.05.404.

[2] ALTINTASI Oruc, TUYDES-YAMAN Hediye and Kagan TUNCAY, *Detection of urban traffic patterns from Floating Car Data (FCD)*. In: Transportation Research Procedia [online]. Volume 22, 2017, 382-391. ISSN 2352-1465. Available: https://doi.org/10.1016/j.trpro.2017.03.057.

[3] HELLWEG Meike, ACEVEDO-VALENCIA John-Walter, PASCHALIDI Zoi, NACHTIGALL Jens, KRATZSCH Thomas and Christoph STILLER, "*Using floating car data for more precise road weather forecasts*," 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium, 2020, pp. 1-3, doi: 10.1109/VTC2020-Spring48590.2020.9129401.

[4] MIKAT, Jürgen, THIESSENHUSEN, Kai-Uwe and Peter WAGNER, *Using FCD-data for measuring route-choice probabilities*, Conference Behavioural Responses to ITS, 2003. Available: https://www.researchgate.net/publication/224987608_Using_FCD-data_for_measuring_route-choice_probabilities.

[5] SUN, Daniel, ZHANG Lihui, CHEN Fangxi and Zhong REN PENG. *Urban travel behavior analyses and route prediction based on floating car data*. In: The International Journal of

Transportation Research. Volume 6. 2014, 118-125. Available: https://doi.org/10.1179/1942787514Y.0000000017

[6] FUSCO Gaetano, BRACCI Agnese, CALIGIURI Tommaso, COLOMBARONI Chaira and Natalia ISAENKO, *Experimental analyses and clustering of travel choice behaviours by floating car big data in a large urban area.* In: IET Intelligent Transport Systems. Vol. 12, No. 4, 270-278, 2018, DOI: 10.1049/iet-its.2018.0015.

[7] HRUBEŠ, Pavel, LANGR, Martin, DERBEK, Přemysl, SAIKO, Dušan and Martin VOLNÝ. *Zmapování služeb a dat v oblasti FCD (Floating Car Data) pro využití v rámci informačních systémů ŘSD.* Praha: ČVUT v Praze, Fakulta dopravní, Ústav řídicí techniky a telematiky, 2010. Research report no. LSS 384/10.

[8] HRUBEŠ, Pavel, PURKRÁBKOVÁ, Zuzana, KOVALJOV, Michal, KOVALJOVÁ, Kateřina, HAMPL Josef and Tomáš NOVOTNÝ. *Kontrolní a expertní činnost při zavádění a následném provozu FCD – kalibrace výpočetního modelu II*, 2020, research report č.: 97/2019a.