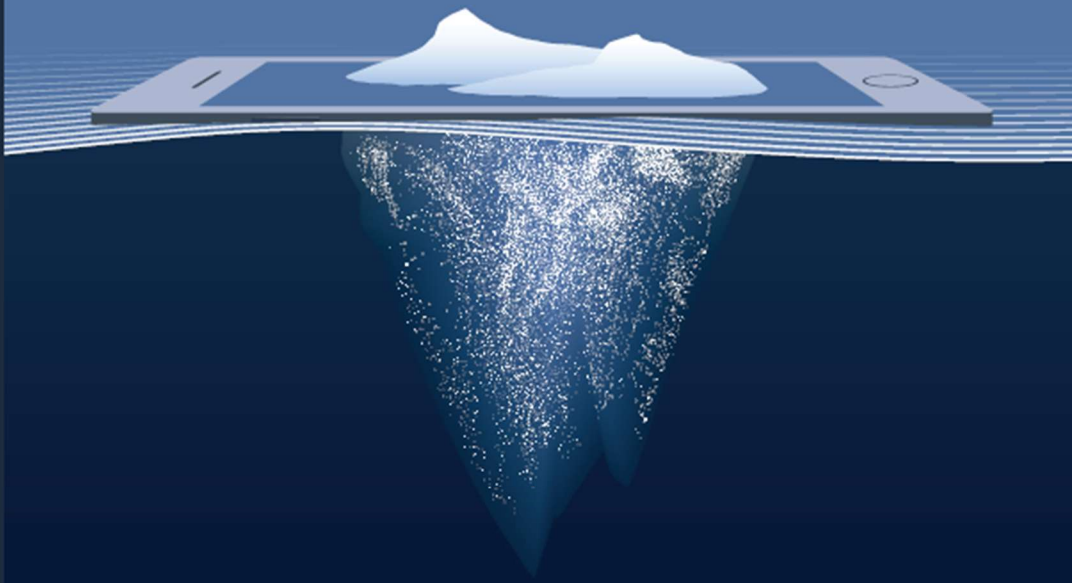
Population Mobility Data Retrieval

from Cellular Networks and its Use Case in Public Transport



MIROSLAV VOZŇÁK AND JIŘÍ HYLMAR

Vozňák & Hylmar: Population Mobility Data Retrieval from Cellular Networks and its Use Case in Public Transport

T A

Č R

Technology
Agency
of the Czech Republic

Results published in the brochure are the output of research project no. TB0500MD011, realised under the framework programme of public procurements in research, experimental development and innovations for the needs of public administration “BETA”, Technology Agency of the Czech Republic.

Authors:	© Miroslav Vozňák and Jiří Hylmar
Title:	Population Mobility Data Retrieval from Cellular Networks and its Use Case in Public Transport
Issued:	April 2017, 1st ed.
Pages:	81
Cover design:	Lukáš Orčík
Editor:	Robert Bešťák
Published by:	VŠB-Technical University of Ostrava
Impression:	200

ISBN 978-80-248-4039-0

Acknowledgement

The authors would like to thank Dr. Robert Bešťák, Dr. Petr Mazouch, Mr. Lukáš Orčík, Dr. Jan Platoš, Dr. Michal Prílepok, Dr. Jan Růžička and Mr. Marek Sedláček for their contribution in implementing individual components and tools in the Apache Hadoop environment and also Mr. David Blagodárný, Mr. Jindřich Borka, Mr. Jan Dušek, Mrs. Martina Jakubová, Mr. Lukáš Kapičák, Mr. Jan Karšňák, Mr. Milan Kříž, Mr. Jiří Nekovář, Ms. Jana Nowaková, Dr. Ondřej Přibyl, Mr. Jan Přikryl, Dr. Jan Rozhon, Prof. Václav Snášel, Mr. Jakub Staněk, Prof. Miroslav Svítek, Mr. Karel Šafr, Prof. Boris Šimák, Mr. Karel Uhlíř, Mr. Tomáš Vacek and Mr. Miloš Žák for their efforts and valuable help in contributing to the achieved result.

Foreword

We have invested additional effort to translate and edit the original methodology, “Specific system of passenger´s check-in and number of transported passengers” in order to provide an insight into the objectives and outputs of the project.

The abridged version named “Population Mobility Data Retrieval from Cellular Networks and its Use Case in Public Transport” was drafted to provide an insight into the methodology which enables to determine transport performance based on localisation and traffic data from mobile networks. The abridged version does not supersede the unabridged methodology, the Czech version, ISBN 978-80-248-4018-5.

We believe that the abridged methodology provides a solid, practical and updatable reference guide for the interpretation of the indicators gained during the processing of mobility-related information by adding mobile network data. It should be emphasised that the primary purpose of the methodology is to create the basis on which a further analysis and interpretation of information can build its ground, so it can be become meaningful to decision makers.

The TB0500MD011 project of the Technology Agency of the Czech Republic entitled “Specific system of passenger´s check-in and number of transported passengers” was realised between June and December 2016. To realise such a project, including the certification process, within 6 months, with a budget of CZK 1.9 million was only possible thanks to 3 years long experience and prior results of the implementation team in the same domain and the cooperation and involvement of the ordering party and the guarantor of the research need.

Editor´s Preface

We believe that where methodology users realise their respective projects with an equal dedication towards improved transport services as the team engaged in methodology preparation, it is probable that the enhanced awareness about the behaviour of transport network participants could improve the quality and the extent of the services provided by a public authorities and a private sector.

Consequently, more passengers could use public transport. This will positively affect the increase in the efficiency of means allocated for public passenger transport, reduced congestions within road network, increased comfort of travelling, less stress for transport network participants, less polluted environment and last but not least increased attractiveness of the disadvantaged regions.

Authors´ Preface

The methodology shows that the available technologies can be used to efficiently determine passengers assigned to a particular vehicle at a particular time in a particular point on a given route.

It could be assumed that when such information is available, the management of the public budget resources allocated for the subsidies of the public transport could improve. In 2015, these resources totalled almost CZK 19 billion, of which CZK 5.1 billion were subsidies of bus passenger transport excluding public transport, and CZK 13.8 billion were subsidies into passenger railway transport.

The principal areas of application can be summarised as follows.

a) Demand for the provision of public service in railway passenger transport in the public interest, i.e. the specification of input parameters for the opening of tenders for transport services, and the determination of other contract elements inducing the increase of the share of public transport on the overall transport demand.
b) Review and management of existing contracts for the provision of public services, i.e. monitoring real and reported performances or monitoring the ratio of the public transport to overall transport demand.
c) Bottlenecks inhibiting the reinforcement of the transport infrastructure, i.e. defining the priorities for the strengthening of the transport infrastructure considering not only the occupancy in graphs but also the overall potential of transport demand and its development trend in time.

The pre-requisites for the project implementation stated above need to be accompanied with technical pre-requisites which are crucial to improve the society's quality of life in terms of mobility.

a) Projects should be launched sufficiently in advance to generate time series of values for the monitored indicators.
b) A wide political consensus on the application of the certified methodology in projects should be achieved since the time framework of the projects exceeds the election periods of the Chamber of Deputies and regional and municipal councils.
c) The contracts for the provision of public transport services concluded between the carrier and the ordering party, both on the national and municipal level, should stipulate the obligation of mutual exchange of machine-processable and freely publishable data about vehicle traffic, data obtained from the check-in systems or the number of tickets sold.
d) The public administration should enable open-data access to data sources which had been purchased or funded by the state. For the purposes of the methodology, these include for instance data from the toll gates operated by the state based on various tenders since 2007, output of the RODOS project in the form of a comprehensive database of persons and goods mobility in the Czech Republic (realised between 2012 - 2018 with a total amount of eligible costs CZK 212 million), the output of the system for continuous area monitoring of traffic flow dynamics within the Czech Republic's strategic network of roads (the announcement to open a call for tenders from August 2015, estimated value of CZK 290 million).
e) To ensure a suitable data format for sharing the existing data about the railway network in coordination with the Railway Infrastructure Administration.
f) As regards road network, the Ministry of Transport should secure a wide coordination of data formats and records about the road infrastructure given the fact that road networks are administered by various entities (regions, municipalities, Road and Motorway

Directorate), resulting in heterogeneous data sources to be processed.

Both opponent reviews agree that the solution presented in the methodology is innovative and yields benefits. At present, the acquisition of data sources could be funded from EU Funds as the expected outcome achieved by means of the projects envisaged by the methodology fulfil the objective of EU policies. To achieve the highest efficiency possible, the public administration needs to maximise synergic effects from existing investments. An example of major synergic effect of the investments into data processing infrastructure could be the exploitation of the capacity of national supercomputing centre.

The volume of information and data is growing geometrically. We maintain that the key is that the output of the existing projects is understandable both to professionals and the wider public, justifiable and publicly available. The methodology contains several examples which require in particular the knowledge of the dynamics of the transport demand and its individual components. The examples have not been chosen randomly. Their selection was preceded by an analysis of dozens or hundreds of possible scenarios. The qualification criterion for the inclusion in the methodology was their applicability and potential benefits in practice. Clearly interpretable information was selected to facilitate answering particular questions and adopting tailored measures. For instance the knowledge of hourly, daily or monthly elasticity of performed transport demand using the railway and individual passenger car transport enables identifying the group of transport network participants which contributes to the peak, where they are from, what is their transport behaviour and subsequently streamlining group's behaviour towards sustainable means of transport. It also enables to assess whether the measures adopted fulfilled the expectations. Similarly, in combination with other information, other factors relating to the occupancy of the transport network (weather, public holiday, weeks

and days in month, season, etc.) could be identified. Such information can later be used to predict the transport network occupancy and to adopt measures to increase transport fluency and safety and to shorten the time spent commuting.

The economic benefit of the examples listed is measurable. We believe that only such comprehensible basis enables launching the dialogue between the public administration, businesses and consumers about viable measures from which the society as a whole could benefit. We are willing to continue contributing to this dialogue.

Ostrava, April 2017

Miroslav Vozňák and Jiří Hylmar

Contents

List of figures

List of tables

1 Introduction and summary

1.1 Objective of the methodology

The promoter and the guarantor of the research need TB0500MD011 is the Ministry of Transport of the Czech Republic. České dráhy, a.s. is the special guarantor of the project. The output of the research project is the certified methodology. The methodology also comprises the design of the framework system for sharing information about passenger movements in public transport. The objective of the methodology is to define an efficient use of available technologies to determine the number of passengers assigned to a particular vehicle at a particular time in a particular point on a given route.

The methodology provides an economically beneficial alternative model to monitor the passengers on individual fixed routes (with a specific system of passenger´s check-in) and to determine the number of passengers. The economic benefit lies in the substitution of the existing passenger monitoring systems and systems determining passenger number, while including previously unavailable data. In both cases, data relating to the number of passengers transported may be utilised in practice. When processing the methodology, the assumption that the currently available data are not sufficiently mined by the professional public proved valid. The reasons are both technical (actually poor and difficult-to-interpret content) and procedural and administrative (licence agreement, confidentiality agreement, etc.).

Areas with specific system of passenger´s check-in and number of transported passengers in which the technology could also be implemented.

1. Determining the use of particular fixed routes and their reconciliation against the report of specifically checked-in passengers.
2. Planning the activity of carriers' inspectors or the reassessment of passenger check-in.
3. Initial data to support information for infrastructure administrators to establish the utilisation of a particular fixed route.
4. Efficient passenger check-in with the emphasis on mitigating revenue loss and optimising passenger check-in and transport type expenses.
5. Addressing accidents within the Integrated rescue system.

The data may be applied to the structure of transport zones established in accordance with basic territorial units and networks of individual transport modes of the provided module. The results may be stored as matrixes processable in the given module; the module could also comprise information about the individual fixed routes within the networks. The data may also be displayed in charts.

The technology for processing mobile network signalling data does not enable ascertaining whether a passenger has been checked-in specifically or otherwise. This aspect is not relevant in terms of technology's aims and purpose as it enables to ascertain the number of passengers on any route in any place and at any time, provided all defined pre-requisites and limitations are adhered to.

The data and visualisations within the methodology are solely illustrative. No particular conclusions about values or situation in a given territory can be derived. The methodology cannot be applied and interpreted otherwise than as a guideline to process this type and quality of data.

1.2 Description of the application of the certified methodology, its target audience and its application

1.2.1 Project benefits and impact

Efficient planning and realisation of public transport with specific passenger check-in. Additional information about the utilisation of such transport throughout the year and not only as a result of ad-hoc passenger surveys. It also allows for better tracking of the transport flows for this type of public transport, both from the point of view of passenger volume and the length of the trip from various points of view (time-wise, weather-wise, etc.). The project output shall enable quantifying the total transport demand and the share covered by the public transport. The output can also be referred to by Integrated Rescue System in case of emergency.

1.3 Expected application of the results and project output users

1.3.1 Knowledge-based decision-making

The current social prerequisite is that the public administration measures and policies should lead towards the new economic growth in the Czech Republic and the entire EU. To successfully face up the innovative challenges of the so-called intelligent and sustainable economy which supports social inclusion requires implementing new intelligent decision-making solutions. In selected cases of application, the new technology reduces the operating decision-making deficit consisting of non-availability, obsoleteness or disproportionate costs of the provision of the underlying materials.

From the point of view the methodology, mapping the spatial and temporal mobility can either be performed deductively where theory-based hypotheses are constructed, empirically verified and adopted

or refuted; or inductively where based on data analysis, regularities are sought based on which new conclusions and theories can be formed.

From the point of view of the public administration, the substantial advantage of spatial and temporal data obtained from the mobile networks compared to the existing methods of establishing relevant indicators is the high-level standardisation, method continuation, quality, continuous calibration and ultimately more reliable information values.

The principal users of the methodology are the public transport commissioners, regulators and providers, the Integrated Rescue System which can use it for operative, short-term and middle-term planning and strategic planning.

The transport providers will mostly find use for operative/short-term opportunities. Individual components of the public administration – cities, regions, ministries and the components of integrated rescue system may exploit data for particular short-term measures relating to the public transport organisation they provide; to set the infrastructure of the region they govern in the middle-term; and to realise the master plan in the strategic long-term horizon.

1.3.2 Statutory cities, regions

Thanks to the spatial and temporal data from the mobile networks, the public transport providers will gain detailed information not only about the utilisation of routes and individual lines, in particular from the point of view of changing, boarding and disembarking stations, but also about their availability in time. This enables to create better and more operative schedules, streamline the routes, capacities and stations. Based on dynamic and real-time updated maps of traffic networks utilisation, the passengers will be able to plan better, thus using the public transport services more comfortably.

As regards the transport infrastructure, the spatial and temporal data from the mobile networks can be used as a supporting information source for the tasks addressed by the current projects.

1.3.3 Integrated Rescue System (IRS)

Based on the availability of the spatial and temporal data and thanks to their flexible deliveries, we assume that they would enable monitoring population movements in case of emergency, public events or natural disasters. We expect that such information should enable better organisation and exit routes management, better managed and planned population movement during a public event, etc.

1.3.4 Cutting down expenses on existing passenger controls in public transport

This involves the online passenger monitoring based on the number of tickets sold and the number of passengers established by the methodology per transport type. The inspectors could be sent directly to places which have been actively identified, and where a difference between the number of passengers onboard and the number of checked-in people has been established.

Such an application of the methodology presupposes that the number of actually checked-in passengers is known, setting the basis with which the data obtained by means of the methodology could be compared. In practise, this requires the use of the check-in/check-out systems (the principles for the services of the public implementation of systems of the automated passenger check-in should be incorporated into the tenders in 2019/2020).

1.3.5 Application in other areas

1.3.5.1 Land-use planning

The public participates in the land-use planning procedure, monitors the results of the measures and the circumstances of their implementation, and takes into account the information from underlying territorial analytical materials, in particular when making housing decisions. The municipality, or the region, is responsible for the land-use planning

documentation based on which decisions are adopted within the given territory. The businessmen and businesses take into account information included in the land-use plans when assessing the investment and operation intentions.

Information about transport demand and offer enable adopting objective and rational decisions about the placement of activities in the space, comparing spatial and organisational measures in localities and gaining additional underlying materials for decisions about project alternative scenarios when having limited budget resources. Another major issue which could be solved is the prevention of the occurrence of socially excluded areas and monothematic localities where people only commute for entertainment, work and the locality is not organically bound.

At present, data for land-use planning rarely include information on actually territory use by the inhabitants and the traffic burden, or, if so, in a hard-to-measure, mostly descriptive format. For the local and regional self-government, it is difficult to efficiently, objectively and operatively answer questions as to how many, what kind of inhabitants and at what time are present on a given territory, or whether the territory use pattern by the inhabitants changes during the day, week, season, or in relation to the measures adopted within a territory, and impact it has on the transport network utilisation and performance etc. when planning the territory use.

1.3.5.2 Planning and review of subsidies

The pre-requisite of an efficient measurement is the explicit definition of the objective and its measurability. Spatial temporal data about mobility could be used to measure specific objectives in individual areas of public administration and to present the measure outputs to the public. Spatial temporal data offer new, objective and easily readable indicators both for the allocation of the public budget resources and for the measuring of the results. It is for instance possible to operatively establish the increase in the number of inhabitants in the project locality, for instance new sport venues, cycling path, parks, changes in traffic

intensity relating to the changes in transport infrastructure and pertaining environment pollution by noise and CO_2, shortening of the span of public service availability by its relocation or by changes the areas which use the service, etc.

1.3.5.3 Tourism

In tourism, spatial and temporal data are important for the so-called destination marketing. Visit rate can be assessed significantly more precisely and operatively from the point of view number, time and origin of visitors, movement in locality, the share and origin of foreign visitors, etc. In this respect, the technology could also be applied to provide targeted services in state and public administration, publicly beneficial companies, associations, non-profit organisations by means of smart phones: thematic maps, information and guides, provision and storage of feedback in respect of measure and activities.

Last but not least, tourists may, based on the intensity maps and destination use mapping, better plan and choose their travel destination.

1.3.5.4 Identifying remote areas and countryside

Spatial and temporal data may provide supporting information for operative and strategic establishment of conditions within remote areas and countryside from the point of view temporal availability of individual municipalities in relation to the administrative centres and working micro-region centres.

1.3.5.5 Social relations and social network structure

Spatial and temporal data may contribute to map social network structures and micro-region interactions. Thanks to the diverse communication in regions, it will be possible to create coherence indices, or indices of region inhabitant deprivation relating to the social interaction. When a sufficiently long-time series are available it is possible to determine the impact of migration from cities into the countryside and vice versa on social relations and society coherence. Thanks to such information,

it will be possible to better target the rural development policies both at regional and national level.

1.3.5.6 Planning and reviewing the impact of interventions

Data on mobility could also be used to measure specific objectives in individual areas of public administration and to present the measure outputs to the public. Spatial and temporal data also offer new and objective indicators both for the allocation of the public budget resources and for the measurement of the results. It is for instance possible to operatively establish the increase in the number of inhabitants in a particular locality as a result of business and leisure time visits to the region, changes in time availability of regional centres, public services, employment, etc.

1.3.6 Research issues

In the transport research, localisation data can be used as follows.

1. O-D matrixes (relation information about trip's origin and destination).
2. Current traffic intensity.
3. Movement speed.
4. Trip length.
5. Traffic density in certain areas.

The application example described in the methodology covers in particular relation information about the trip's origin and destination, current traffic intensity and traffic intensity in certain areas. Namely.

1. How many residents in the municipality can use the route in days and daily hours?
2. How many people present in the municipality for at least 60 minutes may use the route in days and daily hours, and how many have actually used it?
3. What is the stretch route occupancy and available line capacity?
4. What is the occupancy compared to the total possible route occupancy?

Possible issues to be further investigated and for which the example of use provide certain underlying information include, among other, the following issues.

1. What kind of participants of the transport networks contribute to the rush hours, where are they from; and what is their traffic behaviour? Is it possible to work with this group and if so, in what manner (even/odd tickets, marketing policies, "My train" application, etc.)?
2. Are the rush hours the same across the mobility, modal split?
3. What is the relation between route occupancy and stretch route occupancy in the modal split (public transport capacity, congestions)?
4. Is it possible to identify factors which may impact capacity (weather, national holiday, weeks and days in a month, season, etc.)?

1.4 Linking the methodology up to the existing strategies

1.4.1 Fulfilling the existing strategic objectives

The results of the research fulfil the strategic objective no. 8 as defined by the Action plan for intelligent transport system development ("AP ITS") in the Czech Republic until 2020 (with the prospect of 2050) towards 'Intelligent Transport System' (ITS) development in particular as regards passenger transport and (among other things) systems.

The project shall contribute towards the implementation of AP ITS, the fulfilment of the Czech Republic's Transport Policy's objectives in the area of passenger transport, and the following objectives of other areas of the Czech Republic's Transport Policy:

- To plan the development of the transport infrastructure with regard to ensuring high-quality reachability of all regions,

to plan the development of technologies on the basis of satellite and ITS systems with regard to transport needs;

- To build high-quality transport infrastructure and to equip it with modern ITS technologies for all types of transport, including the infrastructure for multimode interconnection of individual passenger transport types;
- To implement measures to increase the use of railway and waterway transport.

1.4.2 Major references

1.4.2.1 The Transport Policy of the Czech Republic for 2014 – 2020 with the Prospect of 2050

On 12 June 2013, the government of the Czech Republic adopted 'the Transport Policy of the Czech Republic for 2014 – 2020 with the Prospect of 2050". The Transport Policy is a superior conceptual document for the 'Transport Sector Strategies, 2nd phase' (TSS-2), approved by the government of the Czech Republic on 13 November 2013, defining the country's key transport structures and the priorities as regards the newly built European networks. Accordingly, the ITS development in this paper focuses mostly on ITS systems which make part of the transport networks on motorways, dual carriageway roads, 1st class roads and the urban road infrastructures. Transport Sector Strategies also define the overall financial needs for the development of ITS systems embedded in transport networks and form an inherent part of the transport structure. Since TSS-2 do not cover the entire ITS issue, the Action Plan for the Deployment of Intelligent Transport Systems (AP ITS) had to be drafted. TSS-2 state that individual measures shall be defined in the Action Plan.

1.4.2.2 Action Plan for the Deployment of Intelligent Transport Systems (ITS)

Action Plan for the Deployment of Intelligent Transport Systems (ITS) in the Czech Republic until 2020 (with the prospect of 2050).

The Resolution of the Czech government from 15 April 2015 no. 268 on the Action Plan for the Deployment of Intelligent Transport Systems (ITS) in the Czech Republic until 2020 (with the prospect of 2050).

The Implementation Plan to the Action Plan for the Deployment of Intelligent Transport Systems (ITS) in the Czech Republic until 2020 (with the prospect of 2050). The Resolution of the Czech government from 15 June 2016 no. 538 on the Implementation Plan to the Action Plan for the Deployment of Intelligent Transport Systems (ITS) in the Czech Republic until 2020 (with the prospect of 2050).

1.4.2.3 ITS Implementation Plan

The ITS Implementation Plan provides the strategic framework for the fulfilment of the follow-up measures of the Action Plan for the Deployment ITS's objectives. Its aims include in particular the following.

- To identify the set of project intents fulfilling the measures of the Action Plan for the Deployment ITS's objectives and continuously update and adjust them.
- To assess the readiness of project intents.
- To assess the project intents from the point of view of their potential to fulfil follow-up measures of the Action Plan for the Deployment ITS's objectives.
- To provide an overview of the funding of project intents until 2020.

When implementing the measures of the Action Plan for the Deployment of Intelligent Transport Systems, the departments of the Ministry of Transport play a crucial role since they are responsible for the design and implementation of the project intents.

2 Methodology for the control of the number of passengers with a specific system of check-in

2.1 Recommendations for working with spatial temporal data from the mobile networks

2.1.1 Introduction, interpretation of results

This methodology does not aim at substituting or casting doubt on the existing indicators currently used in the transport practice. The methodology introduces new indicators enabled by the technological progress which are meaningful on their own. The existing methods and indicators remain important, in particular in the early stages of the implementation of new assessment procedures. They form the basic control mechanism, a certain 'gold standard', and any deviations from it have to be justified. It is not until the new approaches have been implemented and verified in practice that certain existing methods could be abandoned and replaced with other, more efficient and comprehensive ones.

2.1.2 Daily cycle

The prerequisite to maintaining the possibility to interpret the spatial temporal data in relation to social economic phenomena in a territory is the definition of a DAILY CYCLE. The daily cycle starts at 0:00:00 and ends at 23:59:59.

The interpretation of any phenomenon in a territory from the point of view of inhabitant mobility assumes the so-called daily cycle, which in practice, in particular in the mobile network, is represented by the fact that the participant leaves in the morning and returns to the station which thus becomes is home base in the evening.

One should, however, be aware of the fact that a certain group of participants fails to fulfil this prerequisite. These participants may be included in the statistics of the residents based on various input conditions (e.g. 5 times a week in the same location, etc.). From the point of view of transport demand which monitors the daily status, this fact should be disregarded in order to maintain the comprehensibility of the classification applied.

The share of employees working shifts in the Czech Republic amounts to 30%. Shift work, however, can have different regimes. Shift work not only includes night shifts but more importantly also weekend shifts. Detailed analyses and data from the mobile network indicate that 95% of shift-working employees fall into the pre-defined 'daily cycle'. Most shift-working employees work in the manufacturing industry (457 thousand.), in trade (148 thousand) and in health care and social welfare (124 thousand).Other reasons for exceeding the daily cycle (such as entertainment, travelling, etc.) are, from the point of view of the interpretation of spatial temporal data, statistically immaterial. When necessary, it is possible, based on a mobile network data, to perform specific analyses focused on a small group of transport network users who exceed the time cycle and to determine its size, periodicity of occurrence and other circumstances.

2.1.3 Uniqueness of an individual in time and space

Every participant is classified as a unique individual on that given day in terms of the day and the station and cannot be included in more than one category. The primary classification is made based on spatial temporal behaviour of the individual by means of a retrospective analysis of the movement and occurrence of that particular person during 24 hours.

The basic classification is as follows.

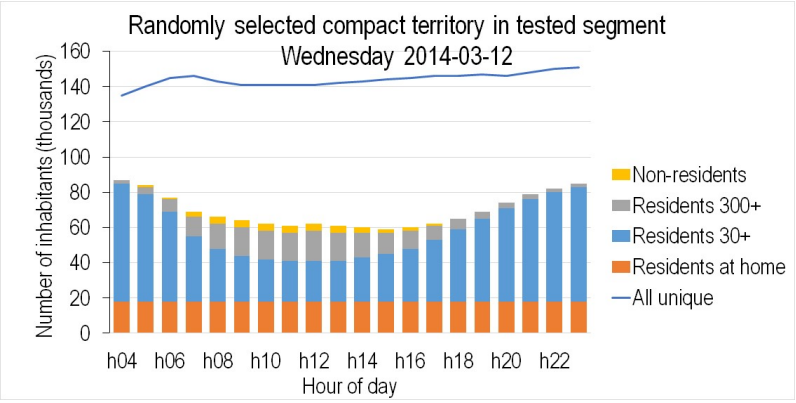
RESIDENTS

OUT-COMMUTERS

NON-COMMUTERS

IN-COMMUTERS

TRANSITING



Note: Non-residents: non-residents in-commuting to the territory; recorded in destinations (300+ min. spent time) in the territory. Residents 300+: residents of the territory out-commuting to destinations (300+ min. spent time); recorded at destinations (300+ min. spent time) located in the territory. Residents 30+: residents of the territory out-commuting for 30 and more minutes; recorded at home before departures and, or after arrivals. Residents at home: residents of the territory not commuting; recorded at home. All unique: all unique; recorded in the territory.

Figure 1 *Example composition of participants in space and time– separate localisation data*

Unless stated otherwise in more specific cases, each person may only be included once in a particular time slot. **Figure 1** demonstrates the example of a possible presentation.

The classification applied does not, however, inhibit the application of a different classification covering a longer time period. In fact, quite the contrary. This is the corner stone which enables monitoring of the behaviour of the population in time and space during a day, a task highly-sensitive to the ability to interpret and verify the results. Other tasks, supplementing the existing statistical indicators (e.g. commonly residents) also derive from this universal platform.

2.1.4 Averages applied

One of the methods to address the phenomena which occur in relations to the natural dynamics of the monitored indicators is to define suitable averages and corresponding variation coefficients. Naturally, certain fluctuations may occur. Yet certain patterns in the behaviour of residents (out-commuting and in-commuting), both steady and less steady, can be identified.

In general, working with the averages for a particular period such as a day in a week, season of the year, days with a specific weather, employment rate, etc. should be regarded as one of the key aspects of work with spatial temporal data. This enables creating multiparty correlations and obtaining a rational basis for the clarification and causality of hitherto unknown or un-quantified social-economic phenomena.

2.2 Input parameters

2.2.1 Introduction

The aim of this section is to define the areas which need to be taken into account when assigning the task.

2.2.2 Time period and time slots

The period is defined in calendar days. A calendar day comprises the shortest time interval set for the monitoring of spatial temporal phenomena – most frequently this is 60 minutes.

Admissible values for time intervals include longer time slots consisting of entire hours or days. Similarly, time slots can be shortened. The recommended shortest time slot is 15 minutes.[1]

When using an average or mean value, the referential period and time basis is stated.

2.2.3 Transport system participants

The selection of transport system participants, the mobility of which (change of place for the purpose of fulfilling a need)[2]the task analyses, is made according to general guidelines for the processing of spatial temporal data from the mobile network, i.e. by means of the request consisting of the 'time– station'. The basic classification of participants with mobility within the daily cycle:

- Residents (i.e. started and ended the daily cycle in the station);
- In-commuters (i.e. stayed in the station for time t > x);
- Transiting participants (i.e. stayed in the station for time t< x].

2.2.4 Types of transport

The selection of the transport type is made in accordance with the APITS terminology (chapter 1.4.1). The basic classification is as follows:

- Road transport;
- Public passenger transport;
- Railway transport.

[1]This takes into account the existing technology options based on the processing of signalisation data from the mobile network. This relates to the so-called periodical update.
[2] In a broader sense, the interaction by means of communication networks (information service and electronic trading service) may also fall within the definition of the mobility. The analysis of this type of mobility is not the focus of the methodology.

Railway freight transport, cargo road transport, inland waterway transport and air transport are not covered by the methodology.

Note: The example of the methodology application in chapter 2.5 is provided for the railway transport, taking into account the volume of road and railway transport within the transport network, i.e. the total transport demand, and the overall amount of people within the territory.

2.2.5 Stations

The stations are perceived as spatial boundaries for the task. Stations can also be interpreted as the place of the realisation of activities, projects and passenger routines. From the spatial point of view, the stations are fixed.

As regards railway transport, the basis for the selection of a station is the places in which the monitored lines stop at least once within the reviewed period. These places are defined by the stops and represented by a unique stop number and coordinates[3].

From the spatial point of view, the stations are defined by a polygon. For the purpose of processing data tasks, the polygon is represented by a definition point, either significant or geometric (so-called 'centroid'[4]).

2.2.6 Territory

The territory for the completion of the task is defined by the boundaries of administrative of functional territorial units and the existence of signalisation records from the section (cell) of the mobile network covering the unit.

[3]The current public form of CISJR does not clearly determine the position of stops.
[4]See for instance significant focal points of territorial units
https://www.czso.cz/csu/rso/ekatalog_gp

[Certified translation from the Czech language]

Ministry of Transport
of the Czech Republic

nábřeží Ludvíka Svobody 1222/12
PO BOX 9, 110 15 Prague 1

Prague, 21st December 2016
File no. 133/2016-710-VV/1

Has issued this

CERTIFICATE

Accrediting the applied certified methodology
in accordance with the "Methodology for the Assessment of Research and Development Results" guidelines
titled

"Specific system of passenger's check-in and number of transported passengers"

RESOLVER

VŠB – Technical University of Ostrava

Author:

Doc. Ing. Miroslav Vozňák, Ph.D.
Bc. Jiří Hylmar, MBA
Bc. David Blagodárný
Ing. Jindřich Borka
Ing. Jan Dušek
Ing. Milan Kříž
Ing. Lukáš Orčík
Doc. Ing. Jan Platoš, Ph.D.
Ing. Jan Růžička, Ph.D.
Ing. Marek Sedláček

Developed in the framework of *research project* no. TB0500MD011 - *"Specific system of passenger´s check-in and number of transported passengers"* supported from the funds of the Technology Agency of the Czech Republic, within the BETA programme.

Authors of 2 independent opponent reviews:

- Prof. Dr. Ing. Miroslav Svitek, dr.h.c., CTU of Prague, Faculty of Transportation Sciences
- Ing. Jan Přikryl, Ministry of Transport of the Czech Republic, Public Transport Section

JUDr. Václav Kobera

[illegible signature]
Director
Intelligent Transport Systems, Space Activities and R&D Section

Translator's Clause
As a certified translator of the English language appointed by the decision of the Regional Court in Prague from 14/12/2011, ref. no. Spr 4078/2011, I hereby certify that this translation corresponds with the original document attached.
I have made the following corrections in the translation: .....---.....
The certified translation is registered in the certified translator's book under the serial number of 72/5/2017.
Michaela Buchtelová
Lety
18 April 2017

The results of our research presented in this publication are based on in-depth technological knowledge of cellular networks; data distributed computing in the Appache Hadoop platform as well as other technical and economic disciplines. The scope of the book focuses on population mobility data that become available once the above mentioned knowledge has been mastered. It ought to be said that our intention was to introduce the basic building blocks on which subsequent analysis could be applied in a simple way. The publication contains several examples of application and describes some previously unknown phenomena discovered, such as the need to consider a dynamic number of residents in the locations. Nevertheless, we intentionally leave up to the readers to consider the possibilities for an even better fulfilment of their duties. We believe that only by implementing such information into our daily routines can we achieve our ultimate goals - enhance citizens' quality of life and protect our environment.

Miroslav Vozňák and Jiří Hylmar

ISBN 978-80-248-4039-0

Vozňák & Hylmar: Population Mobility Data Retrieval from Cellular Networks and its Use Case in Public Transport