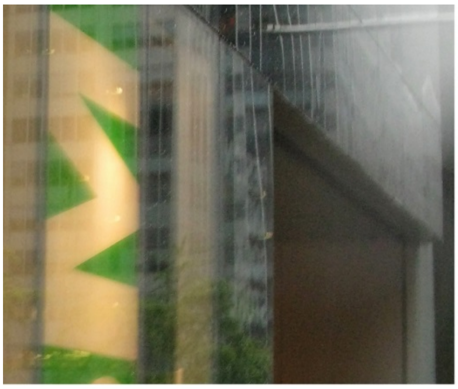
Trivector

Trivector Traffic

Rapport 2017:32, Version 1.2

Detailed Mapping of Tools and Applications for Travel Surveys



Trivector

Document information

Title:	Detailed Mapping of Tools and Applications for Travel Surveys
Series no.:	2017:32
Project no.:	17046
Authors:	Anna Clark Emeli Adell Annika Nilsson Lovisa Indebetou
Collaborators:	Erik Stigell, Lennart Persson, Luke Hobbs, Thaddäus Tiedje, Johan Kerttu, Astrid Michelsen
Quality review:	Annika Nilsson
Client:	Transport Analysis Contact person: Eva Lindborg, 010-4144244, 072-5202460, Eva.Lindborg@trafa.se

Document history:

Version	**Date**	**Change**	**Distribution**
1.2	2017-08-23	Minor supplementations	Client
1.1	2017-07-13	Minor adjustments	Client
1.0	2017-07-07	Final report	Client
0.9	2017-05-31	Draft of final report	Client
0.8	2017-04-13	Preliminary version	Client

Trivector Traffic Åldermansgatan 13 · SE-227 64 Lund / Sweden
Telefon +46 (0)10-456 56 00 · info@trivector.se

Foreword

A need has been identified within the *Next generation travel and transport* co-operative programme to find and develop new solutions for travel surveys (TSs). A project plan concerning new methods for TSs, for which Transport Analysis was the primary coordinator, was submitted to Sweden's Ministry of Enterprise and Innovation in January 2017. The project plan was divided into five work packages: Coordination (1), Stakeholder requirements (2), Detailed mapping of tools and applications (3), Testing and analysis of tools and applications (4), and Development potential and recommendations (5). This report pertains to work package 3, which has been carried out by Trivector. Work package 2 was carried out by the Swedish National Road and Transport Research Institute (VTI), in parallel with Sweco and Trivector.

Lovisa Indebetou served as project manager at Trivector, with Annika Nilsson performing the quality review on the report. Emeli Adell and Anna Clark shared primary responsibility for the mapping, analyses and report writing, with the support of a number of people, including Erik Stigell, Lennart Persson, Luke Hobbs, Thaddäus Tiedje, Johan Kerttu, and Astrid Michelsen.

A reference group was affiliated with the project to ensure its quality and practical applicability. The group included Abboud Ado, along with client representatives Per-Åke Wikman, Eva Lindborg and Mats Wiklund, plus Andreas Holmström from Transport Analysis. Other reference group collaborators are listed in Chapter 7. The original report was written in Swedish.

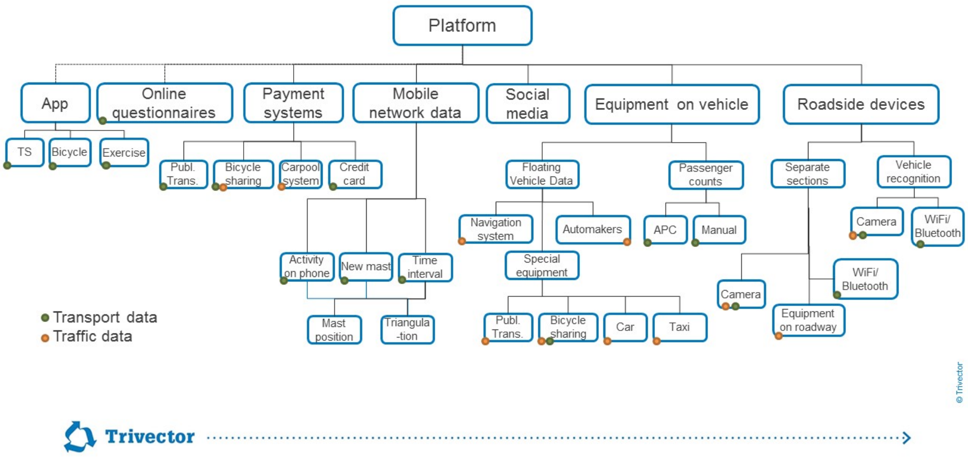
Lund, August 2017

Summary

A need has been identified within the *Next generation travel and transport* co-operative programme to find and develop new solutions for travel surveys (TSs). This led to a project concerning *New solutions for future travel surveys*. This report pertains to one work package within that project: Detailed mapping of tools and applications.

The overall objective is to study new technical data collection solutions and how they could be used, individually or in combination, in future TSs as the basis for official statistics. The purpose of this work package is to inventory the national and international arenas in this field, and to map and characterise the level of technical maturity present among the applications and tools that have been or are soon to be developed to collect data about travel behaviour.

New technical solutions for data collection have been mapped through documentation analysis, a literature study, and contacts with experts/developers and a reference group. The mapping process has resulted in an overview of different types of traffic and passenger measurements (data sources/data collection methods), and a categorisation of tools and applications for collecting data for travel surveys – both big and small data, and passive and active collection.



The data sources/methods measure different parameters, mainly movements of vehicles (traffic) or movements of people (transport). They also measure various subquantities (e.g., all modes of transport or types involving individual vehicles), and entail various sample selections (total survey and/or sample of persons or sample in terms of time and space). The measuring methods are also encumbered by various measurement errors and statistical losses.

The data sources/data collection methods that individually answer most of the questions of importance in TSs are those specifically designed to conduct these surveys, in other words the traditional TS, apps designed for collecting TS data and online questionnaires for collecting TS data. However, the ability to answer different questions statistically requires data sets of differing sizes and data quality, with the result that, in practice, these sources cannot be used in isolation to answer all the questions. Multiple sources can be used simultaneously to obtain a broader view of the travel patterns as a whole.

The scientific literature offers examples of how data from data sources are combined to obtain a better picture of transport or traffic activity as a whole. The literature shows that all data sources suffer from skewness and sample selection problems, but that there is little knowledge that bears on sample selection problems and representativity specifically.

The mapping process identified eight Swedish tools and applications that can collect relevant data for TSs: a platform that uses data from multiple data sources in parallel, three apps and four roadside devices. Roadside devices have come the farthest in terms of development, along with the collecting of Floating Vehicle Data (FVD), and several such commercial tools are available in Sweden, while platforms that collect data from different sources (and specifically mobile network data) are still in the development stage. None of the Swedish solutions currently offers a total solution or provides an overall picture of personal transport due to sample skewness, or because they only offer means of counting certain types of vehicles/road users. The price situation is unclear; passive data collection methods are less expensive but do not offer data linked to individuals, while active data collection methods offer richer data but require more effort, mainly in terms of recruiting participants.

The mapping process identified some 50 international tools and applications capable of collecting TS-relevant data. They include platforms, TS apps, other types of apps, FVD, and roadside devices. Given the search words used, the mapping was more comprehensive for TS apps than for other types of tools and applications.

The international methods reflect new solutions and ongoing research projects in Sweden. The review shows that companies that develop data collection tools also own (and sometimes sell) data that have already been collected for other purposes, but that could also be used to answer TS-relevant questions. It is, however, difficult to ensure the quality of data collected with these new tools, as their data cleansing and processing algorithms are not usually published. The costs are also unclear.

Legal conditions associated with the collection, sale and use of personal data vary from country to country. The rules of the game will also change when the European Union's (EU) new General Data Protection Regulation goes into effect on 25 May 2018. This could have an impact on which tools can be used in which countries, and what adaptations to the law will need to be made.

Internationally there are a number of products (e.g., TS apps, roadside devices) that are similar to the Swedish ones. It is not possible to identify which are best, as their advantages and disadvantages depend on how they are used. There are a few areas in which Sweden has products that are on an equal footing with what is happening in the rest of the world. We also have an active mobile network and automotive industry that could be a conduit for new solutions. Common to all the tools and methods is the fact that more knowledge is needed concerning recruitment, sample selection and statistical losses.

Table of Contents

1. Introduction		**4**
1.1	Background	4
1.2	Purpose and goals	4
1.3	List of definitions	5
1.4	Chapter descriptions	7
2. Method and delimitation		**8**
2.1	Mapping process and delimitation	8
2.2	Literature search method	8
2.3	Analysis and synthesis	9
3. Starting points for the mapping process		**11**
3.1	Various types of traffic and passenger measurements	11
3.2	Current state of knowledge per the scientific literature	20
3.3	Important questions for travel surveys	28
3.4	Quality aspects	33
3.5	New collection methods' potential for overcoming current deficiencies	37
4. Results for national tools and applications		**41**
4.1	Swedish examples – an overview	41
4.2	Platforms	45
4.3	Apps	47
4.4	Roadside devices	48
4.5	Summary and conclusions	49
5. Results for international tools and applications		**51**
5.1	International examples – an overview	51
5.2	Platforms	52
5.3	Apps	54
5.4	Mobile network data	62
5.5	Floating Vehicle Data	63
5.6	Roadside devices	64
5.7	Summary and conclusions	65

6.	**Conclusions and recommendations**	**67**
6.1	Summary and discussion	67
6.2	Conclusions	69
6.3	Suggestions for further studies	69
6.4	Recommendations for pilot testing	70
7.	**References**	**73**

Trivector Traffic

1. Introduction

1.1 Background

A need has been identified within the *Next generation travel and transport* co-operative programme to find and develop new solutions for travel surveys (TSs).

The ability to design a future transportation system that can sustainably meet existing and future societal challenges will require a highly advanced understanding of how people travel and why they travel as they do. Current methods used for TSs have problems in certain respects with regard to coverage, costs and the information burden on the respondents. The trend toward greater digitalisation can contribute in this context, and new methods are being used to collect data for TSs.

Technological advancements, primarily digital in nature, are offering greater and broader means of collecting data to describe travel behaviour and mobility patterns (mobility data). The data collection methods used heretofore for TSs have, at the same time, encountered difficulties, primarily in making contact with the intended respondents. The new data collection methods can offer more data, usually at a lower cost than current methods.

New solutions for collecting mobility data will focus on describing various aspects of people's mobility, for example how they move about in a transport network, or what purposes or activities give rise to trips. In the future it will also be necessary to collect data from individual persons. If more options are available in the form of technical solutions that can be used to collect data, then different ones will work more or less well for different people and in different geographical settings (e.g., rural towns and cities).

Consequently, there is a need to develop methods and processes to enable the combination and integration of different data collection solutions. At the same time, it is important that these methods and processes be designed so that personal privacy is protected. In addition, methods for determining statistical weights for collected data need to be developed so that the data reflect the population that one intends to describe.

1.2 Purpose and goals

The overarching purpose of this project is to study new technical solutions for collecting data, and to study how different data collection methods could be used, individually or in combination, in future TSs as the basis for official statistics.

Our goal is to inventory the national and international arenas in this field, and to map and characterise the level of technical maturity present among the applications and tools that have been or are soon to be developed to collect data about travel behaviour or, more generally, people's mobility. The goal is also for the mapping to be able to serve as the basis for method selection in a pilot project that is planned for the fall of 2017.

1.3 List of definitions

Definitions and abbreviations used in the report are presented below.

Application Programming Interface (API) – An application programming interface specifies the ways in which different application programs can use and communicate with a piece of software. For example, it can specify the format for data that are being provided via a web service.

Automatic Passenger Counting (APC) – Electronic equipment installed on board a public transport vehicle to count the passengers getting on and off.

Base sites – Normally sites where individuals tend to be, for example at home, at work or at school.

Bluetooth Beacon – Small transmitters, usually battery-powered, that are based on Bluetooth Low Energy technology and transmit information about specific sites (Point of Interest) that can be picked up passively by a smartphone or tablet that is within range.

Ticketing and Payment Systems or **Ticketing and Payment Services (TaP) –** Services for paying for items (e.g., credit cards) and/or tickets.

Loss – Loss consists of those elements of a sample that fail to be included in a survey for various reasons. There is also an error source in statistical studies known as "loss error".

Call Data Record (CDR) (sometimes also Call Detail Record) – Data traffic containing information about telephone calls or other mobile phone use, in other words who was called or sent an SMS by whom and when.

Cooperative Systems – Connected and interworking systems. This field concerns systems that interwork with other systems, where connected vehicles with vehicle-borne technology receive information via technical solutions.

Data mining – Refers to tools for searching for patterns, correlations and trends in large data sets with the help of statistical calculation methods and efficient calculating algorithms for computer learning and pattern recognition. The purpose of data mining is to facilitate searches for structures among large number of variables, and to derive comprehensible and usable information from raw data.

Fleet management – The planning, supervision, control, and assessment of the movements and activities of a vehicle fleet, including its drivers.

Floating Vehicle Data (FVD) – Data that are collected by units on vehicles that are connected. This usually occurs in the form of Floating Car Data, which include localisation data, velocity, direction of travel, and time information about the vehicle, but may also pertain to other vehicles.

General Data Protection Regulation (GDPR) – Intended to enhance the protection of physical persons in connection with the handling of personal data within the EU. The regulation will go into effect on 25 May 2018, superseding the Data Protection Directive. It will apply immediately in all the member states, superseding earlier national provisions. In Sweden the GDPR will supersede the Swedish Personal Data Act.

Global Positioning System (GPS) – A satellite navigation system. Twenty-seven satellites (2015) make it possible for anyone with a GPS receiver to determine their position (longitude, latitude and altitude) anywhere on earth.

Interoperability – The ability of various systems, often in a computer context, to operate together and communicate with one another.

Intelligent Transport Systems, Intelligent Transport Systems and Services (ITS) – The use of information and communication systems and services in the areas of transport and transport infrastructure.

Mobility as a Service (MaaS) – A service that links various transport services together into readily accessible and comprehensive mobility services, offering a competitive alternative to the personal automobile.

Media Access Control (MAC) address – A unique identifier for each network card. The address consists of six bytes, each of which contains eight bits, which in turn have two possible positions, meaning that there are more than 10^{12} possible MAC addresses. They are typically used in LANs and WLANs for communication within the network.

Mobile network data – Data derived from mobile phone use.

Origin-Destination (OD) – Travel matrix or traffic matrix, in other words the number of vehicles or individuals moving between the areas, from starting point (origin) to destination.

Point of Interest (POI) – A place of interest for a visit.

Point of Sale (POS) data **–** Information from a sale at a given site where it is possible, via their payment card, to obtain information about the individual, such as their gender and age.

Population – A group of individuals present within a given area at a given time. In a travel behaviour context, the total group of individuals for whom one wishes to map travel patterns.

Swedish Personal Data Act – This act was intended to protect people against violations of their privacy violated in connection with the handling of their personal data. The act regulates how legal and physical persons (government agencies, companies, individuals, etc.) may handle personal data, primarily with a focus on data processing.

Radio-frequency identification (RFID) – A technique for reading information, for example identification, remotely from transponders and memories, which are known as tags.

The Internet of Things (IoT) – This comprises vehicles and everyday objects, machines and buildings that can be connected physically or via a wireless network and then exchange data because they have in-built sensors, computers and Internet connectivity.

Traffic Management System (TMS) – The management of traffic flows (people, vehicles and goods) via demand management, traffic information and other measures. Also known as Real-Time Traffic Management (RTMS).

Coverage – Specification of the chronological and spatial scope of the data collection process.

Sample – A portion of the total population that is selected for study. Formally this may be seen as a portion of all elements. Other words that mean the same thing are "random sample" and "selection".

WiFi – A wireless network technology (WLAN).

Open data – Open data consist of information that is available for anyone to use, reuse or distribute with no reservations other than source identification and sublicensing.

1.4 Chapter descriptions

Chapter 2 describes mapping methods. Chapter 3 describes starting points for the detailed mapping process, as well as different types of tools at a general level and how well they can answer various questions and fulfil aspects of quality. Chapters 4 and 5, respectively, describe in greater detail the national and international tools and applications that can be used to collect travel survey data. Chapter 6 sets forth conclusions and recommendations for the continued work within the project.

2. Method and delimitation

2.1 Mapping process and delimitation

New technical solutions for collecting data have been mapped via document analysis and a literature study. The picture was then supplemented through interviews and contacts with experts/developers to obtain a more complete idea of, for instance, what stage of development each tool is in. The reference group contributed with their knowledge as well.

Delimitation was performed so that only new data collection tools and new methods for utilising older technologies were studied.

Examples of types of tools and applications that were identified as being of potential interest include:

Mobile applications for travel behaviour

New types of travel behaviour questionnaires

Data from public transport

Mobile network data

Congestion tax data

Flow data

Data from travel planners

2.2 Literature search method

A systematic literature study was conducted with a view to mapping scientific articles about new tools and applications for travel surveys (TSs). The search was performed on 16 March 2017 in three databases: Web of Science, Scopus and LUBsearch. The search terms included in the search were as follows:

"travel survey" future “data collection”

"travel survey" new “data collection”

"big data" "travel survey"

"big data" AND (travel AND (behaviour OR behavior))

"data mining" AND (travel AND (behaviour OR behavior))

"data mining" "travel survey"

"mobility patterns" "data collection"

"mobility patterns" "big data"

"mobility patterns" "data mining"

The search yielded a total of 1,142 hits. The time window was restricted to between 2012 and 2017 because the focus of this project is on new methods and applications. Two hundred and four articles were then selected for further review based on their titles. In addition, searches were conducted using the same search terms in Google (in English), and translated into Swedish, Dutch and French. The literature was further supplemented with other reports of interest via contacts and interview subjects.

2.3 Analysis and synthesis

Analysis and interpretation of the collected information were performed with a view to studying the ways in which various new data collection methods could be used, individually or in combination, in future TSs as the basis for official statistics, and to address the following questions and subject areas:

What data are important if the tool is to work? How sensitive is the tool to data loss?

How does data collection occur? Can representativity be ensured?

What types of questions can the tool help to answer?

Assessment of level of technical maturity.

Any problems or risks in terms of data security and personal privacy?

Consequences of who owns the data material?

Efficiency/price picture.

After the collected information was analysed and interpreted, the results were categorised and systematised. The various tools and applications that were identified and described were categorised based on relevant aspects, but mainly with an emphasis on the types of data sets they offer. Other aspects were also elucidated such as:

Ability to answer different types of questions

Usability for travel surveys based on statistical samples

Statistical quality criteria

Resource efficiency

Legal conditions and assumptions

Based on the analyses and categorisation, a synthesis of the strengths and weakness, for example, of the various tools and applications was performed, and the conclusions were summarised. Based on the weaknesses and any question marks, and on the basis derived from the scientific literature, suggestions were made with regard to further studies that could provide greater clarity in terms of the ways in and extent to which the most relevant tools could be used in future TSs.

3. Starting points for the mapping process

The current national travel survey (TS) is a valuable tool for understanding how the transport system is being used in a number of different contexts. With its large and random sample and time-tested questions and collection methods, it has provided a representative picture of how the population of Sweden has travelled for decades. Without this information it would be considerably more difficult to plan, manage, follow up, develop, and maintain the transport system. In brief, the national TSs are important for an efficient transport system, and in terms of Sweden's ability to achieve its environmental goals.

However, the current collecting methods impose a heavy burden on the respondents who provide the information, with high costs and often many years between measurements. In addition, there are also deficiencies in the collected data material, both in terms of the quality of the estimated trip lengths and travel times (people have difficulty assessing this with any great accuracy), and in terms of route choice, which is not currently included in the TS. Certain questions have also arisen with regard to these surveys in recent years. One of the biggest problems is a suspicion of deficient representativity. This is a serious problem and reduces the usability of the TS. These are deficiencies that new methods could potentially rectify.

3.1 Various types of traffic and passenger measurements

In reviewing and compiling various types of tools/methods for collecting data that could help us understand travel patterns, a number of different main types have become clear (see Figure 3-1).

In this section, we will go through the various types and review the types of data they can collect. In TSs we are interested in understanding how people move, in other words *transport* data. However, some data sources provide information about how vehicles move, that is traffic data. The ways in which transport and traffic data relate to one another vary: the traffic data for bicycles can be said to be essentially identical with the transport data, while there are major differences for cars and public transport.

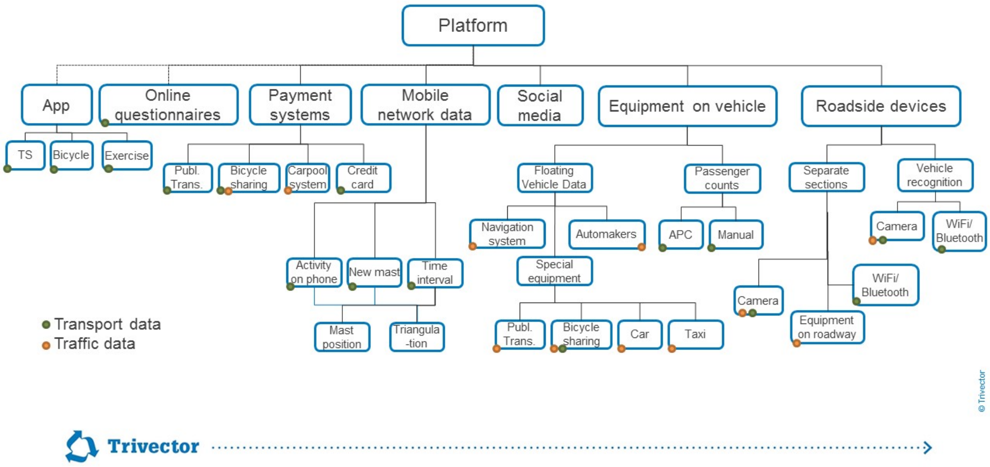


Figure 3-1 Compilation of different types of data sources that can aid in understanding travel patterns.

One way of categorising the various data sources is to consider whether they contain “big data” or “small data” (Chen et al. 2016). Big data refers to large and complex data sets that cannot be handled using traditional analysis methods. Such data are often collected primarily for purposes other than TSs, and without the involvement of the user. One example of big data is mobile network data, where information as to the location of a mobile phone is stored. There are enormous amounts of data concerning how mobile phones move, but they are collected for entirely different purposes than mapping travel patterns. Conversely, small data indicates that the amount of data tends to be relatively limited, often because the data collection process requires the active participation of those providing the data. Small data seldom contains large amounts of geographical information. The traditional TS is a typical example of small data.

A review of various categories of tools and applications for collecting data for TSs, both big and small data, is provided below.

Categories of interesting tools and applications

Apps that collect travel data

We found three different types of mobile phone apps that collect travel data: 1) apps that are designed to collect travel behaviour data, 2) apps that are designed to collect cycling data, and 3) apps whose purpose is to measure and provide feedback to the user with regard to their physical activities. All of these apps collect transport data, in other words person/kilometres at a level such that it is possible to track single individuals.

<u>Apps designed to collect TS data</u>
This type of app uses various sensors in the mobile phone to capture how the phone is moving, which is interpreted as the individual's movement pattern. They provide information about trips[1] and their purposes,[2] and often infor-

[1] The mode of transport, origin, destination, trip length, route choice, start time, end time, and duration are identified at the trip element level. Waiting/transfer times can often be identified as well.

[2] Data collected with regard to purposes: type of purpose, geographical point, start time, end time, duration.

mation about the individual[3] who has the phone as well. Different apps require different levels of user involvement. There are apps that automatically attempt to detect information about both trips and their purposes, and apps that automatically attempt to detect only trips. There are also apps that call upon the user to turn the logging function on and off so that trip data can be collected (albeit mainly in the world of research). Most apps ask that the user monitor (and optionally correct) the data, while certain researchers are developing apps with fully automatic identification of the mode of transport and purpose, so that, while not perfect for certain purposes, the results are "good enough". Data correction can take place either in the app or via a web interface. Information about the user is normally collected via a questionnaire in the app or in the web interface.

There are also apps that have in-built impact mechanisms (contests, etc.) so that the user can travel more sustainably.

Apps that collect bicycling data

These apps concentrate on capturing data about bicycle trips. There may be similar apps for other modes of transport, but we did not encounter any.

Here again the movement of the phone is interpreted as the movement of the individual. These apps are often somewhat simpler in design (compared to those that collect data for all modes of transport), and they usually require the user to turn the logging function on and off in order to collect data about the bicycle trip. The app provides information about trips[4] and their purposes,[5] and sometimes about the individual[6] who has the phone as well. The user is usually required to review the trip and actively "send in" data. Data collected from bicycle apps are usually linked to other services that promote cycling or cyclists (e.g., bicycle trip planners, reporting of problems on the bicycle path). This can affect behaviour, which may be problematic in a TS context.

Apps that measure physical activity

These apps concentrate on capturing physical activity. It matters little whether it is a trip (i.e., with an ultimate purpose) or a workout. The main purpose is to provide the user with feedback about the nature of their physical activity.

The movement of the telephone is here again interpreted as the movement of the individual, but it is also possible to collect data via various watches and armbands that are connected to the app. These apps usually detect automatically when one is running, jumping or cycling, while other movements are lumped together as "transport". Often, the purpose is not included in the data collected. There is some information about the person using the app, but usually this information pertains only to health-related aspects (e.g., age, gender, weight). Information about, for instance, income, education and number of cars in the household is not collected. These apps usually require that the user actively

[3] The data collected about the individual differ greatly from one app to another, and can also be designed based on the aims of a study.

[4] The origin, destination, trip length, route choice, start time, end time, and duration are collected for the stretch on which the logging function is active. The mode of transport is assumed to be a bicycle.

[5] Data collected with regard to purposes: type of purpose, geographical point, start time, end time, duration.

[6] The data collected about the individual differ greatly from one app to another, and can also be designed based on the aims of a study.

approve the movement pattern, but it is also possible to make corrections if the user sees something that is incorrect.

Online questionnaires that collect travel data

Online questionnaires as a method of collecting travel behaviour data are nothing new, but they have evolved with the advent of digitalisation. It is now possible to incorporate a significantly higher level of detail in online questionnaires than had previously been the case. For instance, respondents can now indicate on a map where they have travelled and fill out questionnaires on various devices (computers, tablets, smartphones). Digitalisation offers new means of creating “smarter” questionnaires that are tailored to the respondent, rather than the other way round. Because the survey is digital, recruitment can occur via various types of social media. Online questionnaires collect transport data and make it possible to track trips made by single individuals, including their purpose, the mode of transport and information about the traveller.

Ticketing and payment systems

There are a number of different ticketing and payment (TaP) systems that contain information about where someone has been, and sometimes information about how they travelled there as well. This type of data is currently available from public transport systems, bicycle-sharing systems, carpools, and, to some extent, credit card use. It is also conceivable that Mobility as a Service systems will have this type of information in the future.

<u>Public transport</u>
It is possible to derive a great deal of information from the public transport system, depending on how ticket validation is carried out (which is different in different regions). A ticket is normally scanned (single ticket or time-based travel pass) when, for example, one boards a vehicle or walks down the platform. It provides information about where and when a public transport trip begins. In some systems the same ticket is scanned again when one leaves the vehicle, and in those cases it can also provide information about when and where the public transport trip ended. Combined with knowledge of the conformation of the public transport system (lines and timetables), this makes it possible to form assumptions about how someone has travelled in the system. However, in most cases tickets are not scanned when a passenger leaves the vehicle. Major statistical loss is thus associated with scanning at the point of departure, particularly when the price is not affected.

Data from TaP systems in public transport can only provide information about public transport travel, and usually there is no information about the user. Time-based travel passes sometimes contain a small amount of information about the person who purchased/subscribed to the pass; however, because it is common to share such passes within a household or workplace, for instance, the relevance of such data is doubtful.

<u>Bicycle-sharing systems</u>
The bicycle-sharing systems that are present in many Swedish cities can provide information about how their bicycles are travelling within the transport system. Because it is very rare for more than one person to ride these bicycles, the transport and traffic mileage can be compared in these cases, and both types

of data can be collected. In the vast majority of cases the bicycle-sharing system is based on membership, which makes it possible to collect data about the rider. It is also possible to obtain data about one-time users via credit card data, although that is rarer. Using the ticketing system, it is further possible to link a given membership to where the user picked up and dropped off a bicycle. The length of time that a bicycle can be borrowed is often limited, and if it is short (e.g., 30 minutes) the bicycle will be used for one trip and then left behind, thus providing information about the origin and destination of the trip. If the time permitted is longer, cyclists who have their own bike locks may use the bicycle for multiple trips, in which case only information about where the first trip started and the last one ended will be available. Nor are the precise origin and destination of the trip identical with the locations of the bicycle racks, which lowers the accuracy of the origin and destination data. Normally it is also difficult to determine the purpose of the trip, as the user does not enter information about each individual trip, but rather the membership card is used to check out a bicycle and for debiting purposes.

The bicycles in some systems are equipped with GPS, which makes it possible to track where in the system the bicycles are travelling (route choice). The trips that are made using a bicycle-sharing system are probably not representative of all the bicycle trips in the area, as it is likely that certain types of persons tend to use the system and for certain types of trips. The placement of the bicycle stations also affects the trips that are made in the bicycle-sharing system. However, such trips are a subset of all the bicycle trips that are made.

Carpools

Carpool systems are also based on membership, and thus can also provide information about the individual (and potentially the household) that is a member of the carpool. It is also possible to derive data about when and where the car was picked up and dropped off, and how many kilometres it travelled. However, it is more difficult to determine trip times, as it is commonplace to borrow a car for a longer time than the actual travel takes (e.g., to visit an acquaintance, spend time there and then drive home). The purposes of the trips are also difficult to capture in such data.

All the data collected are traffic data, in other words the number of kilometres the car travels. It is difficult to derive information about how many people took part in the trip. The vehicle mileage driven in a carpool system is probably not representative of all the car trips in the area, as it is likely that certain types of persons tend to use carpools, and that the cars are used for certain types of trips. However, such trips are a subset of all the car trips that are made.

Credit cards

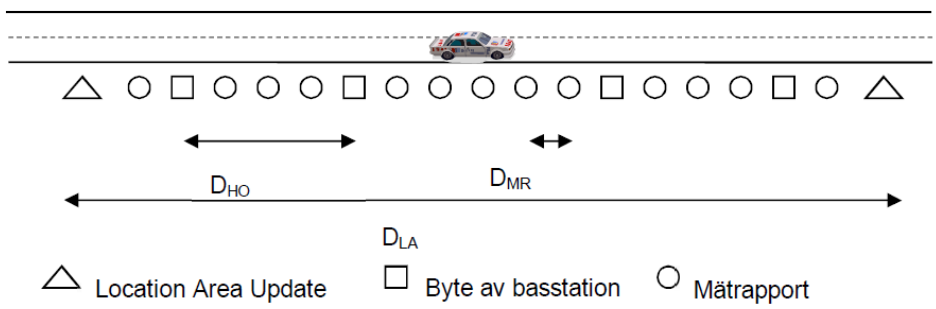
Information about where and when people use their credit cards is collected via the cards, including both the payee and the location. Because credit cards are seldom (but occasionally) loaned out, it is possible to know where a person was at certain times. If this information is to be usable, the credit card companies need to release personal information about their customers, something that it seems unlikely they would be willing/able to do.

Data from credit cards provide no information about how the cardholder travelled to the place where they are shopping (unless the credit card was used to pay for the trip) or how long they stayed. No information is available about trips to locations where the credit card was not used.

Mobile network data

There are three different ways of collecting position information via mobile network data: 1) when the user uses the telephone (SMS, surfing, conversation, etc.), 2) when the telephone comes into contact with another mast and 3) at set time intervals. These different ways offer different levels of precision.

The most detailed positioning information, both spatially and chronologically, is generated when the telephone is being used actively. Positioning by means of such data offers precision of several metres up to several hundred metres. Phones that are not being used actively can be positioned when they change Location Area (LA) (a larger geographical area consisting of a number of base stations/cells), yielding a lower level of precision, that is between several hundred metres and several kilometres (several tens of kilometres in rural towns). In addition, the phone will transmit (when it is not being used actively) signalling information at regular intervals. This interval normally spans several hours and affords positioning precision of several kilometres to several tens of kilometres (see Figure 7).



Figur 7. Avstånd mellan mätrapporter med lokaliseringsinformation i GSM. D_{LA} är avståndet mellan uppdateringar för LA, storlek från flera kilometer upp till flera tiotals kilometer. D_{HO} är avståndet mellan byte av basstation, storleksordning från flera hundra meter upp till flera kilometer (flera tiotals km på landsbygden). D_{MR} är avståndet mellan mätrapporter, storleksordning från flera meter upp till flera hundra meter.

Figure 7 *Distance between measurement reports with localisation information in GSM. D_{LA} is the distance between updates for Location Area, on the order of several kilometres up to several tens of kilometres. D_{HO} is the distance between base station changes, on the order of several hundred metres up to several kilometres (several tens of km in rural areas). D_{MR} is the distance between measurement reports, on the order of several metres up to several hundred metres.*

Key: Byte av basstation = Base station change Mätrapport – Measurement report

Depending on the information transmitted by the mobile phone, one can either use the mast position to estimate where the phone is or triangulate to determine a more realistic position for the phone. Another more advanced method, known as Timing Advance, calculates the round-trip time for the communication with the mobile phone and the three connected cells to enable more precise location determination. In rural towns there is usually not enough information to estimate the position of the mobile phone.

Mobile network data can provide information about how phones are moving, in other words an approximation of transport data. Collected data can provide information about primary trips and base sites (i.e., locations where people spend long times such as the home and workplace). The times when trips start/end are approximate, and depend upon the types of data that are available. Data may also be available between the various start/end points, but the positioning is not precise.

It is difficult (but not impossible if other data are combined) to use mobile network data to achieve a degree of geographical detail that makes it possible to track detailed route choices in a city. On the other hand, it may be possible to identify principle routes of travel in connection with longer trips. Knowing the route of travel may make it possible to draw certain conclusions about the mode of transport, for example if the position points follow a road or railroad. However, it is more difficult to derive information if the trip is made along a road by car or bus.

Information about the purpose of the trip and about the individual is difficult to obtain, as there is no interaction with the person providing the data. If information about the user is sought, it is necessary to assume that the person using the telephone is the one who is registered as its user, and the telephone operator needs to disclose personal information about its customers. Even if this were feasible, a large proportion of phones are used by people other than those who are registered for the subscription (e.g., company phones), providing either no information about the user or misleading information.

Social media

Data are collected via social media primarily through data mining, that is using algorithms to analyse large amounts of data to answer various questions. Postings on a number of data sources (social media channels) contain geographic positions, sometimes GPS coordinates, sometimes at the city level. Data can be used, for instance, to comprehend travel patterns at the aggregate level, qualitative experiences in traffic (e.g., complaints with regard to public transport or "likes" on a bicycle or in a car), and travels to and from a major event. The data may be detailed but they are also skewed, in other words non-representative, and depend on who is using a given platform or on who uses social media at all.

Equipment/measurements on vehicles

Vehicles can be equipped with sensors and/or other measuring devices to collect information about vehicles moving in the transport system. We found two different types of onboard data collection: Floating Vehicle Data (FVD), where the movements of the vehicle are tracked; and passenger counts, where the number of passengers on board the vehicle is counted.

Floating Vehicle Data

FVD are based on the collection of data from vehicles moving in the transport system. Vehicle-to-Infrastructure (V2I) or Vehicle-to-Everything (V2X) can also be used to describe the interactions between vehicles and infrastructure or between vehicles and other entities. If there is a need to choose a vehicle route or monitor idling or eco-driving, for example, then equipment is connected to the vehicle and the driver can be given their own screen and be connected to a control centre in real time. There is also equipment that is placed in vehicles, such as GPS devices in cars, buses or on bicycles, that can deliver data to recipients in the same way. The vehicle can also have onboard WiFi or radio-frequency identification equipment, enabling it to make itself known to the surrounding infrastructure. Depending on what the information is to be used for, it is also possible to choose whether real-time data or statistical data are needed. Real-time data are needed if the application is a travel-time system or fleet management. Only statistical data are needed with regard to TS data, which means that it is possible to upload data via WiFi, for instance, once the trip has been completed.

This type of data can provide information like how the vehicle is moving in the traffic system, when and where the vehicle's trip starts/ends, and times and passenger counts based on how many seatbelts are in use. This is not necessarily identical with the individual's trip, but may comprise a trip element in a trip chain. Nor do FVD provide any information about the purpose of a trip. What can be done is to attempt to guess the purpose based on known locations, in other words Point of Interest. More data could potentially be obtained if there is a navigation system from which to determine the chosen destination.

Nor is there much information about the individuals in/on the vehicles. Data could potentially be linked to various registers (vehicle registers, bicycle-sharing system membership, insurance policy holders) to enable assumptions about who is in the vehicle. However, if that is feasible then the estimates will be extremely uncertain because, for example, the person driving a car is not always the person who owns it.

Passenger counts

Passenger counts in public transport are conducted either manually or automatically (Automatic Passenger Counting). Manual counts normally yield information about how many passengers are boarding a vehicle. The aim is often to determine the ridership on the public transport vehicle and/or the times that people choose to travel on a given line. Little knowledge is generated with regard to the conformation of the individual's travel, as no information is available as to when/where one boarded, nor is there any information about the person (age, gender, etc.).

Automatic counts can be made in the same way as manual ones, that is the number of passengers boarding is counted. There are systems for which new sensors are installed, and systems that use existing surveillance cameras are in the pipeline. The system can also be supplemented by counting how many people depart the vehicle. There is also a camera system that detects when an individual gets on/off, thus enabling enhanced knowledge of the individual's trip (even though it is limited to the trip element involving public transport). In the future, it may be possible to collect estimated age and gender data for individuals captured on camera, but such technology has not been implemented at present.

Roadside devices

Measuring devices deployed along the roadside can collect data about the people who pass by. This can be done with no connections between different sites (individual interfaces/sites) or by combining data from multiple sites (recognition).

Motor vehicle flows are traditionally measured using measuring hoses, which provide information about traffic volumes, timepoints, vehicle types, and, in some cases, velocity. Because it is the vehicles that are being measured, no information is available about how many people are in the vehicle or about the individuals who are travelling. New ways of doing so using camera systems (image analysis) and mobile phone identification have emerged in recent years.

Camera systems

Some camera systems scan vehicle registration numbers, which, when supplemented with the vehicle register, provide good information about the vehicle as well as who owns it (not necessarily the same person who is driving it). Other camera systems can detect different types of traffic and, using image analysis, recognise pedestrians, cyclists, and light and heavy motor vehicles. With the more advanced image analysis techniques it is also possible to track objects in the images, thereby gaining information about how road users are moving at, for example, an intersection.

On the other hand, when using such systems it is not possible to determine how people travel before and after they pass the relevant location, the purpose of the trip, or information about the individual travellers. The data are collected at the vehicle level (i.e., traffic data), and are often identical with transport data for pedestrians and cyclists.

There are camera systems that recognise images when they appear in another camera positioned at another location. This pertains mainly to cameras that automatically scan vehicle registration numbers. Recognising registration numbers at another point in the traffic system makes it possible to obtain information about how vehicles are moving in the traffic system. However, large numbers of cameras are necessary for high precision and to cover large areas.

WiFi/Bluetooth

Systems that measure how many mobile phones pass by a given location can also be used instead of traditional measuring hoses. The mobile phone's WiFi and/or Bluetooth signals are captured, making it possible to count the number of passing mobile phones. This does not mean that the vehicle traffic can be measured, but rather only individuals with mobile phones from which WiFi and/or Bluetooth can be actively measured.

There are also systems that set up WiFi networks (WiFi base stations) in a desired environment, making it possible to monitor how phones are moving, in other words capturing travel patterns (place and time). It appears that it is also possible to identify road user groups (pedestrians, bicycles, cars, etc.), although it is unclear just how this is done. Information about the purpose of a trip can be obtained at an aggregate level by analysing where the collected flows have gone. Information about individuals is lacking, although the collected data can be linked to demographic data by using other sources. All phones within an area are measured, which can result in both over- and underestimation of trips. The loss (individuals without mobile phones) is, however, not skewed, and the estimates are relatively good.

3.2 Current state of knowledge per the scientific literature

Digitalisation is making new methods for collecting data possible, while the Internet of Things (IoT) (the so-called fourth wave of digitalisation) offers many new ways of understanding movement patterns, travel behaviour and how the transport system is being used (Davidsson et al. 2017). The vast majority of data sources are based on data linked to mobile phones (Wang et al. 2017), although there are also new data sources such as TaP systems (e.g., via payment systems for public transport (Poonawala et al. 2016) or credit card data (Sobolevsky et al. 2014)), data from social media (e.g., Rashidi et al. 2017), or from special equipment in cars or carried by individuals (mainly GPS (e.g., Bliemer et al. 2010)).

A review of the various data collection techniques is provided below, along with examples from the scientific literature. The method used to identify relevant research literature is described above in Section 2.2. Four categories of data sources emerge from the literature:

1. Data linked to mobile phones:
    a. Active mobile telephone data: via sensors (accelerometers, GPS, gyroscopes, etc.) in mobile phones.
    b. Passive mobile phone data: via radio frequency data (Bluetooth, mobile network data, WiFi, etc.).
2. Data linked to TaP services.
3. Data from data mining social media.
4. Data from special equipment in vehicles or carried by individuals.

All the data collection methods make use of connected entities — IoT. A brief summary of the various data collection techniques is presented in Table 3-1. The literature with regard to the various data sources is described in greater detail below.

Table 3-1 Summary of new data collection techniques that use IoT according to the scientific literature.

Type of data	Brief description	Sample size	Portion of population	Area of application
Active mobile tele-phone data	App-based data collection using sensors in mobile phones (GPS, accelerometers, gyroscopes)	Small	People who download and use the app (and have it with them when they travel)	Determining the individ-ual's travel pattern (length, time, mode of transport, origin/destination, route (if GPS is used))
Passive mobile phone data	Data on how mobile phones are moving via mobile network data, WiFi, Blue-tooth	Large - Big data	Monitors mobile phones that are connected to a network (mobile network or have WiFi/Bluetooth on)	Primarily determining points between which individuals are travel-ling, estimating OD matrices
TaP	Payment and ticketing data from public transport or credit cards	Large – Big data	Individuals who use such cards	Mainly determining destinations of individu-als' public transport trips; estimating OD matrices
Social media	Data mining social media to determine travel patterns	Large – Big data	Users of social media	Mainly determining destinations of individu-als' trips; estimating OD matrices. Can also provide qualitative data of trips
Special equip-ment	Vehicles or indi-viduals are equipped with units that track their movements (primarily via GPS)	Small	People and vehicles equipped with such units	Determining individual's or vehicle's travel pat-tern (length, time, origin/destination, route (if GPS is used)

Data linked to mobile phones

Many of the new data sources for determining travel patterns are linked to data from today's mobile phones, either via data from sensors in mobile phones, links to satellites, or radio frequency data (e.g., Bluetooth, mobile network data, WiFi) (Figure 3-2). Data from social media and TaP systems are often linked to mobile phones, but these are handled separately, as they are not always linked to mobile phones.

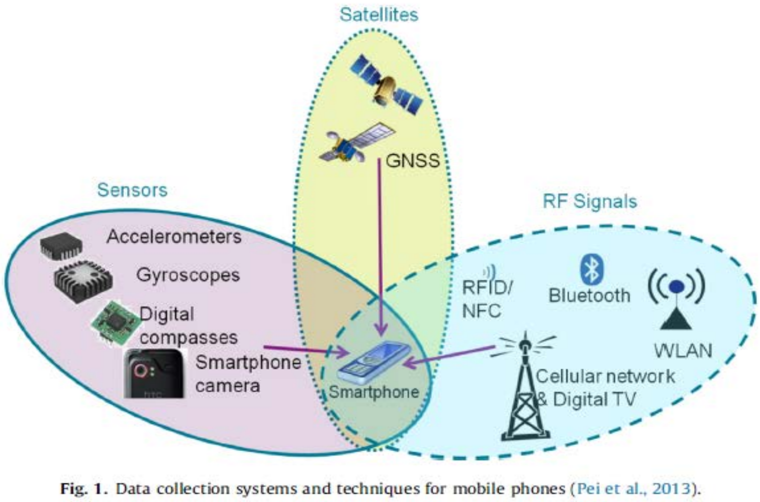


Figure 3-2 Data collection techniques using mobile phones. Source: Pei et al. 2013, in Wang et al. 2017.

Sensors in mobile phones – active data collection

Sensors in mobile phones include GPS, accelerometers, gyroscopes, and compasses. Data from such sensors enable the identification of the trips and activities of individuals who have a mobile phone with them. Apps can be installed in the phones to share in the data from sensors. There are apps that collect travel data in mobile phones in two different ways: totally passively, where the user does not need to be involved (other than to install the app); and semi-actively, where the user corrects data residing in the mobile phone, either in a web interface or directly in the app.

There are few examples of actual tests of working apps in the scientific literature. Most articles concentrate on technical possibilities and details of algorithms for correctly identifying modes of transport and/or activities/purposes. One example of an actual test is a newly published article from Australia (Safi et al. 2017) in which the ATLAS II app is compared with three other technologies: online questionnaires, simple GPS units and GPS collection from mobile phones without active input. The ATLAS II app uses algorithms to automatically identify the mode of transport, origin and destination by using GPS and accelerometers in the mobile phone. The user corrects data suggested by the app directly in the app. The result of the data collection process is a detailed travel diary with GPS tracks. The ATLAS II app delivered the best response rate of all the collection techniques, but it is worth noting that older respondents preferred online questionnaires over app-based methods, while younger ones preferred app-based methods, and that the response rate for apps was highest among the most highly educated. There were a total of 267 respondents, with the lowest number for ATLAS II (n = 73) and the highest for online questionnaires (n = 185). The respondents were recruited via panels and from people who had previously participated in TSs. The research team that developed ATLAS also describes the development of the app in Safi et al. (2013; 2015).

Petrunoff et al. (2013) also compare data collected from accelerometers with online questionnaires and hardcopy questionnaires, and show that there are few differences between these data collection techniques in terms of identifying modes of transport or trip origins and destinations. The study focused on active mobility, in other words physical activity during trips.

GPS has been used for many years to determine travel patterns. Data from GPS, initially from standalone GPS devices and subsequently in mobile phones, have been researched by many people in many parts of the world. Some examples include Japan (Asakura et al. 2014), Singapore (Kim et al. 2014) and Scotland (SiłaNowicka et al. 2016).

GPS via mobile phones has also been used as a means of correcting data from traditional collection methods (Wolf et al. 2013), while Inbakaran (2014) believes that GPS via mobile phones could complement traditional data collection techniques, as multiple collection methods (mixed methods) can yield a higher response rate from various groups of respondents.

GPS data from mobile phones are sometimes combined with data from other sensors (mainly accelerometers) in order to determine the mode of transport automatically (Shen & Stopher 2013; Assemi et al. 2016; Shafique & Hato 2016; Berger & Platzer 2015), or to automatically determine activities (Kim et al. 2014; Cottrill et al. 2013). The accuracy in terms of finding the correct mode of transport is generally around 80%. There are a number of examples of higher accuracy, but they are based on small numbers of people/trips (Lee et al. 2016). The accuracy depends a great deal on the mode of transport, as it is easier, for instance, to identify walking trips than to distinguish between bus and bicycle trips in an urban environment. It is more difficult to automatically determine activities other than those occurring at base sites (home, work) due to the heavy mixing of activities in geographical proximity in certain areas.

In addition to using sensors in mobile phones to determine travel patterns, mobile phones can also be used to collect data by creating new direct contacts, for example via SMS (Gould 2013).

To be able to collect data via an app, the app needs to be downloaded by a person. Unfortunately, there is very little knowledge concerning the recruiting methods that can be used to reach various respondents and induce them to download and (if necessary) interact with the app/web interface.

One report from 2017 describes experiences from Finland (Pastinen et al. 2017). A pilot study was conducted in connection with the most recent Finnish TS (2016), in which the respondents had to collect travel data using an app in their mobile phone or GPS. A random sample of 1,000 people was contacted by phone, 171 of who were recruited into the pilot study, while 195 declined to participate. The bulk of the sample could not be reached by phone. Of the 171 recruits, 45 used the app or GPS unit on the proper measurement day. The experiences from the pilot study indicated that GPS worked best, as there were technical problems with the app (an early beta version for iOS, while the Android users had problems using the app or were unwilling to use it). The GPS must in turn be simple and easy to use, with no buttons other than on/off. The experiences showed that it is not yet realistic to obtain reliable data that reflect the travel behaviour of the Finnish population using apps or GPS. Another conclusion was that very few people would be willing to use the mobile phone app without an incentive of roughly €50, even if the technical problems were to be solved. Other experiences indicated that it would take a great deal of work (algorithm development) to obtain the reliable data on modes of transport and purposes of trips that the national TS requires. The conclusion was that apps or GPS can be used at present as a supplementary method for certain subgroups of the population and for certain survey purposes, for instance to estimate the length of walks or check the participants' ability to estimate trip lengths, timepoints or numbers of trips. Traditional methods such as online/mobile phone questionnaires, telephone interviews and hardcopy travel diaries will be needed until further notice to obtain basic information about modes of transport and numbers of trips. The report also noted that errors may occur if the app or GPS unit is not running (e.g., if the app has not been turned on, the operating system has been shut down or the battery is dead). It was emphasised that it is important in terms of data quality that algorithms for processing raw data consist of open source code.

Mobile network data – passive data collection

There are many ways of deriving benefits from mobile phones via passive data collection methods. Lee et al. (2016) studied how mobile network data, GPS, apps in mobile phones, and Bluetooth can be used as a means of estimating TS data. The conclusions from their article indicate that there are many possibilities, but that skewness in the data may be a problem, even as socio-demographic variables have to be linked to the data in some way. In their view it is important to study, first and foremost, data representativity in future projects, that standardisation of data-cleansing methods is incorporated, and that more practical applications are studied. They also identify a lack of knowledge as an impediment to the use of new data sources. Issues concerning insufficient knowledge and the need for cooperation between traffic and transport experts and computer sciences experts are also addressed in an article from 2017 (Anda et al. 2017). This article again points out that representativity is a key issue that needs to be researched, and the authors state how important traditional data sources (“old-fashioned” TS) are in developing methods and conceptual frameworks.

Jiang (2016) used mobile network data together with WiFi data, household data and traditional TS data to gain an understanding of mobility patterns in Singapore. Traditional TS data have also been combined with mobile network data in Boston (Diao et al. 2016) and other cities in the USA (Alexander et al. 2015). Mobile network data have also been used to determine travel patterns in two major cities in China (Yang et al. 2016), and in the Ivory Coast to assess public transport use (Di Lorenzo et al. 2016). In the USA, AT&T Labs have used Call Data Record data in three major cities to determine travel patterns and derive “carbon footprints” for the cities (Becker et al. 2013), while data from Ford have been used together with mobile network data to gain an understanding of travel patterns in Boston (Mearian 2016). Mobile positioning can also be combined with GPS data in phones and reduce the need for the continuous collection of GPS points. This could enable the development of apps that use less energy, which is a problem for apps that collect GPS data (Wang et al. 2012).

The scientific literature contains fewer examples of the use of mobile network data in Europe, perhaps due to its stricter personal privacy regulations. One example, from Italy, is a project involving data from Vodafone Italy (Tosi et al.

2016). The project focused on identifying trips around underground stations and other Point of Interest (POI), and a means of estimating origin-destination (OD) matrices.

As is the case with active collection methods that use sensors from mobile phones, there are representativity problems with mobile network data as well. This is because mobile network data offer information about the travel patterns of mobile phones connected to networks from a given company and not from individuals. Several of the aforementioned studies use multiple data sources in combination, while others use existing TS data as a supplement to mobile network data in attempting to derive an overall picture of travel behaviour. According to one study from China, mobile network data underestimate the actual number of trips (Lu et al. 2017). According to Bricka et al. (2014), passive data collection methods need to be combined with active and traditional ones. Technology bias, in other words the fact that different technologies attract different types of people, must be studied more closely.

Mobile network data are used mainly in applications for estimating OD matrices. Here there is a problem in that it is difficult to determine modes of transport from mobile network data, and TS data broken down by mode of transport are then used. Mobile network data provide large data sets, while app-collected data entail significantly smaller amounts of data. Data from sensors (apps) in mobile phones are used primarily to track individuals' trips and estimate different modes of transport, while mobile network data can offer an aggregate picture of movements in an area. A summary of some examples is presented in Figure 3-3.

Selected studies applying different sources of mobile phone data for various research purposes.

Data source	Data collection techniques	Dataset description	Research applications	Research methods	References
Cellular network-based data	Triangulation positioning	7989 individuals	Identify activity and types of locations visited	Clustering method, logistic regression model, behaviour-based algorithms	Chen et al. (2014)
Cellular network-based data	Cell-ID positioning	2.87 million users over a month	Develop OD matrices	Frequency method, Scaling method	Iqbal et al. (2014)
Cellular network-based data	Triangulation positioning	No details provided	Identify the temporal and spatial attributes of citizens' trips.	Clustering method	Li et al. (2016)
Cellular network-based data	Cell-ID positioning	4 million samples for12 days	Develop OD matrix	Frequency method, clustering method	Patrick et al. (2015)
Cellular network-based data	Rio de Janeiro: Cell-ID positioning Boston: Triangulation positioning	Boston: 2 million users for 2 months Rio de Janeiro: 2.8 million users for 5 months	Perform a four-step model, generate OD matrices and identify trip purposes	Rule-based classification algorithm	Çolak et al. (2015)
Sensor-based data	Positioning: GPS and WiFi; Motion: accelerometers, gyroscope, magnetometers	560 samples for verifying the positioning accuracy, 1200 samples for motion recognition	Capture human movements, including positioning, motion recognition, and human behaviour modelling.	Bayesian theorem, Histogram maximum likelihood algorithm, Sequential forward selection algorithm, LS-SVM classification algorithms, LoMoCo model	Pei et al. (2013)
Sensor-based data	GPS positioning	658 trips and 457,945 points	Explore methods to improve travel modes inference performance.	Classification models: Naïve Bayes K-Nearest Neighbour; Discriminate analysis, PCA	Nour et al. (2016)
Sensor-based data	Sensor fusion: preprocessed data from the social platform of sports tracking app.	36,757workouts by 2424 users for 31 months.	Derive privacy-preserving heat maps of cycling workouts.	Privacy-preserving kernel density estimation and user count calculation, k-Anonymity model	Oksanen et al. (2015)
Sensor-based data	Positioning: GPS, WiFi, GSM, Bluetooth; Motion: accelerometers	Over 18,000 trips of 600 participators for 2 weeks	Conduct a smartphone-based panel survey to automatically identify trip details.	Bayesian probability statistics	Geurs et al. (2015)
Sensor-based data, cellular network-based data	Positioning: GPS, triangulation positioning Motion: accelerometers, magnetic sensors	15 volunteers for 2 months	Propose a classification technique to automatically detect travel modes.	Randomized ensemble of classifiers with a Hidden Markov Model	Widhalm et al. (2012)
Sensor-based data, cellular network-based data	Positioning: GPS, cellular network-based positioning; Motion: accelerometers, magnetic sensors	15 volunteers for 2 months	Detect trip legs and concerned 8 types of travel modes	Probabilistic classifiers, discrete hidden Markov model	Nitsche et al. (2014)

Figure 3-3 Some examples of how data linked to mobile phone are used to determine travel patterns. Source: Wang et al. 2016.

Ticketing and Payment data

There also always ways of using TaP data such as payment cards for public transport (smartcards), data from ticket purchases and data from credit card purchases. Here again it is commonplace to combine such data with other data sources in order to understand how people are travelling. Public transport data have been used to estimate OD matrices, and validated with the help of traditional TS data (Nassir et al. 2015). Data from public transport passes (smartcards) have been used together with mobile network data (from StarHub) to estimate the number of public transport trips and OD matrices for public transport in Singapore (Poonawala et al. 2016; Holleczek et al. 2013). Data from public transport passes combined with Automatic Vehicle Location (AVL) have been studied to obtain a better understanding of where people travel, information that is not normally found using traditional data collection techniques (Shin & Lin 2013). Data from public transport payment cards can also be used as a means of determining the traffic generated by various destinations (Hasan et al. 2013). Credit card data represent another source that can provide knowledge about travel patterns (Sobolevsky et al. 2014).

Social media

Data from social media can play a role in understanding activities, and can be combined with other data sources to achieve an understanding of travel patterns at an aggregate level (Abbasi et al. 2015, Chua et al. 2016, Chiang et al, 2014; Salas-Omeda et al. 2016; Wang et al. 2014). A tool has also been developed in the USA that integrates Twitter data with GIS layers (Soltani et al. 2016). The languages used in social media can also be studied to determine movement patterns for various cultural groups (Wu et al. 2016). Data from social media have also been studied together with mobile network data to understand travel patterns and changes in travel patterns over time (Chen 2017).

Furthermore, data from social media have been used to understand travel patterns associated with major events (Cesario et al. 2016; Wang et al. 2016). This can also be supplemented with GPS data (Vlassenroot et al. 2015).

Many researchers believe in the feasibility of using data from social media to understand travel behaviour in the future, such as determining purposes of trips, times, problems in the transport system, modes of transport, and even socio-demographic variables (see Figure 3-4). The advantage of using social media is that the data are inexpensive and readily accessible, although there are problems in terms of representativity and the fact that, although the data are inexpensive, they still require processing before being used for planning purposes. Such processing requires specialised expertise, which can be expensive (Rashidi et al. 2017).

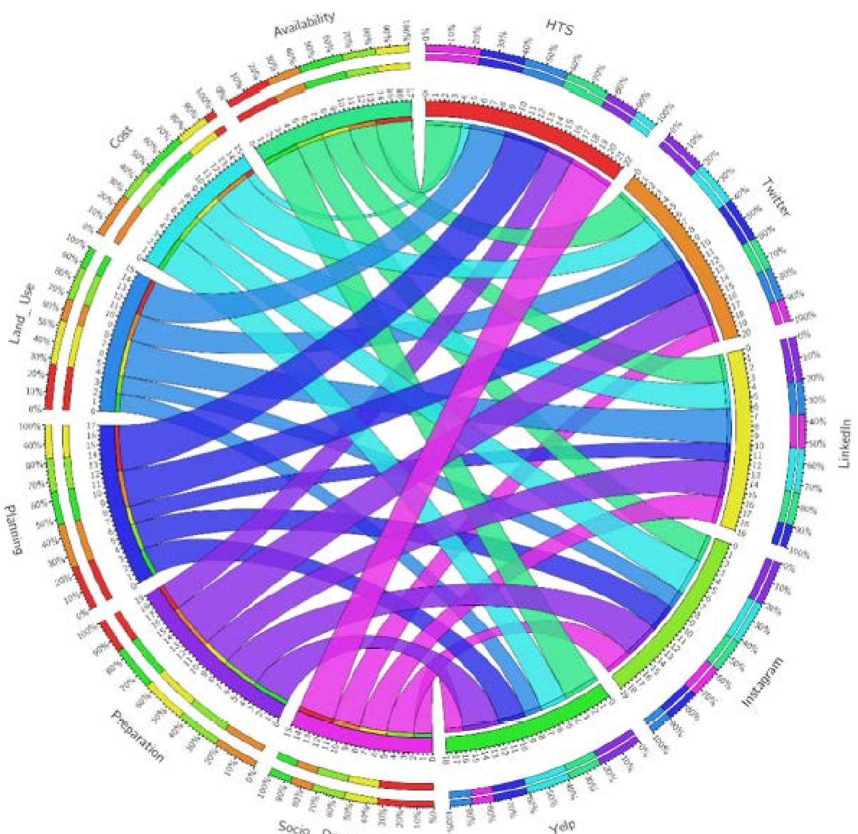


Figure 3-4 Various areas of application (to left) for data from social media (to right). Source: Rashidi et al. 2017.

Special equipment on vehicles/individuals

Vehicles and individuals can be equipped (or carry with them) special equipment that collects data. Before mobile phones with GPS became generally available, collecting position data was commonplace, and there are many examples in the scientific literature. Testing in this area is still ongoing, but it may be difficult to find respondents for this sort of data collection; the response rate in one test in 2015 in Scotland was less than 4% (SilaNowicka et al. 2015).

It is more common to equip vehicles with computers that collect data. In addition to data from automakers, there are examples of aftermarket equipment for collecting GPS data, for example in cooperation with insurance companies in Italy (De Gennaro et al. 2016) and, in one test, an attempt was made to get people to drive outside of peak traffic times in the Netherlands (*Spitsmijden* – Bliemer et al. 2010).

As with other data collection methods, this type of data is combined with different data sources to derive an overall picture. For example, vehicle identification data (AVL) have been used together with data from a travel behaviour app and public transport timetables to determine how many passengers are present in various parts of the public transport network, and how well adapted the timetables are (Carrel et al. 2015).

Summary

The scientific literature offers many different examples of how IoT data have been used to gain an understanding of travel behaviour. Data from various sources are usually combined in order to obtain a better picture of the transport or traffic activity in its entirety, and all data sources suffer from skewness and sample-related problems. However, our knowledge of, in particular, sample-related problems and representativity is very limited.

Most of the articles shed light on research that does not result in the development of tools, but rather concerns general methods for the ways in which data can be collected and used by various actors. This is due in part to our literature review having focused on the scientific literature, where the emphasis is usually not on developing usable tools, but also in part to the fact that these data collection methods are relatively new and the research reporting process is a slow one, which means that there is reason to believe that the results have not yet had time to make it into print.

3.3 Important questions for travel surveys

This section will address the areas of application for data from TSs. It is based on results from the parallel work package *Stakeholder requirements from travel surveys*, which was led by the Swedish National Road and Transport Research Institute (VTI) (Eriksson et al. 2017).

Based on interviews, questionnaires and a compilation of our own experiences, requirements and desires emerged with respect to the following types of information:

- Choice of mode of transport
- Purpose of the trip
- Time for the trip (start and end time)

Travel time (duration)

Geographical location (origin and destination)

Travel distances

Who is travelling (gender, age, income, occupation, driving licence, etc.)

There was also interest in knowing more about:

Vehicle type (e.g., electric vehicle)

Choice of route

Trip chains – to understand the entire trip, not just trip elements and sub-trips

Changes over time (addressed in Chapter 3.5)

It is important to note that the areas of application that have emerged are tinged by the data that are currently available. For example, traditional survey methods offer little information about where people travel (choice of route), and users have adapted their questions accordingly. Consequently, it is important to think in relatively broad terms when it comes to the areas of application and focus on the primary purpose of a TS, so that the possibilities opened up by new types of data can also be incorporated.

The ways in which different data collection methods are able to answer the various foregoing questions are compiled in Table 3-2.

The table includes information about the ways in which each individual data collection method answers the various questions. However, it should be emphasised that the ability to answer different questions statistically requires different amounts of data and different levels of data quality. This means that questions may differ in terms of how easy/hard they are to answer, even if it is possible to answer them using a given data collection method.

It should also be emphasised that a number of sources can be used simultaneously to obtain a broader picture of the travel patterns as a whole. One example is where roadside devices that provide a good representation of the numbers of individuals or vehicles that are passing along a given section can be used together with FVD, which provide information about routes, origins and destinations for trips — that is different sources that cover different parts of the traffic system can be used in conjunction. If a row in the table is mostly red, that does not necessarily mean that the data cannot be used for TS purposes, but rather that the data need to be supplemented with data from other sources. It is, however, difficult to know just how different data sources can be used together without practical experience.

Interpreting the colours in the table

TS Sweden and online questionnaires

TS Sweden and online questionnaires offer all of the information sought, except that TS Sweden does not include geographical routes. At present, the survey contains no information about vehicle type, but such a question could be posed in the questionnaire.

Apps

Apps offer data that answer all the questions. One difference between bicycle and exercise apps compared to TS apps is that the former offer data concentrated mainly on bicycle and/or walking trips, and do not include a full perspective in terms of transport. Information as to who travelled or the purpose of a trip is seldom available from bicycle or exercise apps.

Ticketing and Payment

The purposes of these modes of transport can be linked to POI data to derive an understanding of the purposes of trips. It is possible to gain an understanding of where people are travelling using public transport, but only if they scan the card at the end of the trip (which is not customary in Sweden). Data from carpool vehicle TaP show only when the car was picked up and dropped off, as trip element/purpose data are not generated. However, information about how far one has travelled with the car is included, as the billing charge is usually based on both trip length and travel time, as opposed to public transport and bicycles, where the charge is based solely on time. Trip times can thus be obtained for shared bicycles and public transport. In the case of a carpool vehicle, usually only the total travel distance and trip time for a chain of trips can be obtained, as the car is often kept for the time during which the purpose of the trip is being fulfilled. Information as to who is travelling can be obtained, but it is usually possible for such a card to be loaned out, which then provides misleading information. Information about trip chains is not included.

Mobile network data

There is a great deal of uncertainty surrounding mobile network data, in other words what is possible and what is not. Most of the questions that mobile network data can answer are marked in blue, as it is uncertain where the research will lead with regard to these questions. The data will, for the most part, be rather coarse, with details about base sites (i.e., home and work) and details about lengthy trips. There can also be data about individuals, but it is uncertain whether such data will be available for analysis in Europe.

Floating Vehicle Data

FVD offer excellent information about what vehicle is being used, as well as geographical and chronological information about a trip. Because FVD are traffic data and not transport data, there is little information about who is travelling or about the purposes of trips. Traffic and transport data are identical for bicycle sharing systems, as only one person rides the bicycle. This makes it possible to obtain a little information about who is cycling if the data are linked to a membership card. Information about the purposes of trips may also be available for taxis, cars and bicycle-sharing systems if POI data are used as well. On the other hand, in the case of public transport, FVD track vehicles and not passengers, so that no information about the purposes of trips is available.

Passenger counts

Systems that recognise the same individuals may provide data about public transport trips, but only at the trip element level. There is no information about how one has travelled, the purpose of the trip or trip chains.

Roadside devices

Roadside devices can deliver data about the mode of transport and vehicle information (if the plate number is linked to the vehicle database). It is, however, uncertain whether the correct mode of transport can be identified in most cases, and the measurements include some goods shipments. It is also difficult to distinguish between different modes of transport in section measurements involving sensors in mobile phones.

Table 3-2 Compilation showing which data collection methods (rows) answer which questions (columns). Green means that the method can answer the question. Yellow means that it can answer the question in part. Red means that it cannot answer the question. Blue indicates that it is uncertain whether the collection method can provide the relevant information (research questions). Other designations:
1 = usually not, but it is possible to do if so desired
2 = at the trip element level only
3 = but could be loaned out
4 = with help of POI
5 = long-distance trips
6 = base site (locations where people stay for long times, usually home and workplace)
7 = roughly as the crow flies
8 = general patterns
9 = link to member (not necessarily the person using)
10 = there are systems that recognise individuals and identify start/end
11 = certain goods shipments also included
12= if registration number is identified and run in the vehicle register
13= some information is available about subscriptions, unclear whether it will be used

Data collection method		**transport**	Vehicle information	Purpose	Geo-start	Geo-end	OD matrix	Geo-route	Trip length	Who	Start time	End time	**Duration**	Trip chains
TS Sweden			1											
App	App - TS data		1											
	App - bicycle trips		1							1				
	App – exercise		1							1				
Online questionnaires			1											
TaP	Public transport			4		2				3	2	2	2	
	Shares bicycle			4		2				3	2	2	2	
	Carpool									3	2			
	Credit card													
Mobile network data	Mast position-activity	5		6	6	6		5	7	13	8	8	8	
	Mast position-Mast change	5		6	6	6		5	7	13	8	8	8	
	Mast position-Time interval	5		6	6	6		5	7	13	8	8	8	
	Triangulation-activity	5		6	6	6		5	7	13	8	8	8	
	Triangulation-Mast change	5		6	6	6		5	7	13	8	8	8	
	Triangulation-Time interval	5		6	6	6		5	7	13	8	8	8	
Floating vehicle	Navigation system			4										
	Automaker			6										
	Special equipment – publ.					2	2	2	2		2	2	2	
	Special equipment – taxi			4		2	2	2	2		2	2	2	
	Special equipment – bicycle sharing system			4		2	2	2	2	9	2	2	2	
	Special equipment – car			4		2	2	2	2		2	2	2	
Passen-ger counts	Passenger counts – APC				10	10	10	10	10		10	10	10	
	Passenger counts– manual													
Roadside devices	Section/site – camera	11	12							12				
	Section/site- hose	11												
	Section/site – WiFi/Bluetooth													
	Recognition – camera	11	12							12				
	Recognition – WiFi/Bluetooth	11												

Generally speaking, it is clear that the sources that answer the most questions independently are those that were designed for a TS, in other words the traditional TS, apps designed to collect TS data and online questionnaires for collecting TS data. For example, only these sources can provide a picture of trip chains.

TaP systems can provide information about individual trip elements, but they need to be combined with other data to obtain a more comprehensive picture.

Mobile network data can provide information about how individuals are moving at a more comprehensive level, and are well suited for describing trips at an aggregate level, where links to individual characteristics and high geographical precision are not as important. The extent to which this type of data can capture shorter stops, such as for shopping and picking up/dropping off relatives, is a topic that needs to be studied more.

FVD provide detailed information about the trip itself, but need to be supplemented with information about who is making the trip and why it is being made (purpose). This type of data also provides only information about the relevant mode of transport and relevant user group, and needs to be supplemented with additional modes of transport/user groups to provide a more comprehensive picture.

Passenger counts can contribute information about ridership on public transport vehicles and thus serve as a link between traffic and transport data for public transport. Automatic systems may also come to provide information about trip elements involving public transport.

Roadside devices can contribute to an understanding of total volumes and play a part in scaling up other data. If a higher number of WiFi base stations are deployed, it will be possible to generate data that are similar to mobile network data, but more precise.

3.4 Quality aspects

According to Statistics Sweden's guidelines (SCB-FS 2016:17, see SCB (2016)), the quality concept with regard to public statistics comprises the following five main components:

- Relevance
- Reliability
- Timeliness and punctuality
- Availability and clarity
- Comparability and co-usability

Relevance refers to end purposes, information needs and the contents of the statistics, which were addressed in the preceding Chapter 3.3.

Reliability refers to accuracy or how well a statistical value estimates its target parameter, that is what one wants to measure. Reliability is affected by numerous sources of uncertainty, including sample, measurement, loss, and processing uncertainty.

Timeliness and punctuality have to do with how often recurrent statistics need to arrive in order for the most recent available statistics to be considered up-to-date.

Availability refers to where and in what way the statistics are reported.

Comparability has to do with the possibilities that exist in terms of drawing comparisons between different times or groups, while co-usability concerns the possibilities in terms of combining and analysing different statistical values from the same or different surveys. Uniformity among the definitions of the target parameters and survey methods is decisive in determining whether this is possible.

The quality aspects are considered in relation to restrictions such as the information burden on the respondents, cost, statistical confidentiality, and the handling of personal data.

Figure 3-6 shows the main components and secondary aspects of the quality concept.

● Relevance ○ Purpose and information needs - Purpose of the statistics - Statistics user's information needs ○ Contents of the statistics - Variables - Statistical metrics - Reporting groups - References times ● Reliability ○ Total reliability ○ Uncertain sources - Sample - Frame coverage - Measurement - Loss - Processing - Model assumptions ○ Preliminary statistics compared to final	● Timelines and punctuality ○ Preparation time ○ Frequency ○ Punctuality ● Availability and clarity ○ Access to the statistics ○ Possibility of additional statistics ○ Presentation ○ Documentation ● Comparability and co-usability ○ Comparability over time ○ Comparability between groups ○ Co-usability otherwise ○ Numerical agreement

Figure 3-6 Main components and secondary aspects of the quality concept. Source: Statistics Sweden. 2016. Quality of the official statistics – a handbook.

The quality handbook is prepared mainly for studies based on statistical samples of individuals, but it is applied below in Table 3-3 to assess the reliability of the various types of tools/methods that we mapped. Here we assess how well the various types usually function at present, while in the following chapters (4 and 5) we will describe the individual tools and methods, how far they have evolved, and what they offer in terms of development potential. The sources for this assessment comprise the international scientific literature in Chapter 3.2 and our review in Chapters 3.1 and 3.3 above as the basis for our mapping (Chapters 4 and 5), plus national literature such as Sweco (2013).

Table 3-3 Advantages and disadvantages in terms of reliability for various data collection methods.

Data collection method			Advantages in terms of reliability	Disadvantages in terms of reliability
TS Sweden			Clear target population, statistical sample, measures trips among population (6-84 years of age), telephone interviews make it possible to check responses, and reduce the risk of forgotten trips	Low response rate/skewed response group, uncertain information about trip length and travel time
App	App – TS data		GPS traceability means good measurements of number of trips, trip length and travel time. Corrections can be permitted	Only those with smartphone/computer access are included in the population, recruitment differs, loss often unknown and response group skewed, partial loss when batteries or connections are lacking
	App – bicycle trips		GPS traceability means good measurement of number of bicycle trips, their length, travel time, etc.	Only cyclists with smartphones included in the population, usually crowdsourcing (i.e., loss unknown), partial loss if user has to but forgets to turn the app on/off and when battery/mobile/connection are lacking
	App – exercise		GPS traceability means good measurement of "exercise trips", their length, travel time, etc.	Only exercisers with smartphones are included in the population, usually crowdsourcing, (i.e., loss unknown), partial loss if user has to but forgets to turn the app on/off and when battery/mobile/connection are lacking
Online questionnaires			Enable statistical sampling among those with a postal address, measures trips made by the population, possible to use a travel diary with maps to obtain origin and destination addresses	Only those with computer access/Internet are included, low response rate/skewed response group, uncertain information about trip length and travel time, risk of partial loss if it is a long online questionnaire
TaP	Public transport		Captures public transport users in relevant system, big data – yields absolute figures for public transport travel, little loss/measurement error	High partial loss with regard to end time if the card is not used upon departing the vehicle
	Shared bicycle		Captures bicycle sharing users in relevant system, big data – yields absolute figures for bicycle trips, little loss/measurement error	Captures a skewed subset of all bicycle trips
	Carpool		Captures carpool users in relevant system, big data – yields absolute figures for trips, little loss/measurement error	Captures a skewed subset of all car trips, measures vehicles and not individuals
	Credit cards		Captures credit card users in relevant system, big data	Only credit card users are included/skewed group, loss if one pays cash or purchases something
Mobile network data	Activity		Big data, that is yields absolute figures for movements.	Only individuals with mobile phones from the relevant operator are included, loss/measurement error among those with multiple mobile phones or who lack battery or connection, travel distance measured as the crow flies (possibly with intermediate points)
	Change of mast			
	Time interval			
Floating vehicle	Navigation system		Little measurement error/loss, big data	Vehicles with relevant system are probably few and not representative of all vehicles in terms of OD, measures vehicles and not individuals
	Automakers		Little measurement error/loss, big data	Vehicles (cars and public transport) with relevant system are probably few and not representative of all vehicles in terms of OD, measures vehicles and not individuals
	Special equipment	public	Little measurement error/loss, big data	Vehicles (public transport) with relevant system are probably few and not representative of all vehicles in terms of OD, measures vehicles and not individuals
		taxi	Pertains to vehicles with relevant system within taxi companies. Little measurement error/loss, sampling can be done via taxi companies	Vehicles with relevant system are probably few and not representative of all taxi vehicles, measures vehicles and not individuals
		Bicycle sharing system	Pertains to shared bicycles (i.e., bicycle trips) within relevant bicycle sharing system, little measurement error/loss	Bicycles within relevant system are not representative of bicycle trips in general
		car	Pertains to vehicles with relevant system, little measurement error/loss, recruitment can be done via insurance companies and other companies that map driving patterns	The equipment is intended to encourage a given behaviour; those who use it are not representative, and their behaviour will also be affected

Table 3-3 Advantages and disadvantages of various data collection methods (cont.)

<table>
<tr><th colspan="3">Data collection method</th><th>Advantages in terms of reliability</th><th>Disadvantages in terms of reliability</th></tr>
<tr><td colspan="2" rowspan="2">Passenger counts</td><td>APC</td><td>Pertains to public transport passengers in relevant vehicles, sampling done by public transport operator, little measurement error/loss</td><td>Does not measure all public transport passengers, there may be skewness in the sample of vehicles that use APC</td></tr>
<tr><td>manual</td><td>As above, assuming that one has sufficient resources for counting</td><td>Does not measure all public transport passengers, there may be skewness in the sample of vehicles that use APC</td></tr>
<tr><td rowspan="4">Roadside devices</td><td rowspan="3">Section/site</td><td>camera</td><td>Measures all vehicles or road users at relevant site, selection of site made by the entity responsible for the road, little measurement error/loss</td><td>Site selection decisive for any statistical weights, roughly 10% of reg. nos. missed</td></tr>
<tr><td>hose</td><td>As above, but simpler systems with no reg. nos., measuring cyclists requires the proper equipment</td><td>Site selection decisive for any statistical weights</td></tr>
<tr><td>WiFi/ Bluetooth</td><td>Measures all smartphones at relevant site, selection of site made by the entity responsible for the road, public transport operator or dealers association, little measurement error/loss</td><td>Site selection decisive for any statistical weights, loss/measurement error among those with multiple mobile phones or who lack mobile phone, battery or connection</td></tr>
<tr><td colspan="2">Recognition – camera/WiFi</td><td>Pertains to measurements over selected stretches using recognition, as above, see respective category</td><td>High loss due to deficiencies in repeat recognition</td></tr>
</table>

Various data sources are compared below with regard to other quality elements.

Timeliness and punctuality

How often data are collected and when they are available depend on the data collection method being reported on. In general, there are three different variants: automatic collection, which provides a continuous supply of data (e.g., mobile network data); active collection methods (e.g., TS apps) that collect data more periodically such as (but usually more often than) traditional collection methods; and section measurements that require the installation of roadside devices and can take time to install (e.g., if permission for surveillance is required). Compared to hardcopy questionnaires and telephone interviews, the new data collection methods that process data automatically deliver fast results while measurements are in progress.

Availability and clarity

Private actors are performing the data collection with these new methods. It is not currently possible to review the details of how data are collected or which algorithms are being used in cleansing and processing data into usable figures. If data are to be used in official statistics, they should be open. Open data refers to data that are made public so that they can be used by other partners for their own various purposes. One important principle for open data is interoperability, so that a given data source can be used together with others to create value (for society or for profit-driven purposes). Standards underlie interoperability. One example is the DATEX II standard, which is used for real-time data in Intelligent Transport Systems (ITS) and provided in Sweden via Application Programming Interfaces in Trafiklab.[7] Somewhat similar thinking may be needed in future TS work if data from different sources are used. Requirements with regard to the documentation of, for example, data cleansing and processing algorithms may be important in ensuring quality.

It is also important that data be available so that those analyses that are appropriate to perform on the material from a statistical and confidentiality-based standpoint are possible. This means, for example, that data material needs to be available to users so that they can perform the breakdowns they desire, rather than being referred to already prepared tables. This can be ensured either via open data or by entering into agreements with private actors. What is possible to do with the data material depends in large part on the data material being collected and the means available to use the data material for TS purposes.

Comparability and co-usability

There is at present little knowledge concerning how the new data collection methods perform compared with the existing methods. More studies are needed in which data are collected using new and old data collection tools in parallel so that they can be compared. As noted, interoperability between different methods is important. Standards need to be formulated so that data collected via different sources will be capable of comparison. This is made difficult by the fact that different methods collect data from different frame populations due to narrow recruiting and technology bias. More work is needed in this area.

3.5 New collection methods' potential for overcoming current deficiencies

Current TSs are encumbered by certain deficiencies/problems, as noted above:

- Suspicions of deficient representativity
- Deficiencies in data quality in terms of trip lengths and travel times
- Deficiencies in information about routes
- Heavy information burden on respondents
- Long intervals between measurements
- High costs

Table 3-4 presents an estimation of how new data collection methods could fill in the gaps that exist in the current methods. It is difficult to make a comprehensive assessment, as the new methods are still under development in many cases. In some instances the uncertainty has been so great that we chose not to make an assessment; more research is needed in such cases.

[7] Trafiklab offers open traffic data and a platform where independent developers can share in all the information generated by public transport daily, free of cost (www.trafiklab.se/).

Table 3-4 Summary of how new collection methods can overcome deficiencies in current TS. Light green indicates somewhat better. Dark green indicates significantly better. Yellow means somewhat worse. Red means significantly worse. Blue indicates that knowledge is lacking. Grey indicates that there are no major differences compared to current TS. Other designations:
* = representativity is assessed based on tool's target population
1 = depends on whether or not the system includes measurements of both boardings and departures
2 = depends on proportion of operators/vehicles included
3 = measures movements of vehicles, not of individuals
4 = depends on whether the system can identify where an individual gets on and off
5 = in combination with timetables
6 = total survey over a given time at a given site/sites; although difficult to obtain an overall picture, it may be very helpful in connection with statistical weights, etc.
7 = continuous data collection.

Data collection method		Representativity	Data quality in trip length/travel time	Route	Heavy information burden on respondents	Sparse time series	High cost
App	App – TS data					(7)	
	App – bicycle trips	*				7	
	App – exercise	*				7	
Online questionnaires							
TaP	Public transport		1	1		7	
	Shared bicycle					7	
	Carpool					7	
	Credit card					7	
Mobile network data	Mast position – activity	2				7	
	Mast position – Change of mast	2				7	
	Mast position – Time interval	2				7	
	Triangulation – activity	2				7	
	Triangulation – Change of mast	2				7	
	Triangulation – Time interval	2				7	
Floating vehicle	Navigation system	*2, 3				7	
	Automaker	*2, 3				7	
	Special equipment- publ.	*2, 3				7	
	Special equipment – taxi	*2, 3				7	
	Special equipment – bicycle sharing system	*2				7	
	Special equipment – car	3				7	
Passenger counts	Passenger counts – APC	*2	4, 5	4, 5		7	
	Passenger counts – Manual	*2					
Roadside devices	Section/site – camera	3, 6					
	Section/site – hose	3, 6					
	Section/site – WiFi/Bluetooth	6					
	Recognition – camera	3, 6					
	Recognition – WiFi/Bluetooth	6					

The new methods have major potential to reduce costs and the information burden on respondents, and to expand the means of collecting data more often/more continuously. Certain methods also have the potential to improve the quality of the collected data. On the other hand, many questions remain with regard to the representativity of the new methods. More research is needed to clarify how representative data material is to be collected. This is also a key issue, as unless representativity can be achieved it will be difficult to use the data material, despite the fact that it is of higher quality and easier and less expensive to collect.

Interpreting the colours in the table

Apps

Generally speaking, apps offer better data in terms of the deficiencies present in current TSs. Representativity remains an unanswered question, and research is needed to study this. Data can be collected continuously or at least for longer periods than is the case with current TSs.

Online questionnaires

Online questionnaires offer slightly better data in most respects than do traditional TSs, as contact by e-mail or links in social media offer superior means of finding respondents; the costs are lower and data quality can be higher if the right trips can be included in a map function. However, there are no improvements with respect to sparse time series.

Ticketing and Payment

Data from TaP are of lower quality in terms of trip length and choice of route, as it uncertain when the trips begin and end, even though it is known that someone has travelled. In the case of shared bicycles, the data are better with regard to trip length/travel time, as these data are linked directly to the payment system. The same is true of public transport, but only if the card is “blipped” at the end of the trip. All of these data are passive, so that the costs and the information burden on respondents are lower, although it is uncertain how easy it will be to gain access to the data. The degree to which the data are representative is also uncertain.

Mobile network data

A number of questions remain unanswered for mobile network data. It is uncertain whether this data are representative of the Swedish population, and this remains a topic for research. The representativity also depends on the mobile operators' market shares, and whether data from multiple mobile operators are being combined. It is also uncertain whether the trip length, travel time and route data are better or worse than data from the current TS, given that the generated data are often rough in terms of their accuracy, but researchers indicate that it is possible to obtain more accurate data. This is yet another research question. Because mobile network data are collected passively, the information burden on respondents is low, the cost is low, and data are/can be collected continuously.

Floating Vehicle Data

FVD measure vehicles and individuals; the representativity can be better than the current TS in the sense that the data represent the vehicle being used. This is true of all types of FVD except for special equipment on cars, as in most cases there will not be enough such devices to reflect the “population” of all cars in Sweden. However, FVD do offer inexpensive, passive and continuously collected data.

Passenger counts

Automatic passenger count data are of higher quality in terms of trip length, travel time and geographical route than traditional TSs, but only if the system recognises the people departing the vehicle. The data are collected passively and at less expense than in traditional TSs. We find that it is difficult to recognise people who are getting off public transport vehicles using manual passenger counts, which are also lower in terms of data quality, but they are also less expensive and can be performed more often than traditional TSs.

Roadside devices

Roadside devices offer data that are good for combining with other data sources, but it is difficult to derive an overall picture of traffic or transport activity from them. The data are collected passively, so the information burden on respondents is lower, but the quality is also lower compared with regular TSs. The representativity is good, but only for a given time and location. Movements of vehicles are measured rather than movements of individuals. The costs of recognition tools are unclear, as they are not yet as advanced as section/site measurements. The costs will presumably come down, but this issue is unclear and remains a topic for research.

4. Results for national tools and applications

This chapter describes the Swedish tools and applications used to collect travel survey data. The tools and applications are based on the description in Chapter 3 above.

4.1 Swedish examples – an overview

Interviews and document analyses/literature studies have identified a number of new national tools and applications that are of interest for future travel surveys (TSs).

We have identified eight Swedish tools and applications that can collect TS-relevant data. There is a great deal of data that is of interest from a TS perspective, and there are a number of applications and tools that can collect such data, but in this project we have concentrated on identifying new ways of collecting data that are sufficiently developed to be able to collect data later in 2017 and in the near future. Among the tools identified is a platform that co-uses data from different data sources (both “new” and “old”). We will not include a description of the old data sources here (such as hose measurements and the like, plate number recognition, which has already been in use for many years, and radar). The emphasis in the searches was on data sources or tools that offer TS-relevant data, and we have concentrated on information with regard to trips. As a result, we have not concentrated specifically on flow measurement tools in the searches, and believe that examples of new ways of collecting flow measurements (which may be of interest in combination with other data sources) that we did not find may be missing.

Table 4-1 provides a brief description of the various tools and applications that we found. We will discuss these tools in greater detail below, along with the strengths and weaknesses of the various solutions. Note that we have included only Swedish-owned solutions. There are other international companies that are operating in Sweden as well.

A number of research projects and testing of relevance are underway in Sweden in addition to these tools and applications. These projects have not yet produced any final results or concrete tools that could be used today, but they are of interest to monitor:

Mobile Network Data in Future Transport Systems project

Mobile Network Data in Future Transport Systems is a project that is doing research on the ways in which transport demand can be estimated based on mobile network data. Within the project is a WP that is intended to develop analysis tools. Linköping University is coordinating, while Sweco, Tele2, Ericsson, RISE SICS (previously Swedish Institute of

Computer Science), Swedish Transport Administration, City of Stockholm, Nobina, Swedish Railways, and Samtrafiken are also participating. This is a four-year project that began in early 2017.

Swedish Mobility Project

Samtrafiken is running a programme with the following vision: “We are enabling the emergence of simple, sustainable and profitable Combined Mobility Services”. One working track and project within the programme is designated as SMP – Establishment West, and is intended to facilitate and promote combined mobility services on a large scale, and third-party sales in general, by making producers' offerings available via a national access point and ensuring a shared regulatory framework, where applicable.

Various actors that market, bundle and sell Mobility as a Service services will be able to access open data sources (such as traffic data, real-time information and vehicle information) via the national access point, and will be offered a means of being resellers of offerings from various producers. The starting point is the Länstrafiken's (County Public Transport Authority) offering, but other traffic services involving shared resources could also be available, including carpools, taxis and shared bicycles. Some of the data sources may ultimately reside centrally, but the point of departure in most cases is that the integration platform will instead retrieve data from various local databases, and in that sense may be characterised more as an “exchange”. The services, regulations and technical solutions are being developed in close dialogue with potential actors and potential suppliers of technical platforms. The work is an extension of a project known as White Paper, which is also being conducted by Samtrafiken. A schematic diagram of Samtrafiken's objectives is presented in Figure 4-1.

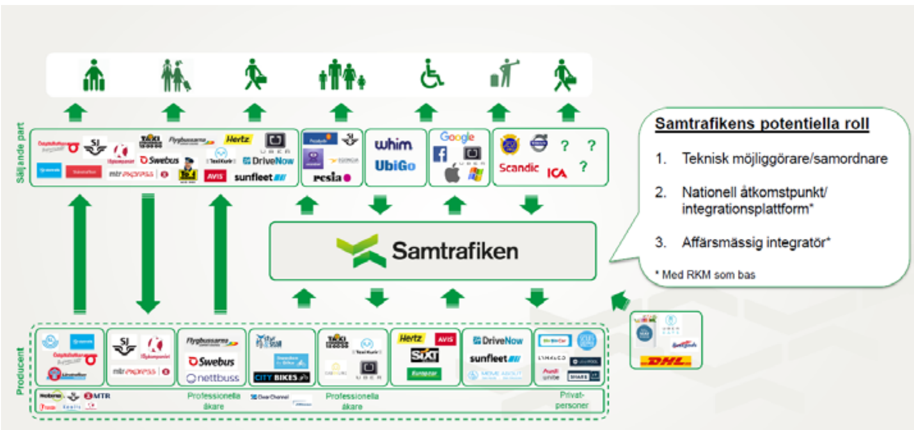


Figure 4-1 Objectives of the platform that Samtrafiken and its cooperative partners intend to build up.

Samtrafiken's potential role:

1. Technical enabler/coordinator

2. National access point/integration platform*

3. Business integrator*

*With RKM (Regional Public Transport Authorities) as base

Real-time system in Gothenburg

The Swedish Transport Administration and the Traffic Administration Office have been measuring travel times on a number of major roads in and around Gothenburg since 2006 in order to describe how accessibility by car is changing over time, and to monitor the effects of implemented measures. Travel times and congestion metrics are needed for traffic and social planning for the Traffic Administration Office's traffic management operations and by the municipal road authorities in the major cities. The documentation for the travel times derives from vehicles (FVD). Given the current solution, the system is not considered capable of meeting the requirements imposed by a TS, although an existing agreement enables more detailed data collection when necessary.

More advanced online questionnaires with map function

Other tools used to map where people have travelled are also available for online questionnaires, such as pointSurvey, which was developed by Entergate. Data were collected in various ways in PASTA, a four-year active mobility research project, but mainly via an online questionnaire that included a map function in which the user could click on where they had travelled. Data were also collected via mobile phone apps and accelerometers during the project, and much of the recruiting was done on social media.

Automatic passenger counts using beacons

Skånetrafiken cooperated in a pair of Vinnova-funded projects (along with several other actors) that involved the development of methods for automatic passenger counting:

- The Minimum Viable Device (MVD) project, which studied the feasibility of using beacons to induce a mobile phone app to record whenever a passenger gets on or off a vehicle.
- Point of Interest, which was based in part on guiding the traveller to the correct bus stop and notifying them once they had reached the proper location, but also included a section in which various techniques that could work together with a mobile phone to record a trip automatically were also studied.

The basic idea is to create a fully automated solution in the long run. The results of the project indicate that it is probably possible to find a solution, but that functions that complement a beacon solution would be needed to achieve a sufficiently high degree of detection probability.

It is important to note that there are also other data sources that could be of interest with regard to TS data, and that could be included in any platform that bundles different data sources. They include data from different companies that are developing vehicle data in public transport traffic and cars, as well as old-fashioned flow measurements. These data collection methods are continuously being developed further as well, but they have not been included in this report because they are already seeing extensive use. Traditional data collection methods that could be relevant in a TS context have been described in Sweco (2013) and elsewhere. There may also be other solutions in Sweden that offer TS-relevant data, but because our searches were concentrated on collected TS-relevant data, we may not have come across all the solutions that could measure traffic flows.

Table 4-1 List of the identified Swedish tools and applications with brief descriptions.

Name of tool/app	Company	Type	Brief description
Platform – Ericsson	Ericsson	Platform	Platform that consolidates various data sources (development phase). Raw data are not shared in order to protect personal privacy, but questions can be posed to the database. The platform is to be designed to receive data from multiple sources, for example vehicles, mobile network data, roadside ITS. Ericsson is building and operating the joint infrastructure in Drive Sweden, in which several pilot projects are also underway.
TRavelVU	Trivector Traffic AB	App	App that semi-automatically collects trip data (including mode of transport and purpose). It is also possible to obtain data about the user via a questionnaire in the app.
MEILI	KTH Royal Institute of Technology, but is Open Source		App that automatically collects trip data (including mode of transport and purpose). Possible to make corrections via web interface. Data about the user's profile can also be obtained via a questionnaire in a web service accessible via the app.
Cykelstaden – Göteborg	Traffic Administration Office, City of Gothenburg		App where all services included in the Traffic Administration Office's bicycle network can be found such as Styr & Ställ (a system for renting bicycles), pump stations and bicycle parking. It is also possible to report defects discovered on the bicycle path and submit suggestions for improvements. The app also includes a function in which a trip can be entered to view travel times and distances. The actual route is traced, although it is available only to the Traffic Administration Office. The user needs to turn the logging function on and off manually in order to log a trip.
Bumbee Labs	Bumbee Labs	Roadside device	Flow measurements of mobile phones are made using WiFi signals. This is done both indoors and outdoors, and offers flow measurements for pedestrians, cyclists, etc.
Meltspot	Meltspot AB		Measurements of mobile phones based on WiFi signals. This offers measurements of visit counts, but can also be adapted for flow measurements of pedestrians in a selected location, plus their movement patterns where a network of measurement points is available.

Facility Labs	Facility Labs AB		The tool consists of various measuring systems (vehicle registration numbers and laser-based) and a central unit that collects information about all vehicles passing a given location. Information is derived for traffic flows, properties of all vehicles and the owner's postal code, which can be used in combination with population statistics.
OTUS3D	Viscando		The flows (plus behaviours and interactions) with pedestrians, cyclists and vehicles are detected by analysing images collected using dual video cameras. The numbers, velocities and directions of road users (selectable, pedestrians, cyclists and/or motor vehicles) at a chosen site are obtained.

4.2 Platforms

Types of data sets

The platform being developed by Ericsson is intended to collect data from different sources and serve as a central data hub. Various ways of handling data are being developed here.

One such way is in development in the Mobile Network Data in Future Transport Systems project. Within the project the solution is known as a collaborative platform, where the code is brought to the data, rather than bringing the data to the code. Using this process enables the data supplier to ensure that those who are working with the data follow data rules and policies (e.g., to protect personal data). TS-relevant data cannot be retrieved directly from the platform, but questions can be posed to the platform, which the supplier will then answer. For instance, if one wishes to know what the mode of transport breakdown looks like for various groups in Gothenburg, the question can be posed to the platform and an answer will be received, even though no access to the raw data is granted. This is a solution that enables the utilisation of personal data (e.g., from mobile operators) without releasing raw data. Mobile network data will be included in the platform in an initial phase.

In addition to this, Ericsson is developing a platform within Drive Sweden. The platform is a cloud service that can provide data from multiple different sources like vehicle data, mobile network data, and data from traffic signals. In the Drive Sweden project, using the platform is mainly a means of ultimately automating the transport system. The idea is to be able to include different types of data in the future, and this also opens up the possibility of using data for TS purposes.

A collaborative platform enables the use of data from mobile operators without having to face problems involving personal privacy. Personal data are never released, which is an advantage compared to other methods. If multiple different data sources can be included, it becomes possible to use data for different purposes, and also offers the option of comparing different data sources against one another in order to validate them.

> The development of the platform is in its infancy, and many questions remain as to how it will be used in practice, including how quality assur-

ance will be performed, how representativity will be ensured, and how much it will cost to use the platform. These key issues are discussed below.

Ability to answer different types of questions

The questions that the platform can answer will depend on the data that are included. In an initial phase the collaborative platform will contain only mobile network data, and the platform will thus be able to answer those questions that mobile network data can answer (see Section 3.3). Its strength is that large amounts of data (big data) can be processed, thereby providing a picture of the flows in the transport system. Mobile network data can be used first and foremost to determine travel behaviour at a rough level, that is origin-destination (OD) movements and possibly different modes of transport in regional traffic. To answer more TS-related questions, mobile network data will have to be used together with data from other sources, which is the idea for the platform in the future. The cloud services being developed within Drive Sweden contain multiple data sources, with the current emphasis being mainly on real-time data for vehicle traffic as a means of preparing for automation. However, it is also possible to include more modes of transport, which may be done in future development efforts.

Usability in travel surveys based on statistical samples

Representativity is a problem with all data collection methods. If the platform contains only mobile network data, the sample will be skewed. Skewness derives in part from the fact that mobile network data provide information about the movements of mobile phones connected to specific networks. Even if all these networks are included in the platform, the problem still remains that people may carry more than one mobile phone, or none, and at present it is uncertain who these people are or how to correct the data. If the platform comprises multiple data sources, several different data sources can be used together to validate one another. For reasons of personal privacy, the platform will not release raw data, but that makes it difficult to know how representativity is to be ensured or how data are to be quality assured.

Statistical quality criteria

Because the development of the platform is in its infancy, no information about data quality assurance is currently available. However, it is worth noting that the platform will (for reasons of privacy) come to function like a black box, which means that quality assurance cannot be performed by a customer who purchases services from the platform.

Resource efficiency

No information on costs is available at present.

Legal conditions

Developing a platform that obtains data from various sources (e.g., different mobile operators) enables it to utilise personal data without violating laws against the handling of personal data. The platform also enables the use of data from multiple competing companies (e.g., mobile operators) without violating competition laws.

4.3 Apps

Types of data sets

Three different apps are described here: Cykelstaden, MEILI and TRavelVU. The first app collects bicycle trips from volunteer respondents in the City of Gothenburg. The app is part of an initiative to promote cycling in the city and contains other applications, such as a trip planner, that also support cycling. Cykelstaden collects GPS tracks for individual bicycle trips. No background information about the respondent (such as age, gender) is collected.

MEILI and TRavelVU are apps that collect TS data. Both apps determine modes of transport and stops automatically. MEILI also automatically determines what activities are being done. The apps are downloaded by the respondents, who answer a background questionnaire that provides information about the user. The user reviews how they have travelled and makes corrections when necessary, either directly in the app (with TRavelVU) or via a web interface (with MEILI). Both apps offer geographically coded TS data, as well as background information about the respondent.

Ability to answer different types of questions

Data from Cykelstaden can be used to derive information about routes chosen by cyclists in the city, as well as length, time and velocity data, and including delays in the bicycle trips. MEILI and TRavelVU provide rich data that can answer all relevant questions posed to TS data (see Section 3.3).

Usability for travel surveys based on statistical samples

There are no data indicating who is using Cykelstaden, so that without a link to another questionnaire it is difficult to know who is included in the sample, and it may be presumed that the sample is skewed.

Respondents for MEILI and TRavelVU should also represent a skewed sample, although no studies have been conducted to investigate this more closely. Because the apps collect background data about the respondents, it should be possible to determine the skewness present and weight the data.

Statistical quality criteria

There are no quality metrics for data derived from Cykelstaden. In TRavelVU and MEILI, the respondent must review and approve data before they are used in analyses. Anomalies in the data are cleansed in TRavelVU and MEILI.

Resource efficiency

Cykelstaden is free, while the costs of surveys with TRavelVU correspond roughly to the costs of online questionnaires. MEILI is not a commercial product and cannot be purchased. On the other hand, interested parties can use the code from the app and analysis services, which is open source.

Legal conditions

There are problems associated with the handling of personal data from apps. This is not the case with Cykelstaden, but MEILI and TRavelVU handle personal data. The data communication is encrypted, and the users sign an agreement in the app and approve the use of their data in accordance with the Swedish Personal Data Act. Preparations for the General Data Protection Regulation are underway with TRavelVU.

4.4 Roadside devices

Types of data sets

These roadside devices offer new means of making flow measurements. Here we will describe four different tools that are available:

Bumbee Labs makes flow measurements of mobile phones based on WiFi signals. This is done both indoors and outdoors, and offers flow measurements for, among others, pedestrians and cyclists.

Meltspot measures mobile phones by using WiFi signals. This offers flow measurements mainly for pedestrians at selected locations, plus their movement patterns if a network of measurement points is present.

Facility Labs makes flow measurements of vehicles by scanning registration numbers and making laser measurements. The registration number is compared against the vehicle register and can thus provide information about the owner.

Viscando uses cameras to perform flow calculations for vehicles, pedestrians and cyclists in specific locations. Data with regard to behaviour and interactions can also be collected.

All of the foregoing methods offer flow measurements, in other words the number (units) passing by a given point/points.

Ability to answer different types of questions

These data sources offer mainly measurements of the number of units passing by a given point and a specific mode of transport. The data can be used mainly in combination with other data, primarily for purposes of validation in, for example, modelling, or for other purposes where point measurements are relevant.

Usability for travel surveys based on statistical samples/WiFi

Bumbee Labs measures mobile phones equipped with a WiFi transmitter (however, the mobile phone needs not be connected to a specific WiFi network). According to their estimates, they capture up to roughly 85% of the population. Meltspot achieves average coverage of 83% of the number of visitors and is precise down to 0.5 square metres. Facility Labs measures the number of motor vehicles passing by a given point, while Viscando measures the number of vehicles, cyclists and pedestrians passing by a given point.

Statistical quality criteria

All of the methods compare results with other measuring methods — manual counts in the case of Meltspot and Viscando — but no information is available for the data cleansing or analysis methods used by Bumbee Labs or Facility Labs. Viscando and Bumbee Labs offer data with confidence intervals.

Resource efficiency

All the solutions require the installation of roadside devices, with cameras being the most expensive such equipment. The cost rises as more of a city has to be equipped with measurement points. All the solutions process data automatically and deliver an answer that requires no further processing.

Legal conditions

All the tools collect data in compliance with applicable regulations. Bumbee Labs initially saved Media Access Control (MAC) addresses in order to be able to track the movements of mobile phones via different measurement points, but they had to alter their service following an audit by the Swedish Data Protection Authority and can no longer collect MAC addresses, a problem that they have now solved. Meltspot adapted its system based on the new European General Data Protection Regulation (EU 2016:679). Facility Labs also has detailed vehicle data linked to vehicle registers, but no personal data.

4.5 Summary and conclusions

There are eight Swedish tools that have been identified and studied in somewhat greater detail in this project. These new tools offer new ways of collecting relevant TS data, and they may be supplemented by a number of ongoing initiatives and research projects focussed on new ways to collect data.

The studied tools vary in terms of their level of maturity. Roadside devices have come the farthest in terms of collecting Floating Car Data and have yielded several commercial tools that have been used in Sweden. The development of apps has progressed relatively far, and they are also being used in Sweden at present. Platforms that bundle different data sources (and specifically mobile network data) are still in a development phase.

None of the solutions described here offers a total solution or provides an overall picture of personal transport activity. This is due in part to skewness in the samples or to the fact that they can only count certain types of units. The price situation varies as well, and it is difficult to estimate the price situation for certain solutions. However, it is clear that passive data collection methods are less expensive but do not offer data linked to individuals, while active data collection methods offer more information but require more initiative, primarily in terms of recruiting participants.

5. Results for international tools and applications

This chapter describes the international tools and applications that can be used to collect travel survey data. The tools and applications are based on the description in Chapter 3.

5.1 International examples – an overview

Interviews and document analyses/literature studies have identified a number of new international tools and applications that are of interest for future travel surveys (TSs).

Some 50 tools and applications that can collect TS-relevant data have been identified within the framework of this project. However, the list is not exhaustive, and we know that there may be more applications outside of Sweden, particularly in countries with whom we have a language barrier (e.g., in Asia). The compilation was made based on discussions with experts around the world plus an international literature search, which provided excellent documentation for the types of tools that exist, although the list is not necessarily exhaustive. There is a great deal of data that is of interest from a TS perspective, and there are many applications and tools that can collect data, but the focus in this project has been on identifying new ways of collecting data. Old ways of collecting data include measuring with hoses and similar flow measurement tools that have been in use for years. The various tools derive from different parts of the world, and those that have been identified are presented in Figure 5-1 broken down by tool type like platform, TS app, other type of app, mobile network data, FVD or roadside device.

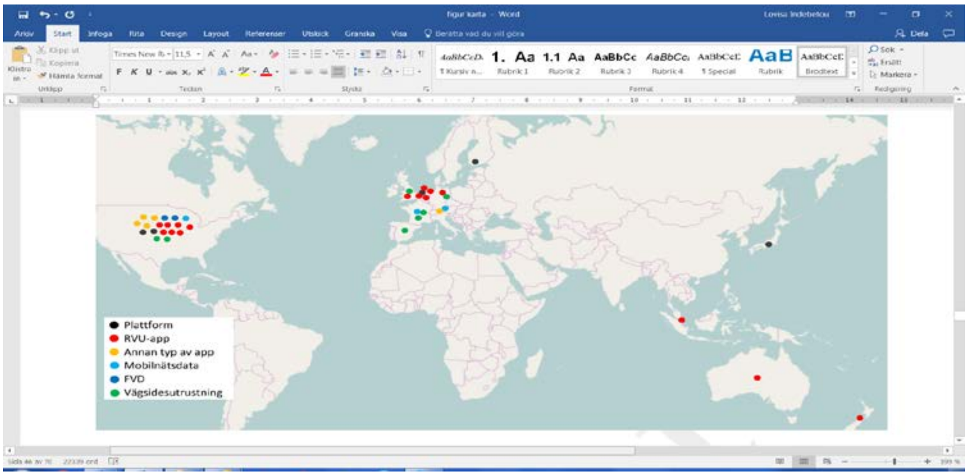


Figure 5-1 Where the various identified international tools come from (note that some tools were developed jointly in two countries, in which case they have been marked in two countries. (Source: basic map: OpenStreetMap)

Key: Plattform = Platform, Annan typ av app = Other type of app, Mobile network data = Mobile network data, Vägsidesutrustning = Roadside device

This description does not provide a complete picture of all development activity, but it does show the tools that we found, which are concentrated mainly in Europe, the USA and, to some extent, East Asia and Australia. We will describe the tools in greater detail below, and discuss the strengths and weaknesses of the various solutions.

Above and beyond these tools and applications, a number of relevant research projects and studies are also underway. One example of a relevant research project is Transportation Tomorrow Survey 2.0[8] in Toronto, Canada. This is a three-year research project to find new solutions for collecting TS-relevant data in Canada, and includes studies of apps, mobile network data and other data sources. Studies of TS apps will be conducted during the fall of 2017 to determine what effects different recruiting strategies can have in terms of collecting TS data via apps.

Relevant research that has already been published was discussed in Section 3.2 – Current state of knowledge per the scientific literature. Interested readers are referred to the references section for more details.

5.2 Platforms

The platforms that have been identified are of various types; some are designed to collect data for a given purpose (e.g., INRIX, which collects data for use in real time in Traffic Management Systems (TMS), and Ring-Ring, a platform for collecting bicycling data), while others collect large amounts of data for various purposes, which can be used to acquire an understanding of travel behaviour (e.g., Google Timeline, Sony's Lifelog).

[8] http://tts2.ca/.

A great deal of research is underway in this field, and many different platforms are being developed in various projects around the world. This is being done, for example, in numerous Intelligent Transport Systems (ITS) projects and products that develop ITS services such as cooperative systems, Mobility as a Service (MaaS), and autonomic vehicles. All of these require data that comply with standards and can be collected in the same platform from which data can be picked or to which questions can be posed.

Most platforms obtain data from motor vehicles, and consequently offer little information about cyclists and pedestrians.

Five different tools that illustrate the breadth of the existing platforms are listed below:

INRIX platform (inrix.com/products/traffic), global, headquarters in the USA

INRIX is an international company that offers a number of ITS services and sells data relevant to the use of the highway system. The data are both historical and real-time, depending on what is of interest to the customer purchasing the data. The data comprise travel times, OD matrices and real-time flows, including up to the city level. The data offered to customers differ in different countries/cities depending on the access to data. INRIX purchases data from multiple sources, integrates them in a platform, and then sells the data in aggregate form to, for example, cities that use data in their TMS.

Google Timeline (support.google.com/maps~), USA

Google Timeline collects data about all trips, their purposes and locations visited at the individual level. This is done using data from multiple different sources, but all are linked to Google services and include Google searches, Android phone use, Gmail, and GMaps. This creates a picture of how an individual travels, but the individual must turn on Location History to view their timeline. It is highly possible that Google saves the data even if the user does not wish to view their own data.

Lifelog (app.lifelog.sonymobile.com), Japan

Sony's Lifelog collects data concerning the time when someone has been active (walking, cycling, etc.) and transported themselves in some way or another. The data come from mobile phones and motion detectors (e.g., armbands). Times when one has slept and been passive are also collected, so that the number of calories burnt can be calculated. GPS tracks are also collected for all trips, although at present there are no aggregate data available for trips made.

Ring-Ring (ring-ring.nu), The Netherlands

Ring-Ring is a platform that collects bicycling data. Cyclists can download an app that records their cycling trips, and they earn points based on the collected data, which can then be used in stores or businesses. The platform was built under the auspices of a cooperative project involving the City of Amsterdam, various boroughs, and cycling and sustainability organisations. The platform offers bicycling data in real time that can be used by the city, while other cities can pay for access to the data as well. The platform is operating in the Netherlands but also records bicycle trips abroad.

Urbanflow (helsinki.urbanflow.io), Finland

Urbanflow was a project conducted under the auspices of World Design Capital 2012 in Helsinki, which was aimed at creating a platform and network to link smartphones, and deployed information screens throughout the city. The screens were intended primarily to furnish visitors with information, but also to give them a means of contributing information to the city with regard to, for example, unsafe intersections and maintenance needs.

Summary – platforms

Platforms are built for a purpose, for instance to facilitate MaaS services, offer better services to customers, collect bicycling data and so on. Many of the platforms were developed by private actors who sell services and sometimes sell data. The data can answer different types of questions, depending on the data available in the platform, but it is important to point out that data may often be lacking for various groups, as it may be difficult to gain access to data from certain ones. For example, vehicle traffic is easier to collect data for, while pedestrians and cyclists are more difficult. Platforms are often black boxes. To date, the focus has been on big data collected passively. Because the data are not approved by the user, there can be problems with errors, loss and the absence of individual data. If a company have individual data (such as Google) it may be difficult, for reasons of privacy/legal issues, to be able to use/purchase detailed information.

The quality of the data is unclear at present. Most of the examples offer coarse data, which cannot describe travel patterns in urban environments in detail. However, this field is unresearched, and it should be possible to produce data of higher quality. The data are also passive and inexpensive. It is possible that such data could play a role in future TSs, but more research is needed to determine their potential in Sweden.

5.3 Apps

There are many types of apps outside of Sweden that collect TS-relevant data. A description of the apps is provided below, with the apps divided into three different categories: TS apps, bicycle apps and exercise apps. The preponderance of the apps identified within the scope of this project are TS apps, due to the search words used in our searches. The purpose of these searches was to find examples of tools that could be used for TS-related purposes. Other apps may not have been used in conjunction with the collection of data relevant to TSs, although they could do so in the future.

TS apps

There are two different types of groups that have developed TS apps internationally. One comprises research institutes and universities, the other companies involved in traffic and city planning. Most universities that work on traffic and transport issues have developed apps or studied how GPS data collected via smartphones could be used to better understand traffic and transport activity. In the USA these universities have gone a step further and are developing apps that can be used commercially (although not always). In other parts of the world the development process often ends with a project that develops the app, with no further sales.

On the other hand, when companies develop apps they usually go the extra mile and develop one that can be sold (usually by a company that has experience with TSs in a traditional format), but with more focus on practical use, operation by and the personal privacy of the respondents.

Eights examples of TS apps intended to collect this survey data are described below, followed by three apps that collect TS data but have a mobility management (MM) purpose, in other words they not only collect data but also attempt to influence the traveller. The apps are presented in alphabetical order.

ATLAS II (Australia/New Zealand)

This app, which the University of Queensland (Australia) is behind, is the result of several research projects and collects trip information with GPS tracks. The traveller reviews the modes of transport suggested by the app based on the automatically collected data. The user fills out a form with socioeconomic information after having created an account on the app. The data collected comprise age, gender, living situation, occupation, and travel behaviour.

Commute Warrior (~gatech.edu/commutewarrior), USA

This is an app that was developed for research purposes by the Georgia Institute of Technology (USA), and is a fairly simple variant in which trips cannot currently be changed, although it is possible to view them on a map. The app automatically identifies the mode of transport, start and stop. It is unclear whether it is possible to purchase or use the app, and at present it is available only to Android users, although an iOS variant is in development.

DailyTravel (dailytravelapp.com), USA

This app was developed by Westat, which has experience with TSs based on traditional methods. To date, the app has only been used in one small study (< 100 participants), and is expected to be used by the City of Chicago later in 2017. The app is intended to serve as a complement to other methods for collecting TS data (e.g., online or hardcopy questionnaires). Trips are identified automatically, but it is important that the participants are able to change and correct their trips, and they fill out an online background questionnaire. In the USA it is more common to have TSs at the household level than at the individual level, so the app makes it possible to link in entire households, not just one individual. Privacy issues are addressed in each individual case, although the users sign an agreement (by clicking in a box) when they download the app and accept that their data may be used for research purposes.

FM Sensing / Mobile Market Monitor (www.mobilemarketmonitor.com), Singapore / USA

Mobile Market Monitor uses an app called FM Sensing, although the name of the company is also the name of the service (use of the app). The app was developed by the Massachusetts Institute of Technology's Singapore affiliate, and the company currently has a presence in both Singapore and the USA. The app was developed by researchers, primarily for research purposes. Its references are currently mainly academic, and it is made clear that the data will be used for research purposes. The app automatically detects the mode of transport and start/end, and the user is able to make corrections in the data. The current version has problems in terms of usability and operation, but development is continuing, albeit primarily in terms of machine learning so that the user needs to make fewer corrections. Users fill out a questionnaire in their web browser, not in the app.

modalyzer (www.modalyzer.com), Germany

Modalyzer is an app that was developed by the German firm InnoZ. The app automatically determines the mode of travel and origin-destination, which the user can then correct. Corrected data are then used to better determine future trips. The user can also flag errors to the company so that it can improve its algorithms. It has been used in a number of studies in Germany, and a study is being conducted in Ukraine in four cities in parallel. The app is currently available in four different languages (English, German, Ukrainian, Spanish). It also offers users a means of downloading their own trips. The app has been approved by the German data protection authority for use in collecting personal data.

rMove (rmove.rsginc.com), USA

rMove is a simpler variant of an app that collects TS data. It concentrates on identifying trip origin/destination, and then asks the user to fill in the mode of transport and purpose of the trip. There is no automatic identification of the mode of transport or activities/purposes. However, the app provides a geographical visualisation showing how one has travelled, plus start/end times and localisation so that the user will better be able to recall how they travelled. There is also a background questionnaire in the app. The app was developed by the RSG, a consulting firm with a number of offices in the USA.

Route Scout (nustats.com/webandmobileapps), USA

Route Scout is one of four tools offered by the American company NuStats that collect TS-relevant data. Route Scout is an app that automatically logs GPS and identifies modes of transport and origin/destination, which are then corrected by the user. The company also developed an advanced tool for collecting TS data online: it uses detailed map functions (NuTripX), and the tools complement one another depending on the purpose of the study.

Sense.DAT (www.dat.nl/en/products/sensedat), The Netherlands

The Sense.DAT app measures, analyses and stores travel patterns and trip experiences. It automatically senses the mode of transport based on the movement pattern, and measures where and when one has travelled. The trips can be corrected after the fact by the user. When the app is launched, one can answer background questions, which differ depending on the particular study in which one is taking part.

The Norwegian Centre for Transport Research (TÖI) and others have used the tool to map velocity discrepancies between regular cyclists and electric bicycle riders, as well as differences in velocity based on gender, age and path steepness. The tool has also been used to generate maps of cyclists' movement patterns in Oslo. The algorithms in the app are also behind a number of MM apps (see below) in the Netherlands and Belgium.

SmarTrAC (smartrac.umn.edu), USA

The University of Minnesota's SmartTrAC app was developed with the support of the United States Department of Transportation. The app collects information about trips and their purposes. It automatically detects and classifies daily activities and trips with the help of the phone's GPS accelerometer data. It lets the user view and correct the automatically identified and classed information, and compiles data at an aggregate level.

SmartSurvey (www.ait.ac.at/~/capturing-travel-demand)

The Austrian Institute of Technology's SmartTrAC app collects information about trips and their purposes, and automatically identifies eight modes of transport. It lets the user view and correct the automatically identified and classed information via a web interface, and compiles data at an aggregate level.

TRAC-IT (~/projects/trac-it/), USA

The app uses phone GPS data to collect travel data via several data processing modules and produces information about Point of Interest (POI), mode of transport and the purpose of the trip (via geographical information systems). TRAC-IT also integrates as a personal travel coach to help the user save time and money by making their travel behaviour more efficient. An application will be developed in a later stage to influence travel behaviour regardless of the mode of transport. However, the development of the app does not appear to have been active for many years. It appears to have been one of the earliest attempts at a TS app.

Summary – TS apps

TS apps are being developed around the world, and many of the most interesting ones have been developed in the USA. Most of these apps are based on the same concept: automatic detection of origin/destination and mode of transport, a means of making corrections when the algorithms are incorrect, and a simple background questionnaire. The differences have to do mostly with usability, whether any corrections are made, and whether questionnaires are filled out in the app or online. Some of the interesting refinements available on some of the apps include data collection at the household level, a means of flagging errors, and availability in several languages.

There is very little information about the extent to which the users have to turn each respective app on/off, or what information on trips made is available to the user (e.g., lists of trips made over a week). It is not possible to try out many of the apps without participating in a study. In many cases it is also unclear whether the apps are being used on a large scale, which could be an indication that many of them have not even been tested in the real world or to any great extent.

Mobility management travel survey apps

In addition to pure TS apps, there are also apps that use the same type of basic algorithms and display the individual's trips, but have the purpose of influencing their travel. These MM TS apps include:

Positive Drive (positivedriveapp.com), The Netherlands

This app was developed by the Dutch firm Ijsberg Holdings. The app automatically identifies the mode of transport being used and offers positive feedback via smileys:

> *Positive Drive is the first gamification tracking app that only rewards positive behaviour in traffic. With precision nudges we aim to solve the problems of the future: the increasing CO2 levels and traffic safety. (From the website: 2017-05-17)*

This is done, for example, to inform drivers that they are driving too fast or to encourage cycling and walking. The app also provides a picture of the travel, although the users do not confirm their trips.

BetterPoints (www.betterpoints.uk), UK

The BetterPoints app logs walking, cycling and public transport trips, and the user earns points so as to encourage more sustainable modes of travel in terms of both health and the environment. Points can also be awarded for carpooling and volunteer work. The collected data material can be used to determine, for example, which parks are most popular, what parts of the city many people are bicycling in, and what bus lines need to be improved.

RouteCoach (www.routecoach.be), Belgium

This app was developed in Belgium by Mobiel 21, which is affiliated with the City of Leuven. It collects travel behaviour data and offers customised tips to users on how they can travel more sustainably. The users do not correct/confirm their trips.

SMART / Sesamo, The Netherlands

Mobidot developed the TS app Sesamo, which automatically collects travel data, including origins/destinations, routes, purposes, and modes of transport. Information about the respondent is collected via introductory questions in the app. In addition to the app, Mobidot also sells the technology behind the app to customers who want to develop their own TS app, in which case it can be adapted based on the customer's wishes. Mobidot's technology is consequently behind mobility apps such as SMART, an MM app that is being/has been tested in Gothenburg and elsewhere.

Summary – MM apps

The MM apps are very similar to the TS apps (above and beyond the fact that they also contain elements that are intended to influence behaviour), and are more common in Europe, where efforts to transition to a sustainable transport system have come further than in the USA. However, it appears that MM apps are intended, to a great extent, to push their users to correct/approve their trips. It happens that the same engine will be used in both TS apps and MM apps.

Bicycle apps

There are a number of apps that collect bicycle traffic data. The European Cyclist Federation website[9] has a separate page devoted to the apps that are available and why they can be important in understanding bicycle traffic. There are additional examples, some of which are elucidated in greater detail here. It is interesting that the developers of these apps tend to have the goal of increasing bicycle use, and they are being developed by cycling enthusiasts and even government organisations to help promote cycling. Data concerning cyclists and their trips are currently lacking (particularly route information), and this is the main factor driving the development of these apps.

Bike Citizens (www.bikecitizens.net), Austria

The bicycling app Bike Citizens is sold by the BikeCityGuide Apps GmbH company in Austria. The company was started by two bicycle couriers and is designed especially for the needs of cyclists in cities. It is currently available in over 350 cities in Europe, Australia and the USA. The app is a navigational tool for cyclists that finds the route, pinpoints a location (it points out POIs while one rides), provides guidance via audio, and records the trip. The activity function makes it possible to record the trip in terms of route and velocity. If an Internet connection is available, all tracks are displayed for the user as a personal heat map. The company has also developed a simple and inexpensive tool that can be used to secure a smartphone to the handlebars (see Figure 5-2).

[9] https://ecf.com/what-we-do/urban-mobility/cycling-data collection.

Figure 5-2 Bike Citizens tool that makes it possible to secure a smartphone to the handlebars.

GPS data for recorded trips from cyclists are provided to partner cities. Insight into the cycling patterns in the city can be obtained by analysing and visualising these data.

CycleTracks (~/cycletracks-iphone-and-android), USA

This app maps bicycle trips in San Francisco, and was developed by the San Francisco County Transportation Authority to understand the needs of cyclists with a view to prioritising cycling infrastructure investments more effectively. CycleTracks uses mobile phone GPS to record the user's bicycle trips and times, and displays maps of the trips. At the end of each trip, data representing the purpose, route, data, and time of the trip are sent to the database at the San Francisco County Transportation Authority. All personally identifiable data are kept confidential. CycleTracks has been used in several research projects, mainly in the USA.

Summary - bicycling apps

There are numerous bicycle apps. Most have been developed to make cycling more attractive, rather than mainly to collect travel behaviour data. This is reflected primarily in the fact that the users often have to actively start/end the logging function, and that navigation systems are used more often for unfamiliar trips than for those that are made daily.

Exercise apps

There are many exercise apps, a few of which are described below:

Moves (www.moves-app.com), USA

This app was developed by ProtoGeo Oy, which is now owned by Facebook. The app is downloaded to mobile phones and uses a pedometer and GPS to map trips made on foot (walking or running) and by bicycle. Data are not collected for anything more than the user’s own use, nor are any background data collected. The connection to Facebook could offer an opportunity for large-scale data collection.

Runkeeper (runkeeper.com), USA

GPS data are collected from the mobile phone via the Runkeeper app and used to calculate travel distances and routes. The main purpose of the app is to enable the person exercising to obtain data about their workouts. The user has to turn the app on and off. However, Mapbox is working together with Runkeeper to analyse trips, and heat maps showing the flows of cyclists and pedestrians on different links can be delivered.

Endomondo (www.endomondo.com), USA

Endomondo, which was formerly Danish, has been acquired by an American company called Under Armour. The company also acquired MapMyFitness, which is a similar app. Endomondo is designed to track the user's workouts, provide audio feedback along the way, and offer guidance for achieving set goals. The app is synchronised with Endomondo.com, where the user can access a workout diary and analyse their workout activities.

Summary – exercise apps

There are many different exercise apps. One common characteristic they all share is that they were developed primarily to collect workout data for the user. However, there are examples where such data are aggregated and analysed in order to try and provide a more comprehensive picture of how (mainly) pedestrians and cyclists move about. There are differences between the apps in terms of whether or not the users need to actively turn the logging function on and off.

Summary – apps

There is a variety of apps around the world that collect TS-relevant data. Apps that are intended to collect this data could be used in Sweden, although it is unclear whether all the apps meet the requirements in terms of handling personal data, as the regulations governing the handling of such data vary from country to country. As of next year, apps developed in the EU must comply with the provisions of the General Data Protection Regulation. It is also unclear whether it is possible to purchase data from services such as Endomondo and Runkeeper.

All apps suffer from problems with representativity, as the data collection process is based on individuals having to download the app and thus be active in collecting the data. TS apps can be supplemented with targeted recruiting and an understanding of which groups have participated in the studies, but none of the services we found studied the problem of representativity to any great extent, although a number of research studies of this problem are currently in progress.

Apps seldom publish exact descriptions of how data cleansing and mode of transport identification are performed. Anomalies in the data need to be cleansed even if individuals have approved the data.

People in many countries have confidence in apps for collecting data in the future, but the problem of comparability with earlier studies and between countries has not been addressed either.

5.4 Mobile network data

There are a number of mobile network data research projects around the world, in addition to several companies that are offering services in cooperation with mobile operators. The projects are developing various forms of cooperation and identifying which data can be produced (e.g., South Pole Group in Nuremberg, Germany in cooperation with mobile operators Telfónica, and Teralytics, as well as NTNU, SINTEF and Telenor Group in Norway). In addition, it is also worth noting that the possibility of using mobile network data is also present in countries that have little or no information about travel, but do have extensive mobile phone use (e.g., in parts of Africa and Asia).

AirSage (www.airsage.com), USA

The American firm AirSage sells various types of OD matrices estimated based on mobile network data. The data are collected from all the mobile phones present within a given area. The information also includes gender and age breakdowns, although it is unclear how these are determined. The focus is mainly on road traffic.

Invenium Data Insights (www.invenium.io), Austria

The Austrian company Invenium Data Insights collects events from all mobile phones that have a subscription with the biggest mobile operator in Austria. It receives time and location data for each event, and can thereby track movements. The tool can calculate mode of transport breakdowns for car, train and subway traffic, and determine the purpose of each trip made.

Mediamobile (www.mediamobile.com), European, headquarters in France

Mediamobile uses mobile network data from Orange, an operator in France, to estimate prevailing traffic situations. The service determines traffic flows and chooses the best route. The emphasis is on road traffic, and there is no breakdown in terms of mode of transport.

Summary – mobile network data

Cooperation with mobile operators is essential to the ability to offer mobile network data to determine traffic flows. There are a few companies that offer such services, first and foremost AirSage in the USA, which offers services specifically for that purpose, while the other services we found are small parts of a larger enterprise.

It is unclear whether data from private actors can be purchased and used by other organisations. Data cleansing routines are generally not published, and it is unclear how the algorithms work in terms of aggregating and/or compiling data.

Mobile network data are passive data and the samples are skewed. AirSage is perhaps the most successful at present, as the laws governing the handling of personal data in the USA are not as strict as those in Sweden and other European countries.

5.5 Floating Vehicle Data

Many companies offer FVD, including automakers. A few examples are provided here:

Octo Telematics (www.octotelematics.com), Worldwide (USA, Europe)

At the behest of major insurance companies, Octo Telematics has been allowed to furnish many vehicles with equipment on their GPS units. The equipment collects data concerning driver behaviour in various traffic situations, and the information is then used to generate models and algorithms for calculating risk, in part as a basis for setting prices in the insurance industry. The company is also active in the ITS branch in general, and heavily involved in cooperative systems that provide various types of FVD.

Strava Metro (metro.strava.com), USA

Strava mainly draws on its users (runners, cyclists, etc.) to collect GPS tracks for exercise purposes, although Strava also includes Strava Metro, a service that sells data from Strava users and offers maps showing pedestrian and cyclist flows (as well as data). It sold its services in numerous cities in the USA and Australia.

TomTom (www.tomtom.com), NL and Waze (www.waze.com), USA

TomTom was acquired by Google and offers GPS information, both historical and real-time data, via navigation systems in private vehicles. This makes it possible to provide advance warnings about, for example, locations with traffic jams and average velocities in advance.

Google also acquired Waze Mobile. Users download the app and, using GPS, Waze is able to derive real-time velocities for their vehicles. These velocities are then compared to the average velocity on the stretch of road in question. The user can then view a map and receive real-time information of the traffic situation and any traffic jams. All trips made are collected and used by Waze in aggregate and anonymous form to conduct various types of flow analyses.

All of this is done as a way for Google to improve its user services such as, for instance, trip planners with real-time estimates and information about road system delays.

Summary – FVD

Data on how vehicles or individuals travel are of interest for various purposes, and there are many examples of FVD outside of Sweden. We described just a few examples here. This type of data tends to be of interest only if collected on a large scale; the aggregate FVD data can be purchased and used to understand travel activity.

It is unclear whether data from private actors can be purchased and used by other organisations. Data cleansing routines are generally not published, and it is unclear how the algorithms work in terms of aggregating and/or compiling data.

Because the user signs an agreement in connection with their use of such a service (e.g., route planning, feedback on bicycle trips), the use of personal data is not usually a problem. However, depending on how these agreements are drafted, they may make it difficult to use the data material for TS purposes.

5.6 Roadside devices

There are many examples of various technologies used to make flow measurements with the help of roadside devices. Here are some examples:

Clearview Intelligence (www.clearview-intelligence.com), UK

Clearview Intelligence developed, among other things, a wireless vehicle-detection system to measure traffic flows (both motor vehicles and bicycles). It uses magnetic sensors in the ground.

DILAX (www.dilax.com), headquarters in Germany

DILAX developed various measuring devices to measure indoor and outdoor visitor flows, data concerning numbers of visitors, and passenger flows in public transport. It is involved in automated flow measurements, automated passenger counts, "dynamic seat management" in public transport, and tracing smartphones, and offer data analysis tools.

The technology is based on cameras and 3D sensors. In addition to calculating numbers of visitors and passengers, the following information can also be obtained:

- Detection and analysis of visitors, for example by gender and age
- Licence plate recognition
- WiFi tracking
- Point of Sale data like data from sales made at a given point, where it is possible, via payment cards, to obtain information about the individual such as their gender and age
- Automatic time clocks in workplaces

Econolite (www.econolite.com), USA

Econolite manufactures cameras, radar equipment and wireless ground detectors for measuring traffic. It also offers systems for compiling and monitoring traffic flows in real time for use in controlling, for example, traffic lights to optimise traffic.

Eco-Counter / Amparo Solutions (www.eco-compteur.com), France

Eco-Counter offers various devices for measuring GC flows. These devices are similar to measuring hoses. Both flow and direction can be measured. There are also variants that have real-time signs that display the number of cyclists/pedestrians who have passed by, which can be used for MM purposes.

FLIR (www.flir.com), headquarters in the USA

FLIR is active in the Intelligent Transport Systems (ITS) field and makes different types of cameras and measurement systems. FLIR has developed various types of products for ITS, including the “all round” sensor, TrafiOne, which uses thermal image processing and WiFi tracking to capture data from all types of traffic (vehicle, pedestrian and bicycle traffic). The data can then be used to measure traffic flows and backups, and in traffic safety analyses. WiFi tracking makes it possible to derive travel times and flow data for the stretch of road where the equipment is deployed.

Quantaflow (www.quantaflow.com), France

Quantaflow measures indoor visitor flows (e.g., shopping centres). It offers various measuring devices and indoor sensors, including lasers, scanners and 3D cameras. The latest device it developed is a 3D camera that is mounted on, for example, the roofs at the entrances of shops of interest so that it can count the number of visitors.

Urbiotica (www.urbiotica.com), Spain

Utbiotica offers wireless sensors to provide guidance in parking lots, so that areas in cities can be used more efficiently with the help of big data technology.

Summary – roadside equipment

There are currently many new methods for making flow measurements among various groups of road users. These tools are already in use in many countries, both indoors and out. All the tools are passive and require no active cooperation from the road users. They all have margins of error in terms of identifying the number of road users, but usually have methods for quantifying errors or comparing against other methods as well.

The price situation for the various methods varies, even as the price also varies depending on the number of measurement points, the data quality and which data are being delivered (i.e., which mode of transport, flow measurements only or recognition as well).

5.7 Summary and conclusions

Many new methods for collecting traffic and transport data are found around the world. The new methods reflect the new solutions and ongoing research projects found in Sweden.

Up until fairly recently it was mainly government organisations that collected data and published official statistics. The situation has changed, and our review of the international tools shows that companies are not just developing tools that collect data, but they also own (and sometimes sell) data that have already been collected for another purpose and can be used to answer TS-relevant questions. Data constitute a new product on the market, and numerous companies are already seeing the potential benefits.

It is difficult to ensure the quality of data collected using these new tools, as the data cleansing and processing algorithms are usually not published. It is also difficult to estimate costs for most of the tools, as the price picture is unavailable in most cases.

Legal conditions vary from country to country, as do the ways in which personal data are handled. In some countries it is easier to collect and use/sell personal data. Sweden (and the EU) have stricter regulations than do many other countries, which can affect which tools are used in which countries, and which adaptations are needed to comply with the regulations.

There are many tools that offer similar products (e.g., many TS apps, many roadside devices). It is not possible to identify which ones are better than others, as their advantages and disadvantages depend on the application in question.

6. Conclusions and recommendations

6.1 Summary and discussion

Problems with traditional methods of collecting travel survey (TS) data are being encountered in many places in Europe and the rest of the world. There are also many people who see the potential in new ways of collecting data. As a result, a great deal of development work is ongoing in this field, both in Sweden and in the rest of the world.

The aim of this project was to inventory the national and international arenas in the field, and to map and characterise the level of technical maturity of the applications and tools that have been or soon will be developed to collect data about travel behaviour or, more generally, about personal mobility. The aim was also that this mapping could serve as the basis for method selection in a pilot project that is scheduled to be conducted during the fall of 2017.

Different types of tools/methods for collecting TS data are under development, both tools designed for the specific purpose of collecting such data, for example more advanced online questionnaires and apps, and methods for using data that have been collected for some purpose other than for use in a TS context such as mobile network data. In this report we have taken stock of what is going on right now, and tried to summarise the key elements. However, there is a risk that using search terms and search engines will not succeed in finding everything that is relevant, for instance due to language barriers. Our starting point was to concentrate on new ways of collecting data and/or new ways of using traditional data. This means, first and foremost, that the numbers of tools described within various areas differ in terms of their extent, but that should not be interpreted as an assessment the importance of a given area is per se.

In many cases the ability to answer questions concerning our knowledge of mobility choices requires that the data material include information about the individual and the purpose of the trip. Such data are difficult to collect using many of the tools/methods that were designed for purposes other than TSs such as mobile network data, ticketing and payment systems, on-vehicle measuring, and roadside devices, and consequently need to be collected using traditional methods, online questionnaires and/or apps designed for TS data. On the other hand, these individual-focused tools/methods are difficult to employ continuously and on a large enough scale to perform all of the breakdowns that certain questions require. If we can succeed in combining these types of data sources, so that interviews/online questionnaires/apps can provide the information at the individual level and other, bigger data, can provide the information about total travel volumes and aggregate patterns, then a very usable TS database could be created.

Generally speaking, there appears to be a trade-off between large volumes of data and information about the individual and the purpose of their trip. These issues are determinative with respect to where the line is drawn.

In future, data should be combined in various ways so as to reflect an overall picture of traffic and transport activity, and to answer the various questions that are posed. For example, individual data can provide a deep understanding of how various groups travel, and this can be combined with continuous and passive data collection that provides a picture of how traffic and transport activity is changing over time. Section or site measurements along stretches of cycling paths could be combined with data from bicycle-sharing systems to derive an understanding of cycling paths, for instance. The possibilities are many and should be based on the questions that are there to be answered.

The data collected using the new methods usually reside with private actors. For a government agency to avail itself of these data, quality assurance must be in place for the various data sources, for example for individual data via user confirmation, and for passive data via comparison with other data sources or via random samples.

The tools vary greatly in terms of their level of maturity, from research projects where a concept is being developed to finished tools that can be used right now to collected data. There are also differences in terms of which tools are available for use in different countries, although developments in Sweden do mirror those in other countries. There are also data sources that are being discussed a great deal, such as travel planners, but it has not been possible to find any tools/methods where these are actually being used. Common to all the tools and methods is the fact that more knowledge about recruiting, samples and loss is needed. The focus to date has been on technical development. Few studies have investigated the differences between traditional collection methods and the new tools and methods in terms of samples, loss and results. There is also still some work to be done with respect to how raw data are to be processed to yield relevant information. The journey from the technical feasibility of doing something to an off-the-shelf solution is one that many of the tools still have to make.

The legal aspects need to be studied more closely in most cases. With respect to data collected for purposes other than TSs, the consent/informing of those who provide the data for the data material may be required for the material to be legally usable. The ground rules will also change somewhat once the EU's new data security law (General Data Protection Regulation) takes effect on 25 May 2018. Privacy protection for the individual will be strengthened, but the new law will also establish consistent regulations throughout the EU, which could facilitate both the import and export of tools/methods within its borders. However, it is unclear what regulations will apply to the importing of tools from countries outside of Europe, as the legal aspects are handled in different ways in different countries.

Sweden has products in some areas that measure up well against what is happening elsewhere in the world. It also has an active mobile network and automotive industry that could be a conduit for new solutions.

We know that much of the research and development work on new tools/methods is being done in certain countries in Asia. This is likely due to the fact that the laws governing privacy protection there differ from those in Sweden and Europe. This is also something to consider for those wishing to import tools from such countries. Europe appears to have one of the strictest sets of privacy laws, stricter than in either Asia or North America.

6.2 Conclusions

In broad terms, we have drawn the following conclusions after reviewing the tools/methods and literature relevant to collecting TS data:

Future TSs will likely comprise different data sources that are integrated in a platform, where different types of data can be integrated.

Overall (all data sources collectively), information is needed about **the trip** (geographical origin and destination, and preferably route; start and end times and the duration of the trip; mode of transport used for the trip), about **the purpose** (why the trip is being made), and about **the individual** (who is travelling and their characteristics).

Different types of data material can answer different questions, and it is not necessary to impose the same strict requirements in terms of information and accuracy in every case.

Relatively **in-depth information about the individual** and the purpose of the trip will continue to constitute a key element in TSs, and this will require the active involvement of the person furnishing the data.

Data that do not contain information about the individual or purpose of the trip can **supplement data** containing such information, but there are also areas of application where such information is not always needed, for example in generating origin-destination (OD) matrices.

GPS tracks of trips can provide **an even better understanding of travel behaviour** than traditional surveys could do, and can broaden the areas of application for TS data.

To derive an understanding of travel decisions, it is important that data for **all of the (primary) modes of transport** be collected, as there is a tendency to focus on car traffic.

Although the technology is in place in many cases, most of the tools/methods are not mature in terms of **representativity, comparability** with traditional TSs, and the **legal** aspects. A step forward needs to be made in order for the tools/methods to be usable in a TS context.

6.3 Suggestions for further studies

General issues that need to be studied further:

Representativity, recruitment, samples, loss, and skewness need to be studied for all types of new tools/methods. It is also important to know whether the frame population is the one desired in a TS context. Knowledge in this area can provide means of addressing the skewness that is present in the various tools/methods.

Data quality. There is no information available with regard to the accuracy or data quality of many of the tools/methods that we studied, due in part to the fact that they are in a development phase and/or new, but also because it is difficult in some cases to find a true value against which to compare.

Comparability with traditional TSs. Further study is needed to maintain comparability with earlier TSs. The very best approach would be to collect data in parallel with different methods in order to be able to analyse any differences.

Interoperability between data collected using different types of tools/methods needs to be studied further. Existing registers should also be included.

The legal feasibility of and restrictions on individual tools/methods, and with regard to platforms that include data from multiple different tools/methods.

In addition to the foregoing, the technical development of the tools/methods must continue.

6.4 Recommendations for pilot testing

In the fall of 2017, the project (of which the work described in this report is a work package) plans on conducting pilot testing to obtain a better understanding of how future TSs can be conformed. A brief description of the pilot testing we consider relevant is provided below, based on the results of this work. The choice of pilot tests does, however, need to take into account more aspects than just the results of this report, and the recommendations below are to be viewed as input in the discussion for the selection of pilot projects, rather than the final solution in that regard.

Generally speaking it is advantageous for all the pilot testing to be conducted in the same location. This creates more favourable conditions for monitoring and comparing the results, and also makes the work within the platform more meaningful. The site where the pilot testing is carried out will hopefully benefit from the data that are collected as well. We suggest that the pilot tests be sited in Gothenburg (if possible), which is playing an active role in the project and has shown a strong interest in these issues, and the City of Gothenburg is also planning to conduct a TS during the fall of 2017 that could serve as a starting point for the pilot testing.

Pilot test 1 – Tools that deliver individual data

Individual data will continue to constitute key information in a future TS. There are relatively few tools/methods that collect such data, and our understanding of the data quality, comparability and representativity for various recruiting methods must be improved. We propose two subprojects in this context:

Pilot subproject 1a – TS app: comparability, data quality and representativity

In this pilot subproject, a TS app would be tested with regard to comparability, data quality and representativity. Data collection would occur in parallel with data collection for a traditional TS using two different recruiting methods for the app (e.g., random selection of postal addresses and crowdsourcing). We propose the following types of analyses:

1) Comparative analyses between trip data approved by the user and the results from the traditional collection process (comparability).

2) Comparison of individual data among the various recruiting methods for the TS app and the traditional collection process (representativity).

3) More in-depth analyses of comparisons between what the app detects automatically and what the user approves being done, as well as a breakdown by different modes of transport (data quality).

Pilot subproject 1b – Online questionnaire: comparability and representativity

In this pilot subproject, different online questionnaire designs would be tested with regard to their comparability and representativity. Data collection would occur in parallel with data collection for a traditional TS using a few different online questionnaire designs (e.g., a simpler, shorter variant, a longer and more detailed variant, and a more advanced "smart" variant). We suggest the following types of analyses:

1) Comparative analyses between trip data collected using the various online questionnaires and the results from the traditional collection process (comparability).

2) Comparisons of the response rate, loss and individual data between the various online questionnaires and the traditional collection process (representativity).

Pilot test 2 – Platform for integrating different types of data

To be able to derive benefit from different types of data sources they need to be integrated. This will probably be an important element in future TSs. Both traditional data collection methods (such as hose measurements) and data from new tools/methods can be included in such a platform. The way in which such a platform is to be designed needs to be studied and developed further. This pilot test is intended to include data from different types of data sources, and to study which possibilities, difficulties and techniques are needed to enable the use of the data material. Legal aspects should be incorporated as well.

The user of the platform should also be able to know which data are available and, in order for the platform to be able to answer the questions posed in a TS, the data material needs to collectively cover all of the (primary) modes of transport, and provide answers as to when, where, how and who is travelling. The pilot test need not include every type of data source, but rather there is room for interested actors to participate and influence the data content of the platform. Types of data of interest could be mobility network data, Floating Car Data or data from roadside devices. The results from pilot test 1 should also be incorporated in the platform.

Pilot test 3 – Mobile network data, OD matrices and long-distance trips

Not every type of analysis imposes heavy demands in terms of detailed individual information. One interesting example is whether mobile network data could be used to create OD matrices. In pilot project 3, this could be tested out in various parts of Sweden, for example urban, suburban and rural areas. The pilot project would study the means that exist to generate OD matrices from mobile network data, and the accuracy with which those matrices could be constructed.

In another part of the testing, the same data could be analysed to identify long-distance trips. It is potentially easier to determine the mode of transport from mobile network data when trips are longer. This pilot test would study whether it is possible to do so, and what types of information derive in this way. The feasibility of obtaining information about the individual should be studied as well.

7. References

List of tools and apps

Name of tool/app	Company	Type of tool	Country
Platform – Ericsson	Ericsson	Platform	SE
TRavelVU	Trivector Traffic AB	App	SE
MEILI	KTH, but is Open Source	App	SE
Cykelstaden – Göteborg	Traffic Administration Office, City of Gothenburg	App	SE
Bumbee Labs	Bumbee Labs	Roadside device	SE
Meltspot	Meltspot AB	Roadside device	SE
Facility Labs	Facility Labs AB	Roadside device	SE
OTUS3D	Viscando	Roadside device	SE
INRIX platform	INRIX	Platform	USA
Google Timeline	Google	Platform	USA
Lifelog	Sony	Platform	Japan
Ring-Ring	Ring-Ring	Platform	NL
UrbanFlow	Nordcap/UrbanScale	Platform	FI
ATLAS II	University of Queensland	App	AUS/NZ
Commute Warrior	Georgia Institute of Technology	App	USA
DailyTravel	Westat	App	USA
FM Sensing	Mobile Market Monitor	App	SPG/USA
Modalyzer	InnoZ	App	DE
rMove	RSG	App	USA
Route Scout	NuStats	App	USA
Sense.DAT	DAT	App	NL
SmarTrAC	University of Minnesota	App	USA
TRAC-IT	University of South Florida	App	USA
Positive Drive	Ijsberg Holdings	App	NL
BetterPoints	BetterPoints	App	UK

RouteCoach	Mobiel 21	App	BE
SMART/Sesamo	Mobidot	App	NL
Bike Citizens	BikeCityGuide Apps GmbH	App	AUT
CycleTracks	San Francisco County Transportation Authority	App	USA
Moves	Facebook	App	USA
Runkeeper	Runkeeper	App	USA
Endomondo	Under Armour	App	USA
AirSage	AirSage	Mobile network data	USA
Invenium Data Insights	Invenium Data Insights	Mobile network data	AUT
Mediamobile	Mediamobile	Mobile network data	FR
Octo Telematics	Octo Telematics	FVD	Global
Strava Metro	Strava	FVD	USA
TomTom/Waze	Google	FVD	USA
Clearview Intelligence	Clearview Intelligence	Roadside device	UK
DILAX	DILAX	Roadside device	DE
Econolite	Econolite	Roadside device	USA
Amparo solutions	Eco-counter	Roadside device	FR
TrafiOne	FLIR	Roadside device	USA
Quantaflow	Quantaflow	Roadside device	FR
Urbiotica	Urbiotica	Roadside device	ES

Information providers via interviews, telephone or e-mail

The people who contributed information or knowledge, the company for which they work, and the tools about which the provided information are listed below. In some cases the person was contacted to obtain a general idea of what is available in their area or tips for tools in a certain area. In those cases, “International monitoring” appears instead of the name of the tool. The list is arranged in alphabetical order by surname.

Behrang Assemi, The University of Queensland (Australia) – ATLAS II

Oliver Bermhagen, Meltspot

Hjalmar Christiansen, Technical University of Denmark, www.transport.dtu.dk – International monitoring

Michael Cik, Invenium Data Insights GmbH, http://www.invenium.io/en/ – Invenium Data Insights

Brett Davis, NuStats, http://www.nustats.com/ – Route Scout, NuTripX, etc.

Ulf Eriksson, EU project PASTA (physical activity through sustainable transport approaches), http://www.pastaproject.eu/home/ – Pasta Online Survey

Martin Fellendorf, Graz University of Technology, https://www.tugraz.at/home/ – International monitoring Enrico Howe, Innoz, https://www.innoz.de/ – Modalyzer

Kees den Hollander, Technolution www.technolution.eu – International monitoring

Adam Laurell, Samtrafiken – Samtrafiken's platform for MaaS services

Staffan Liljestrand, Bumbee labs, https://www.bumbeelabs.com/sv – Bumbee Labs

Magnus Lindhe, Skånetrafiken, The Point of Interest Project

Simon Moritz, Ericsson, www.ericsson.com – Ericsson platform, Mobile network data

Frank Ophuis, Mobidot, www.mobidot.nl – Mobidot/Sesamo/Smart

Lee Peters, Strava, https://www.strava.com/ – Strava Metro

Markus Ray, Austrian Institute of Technology (AIT) GmbH, https://www.ait.ac.at/ – AIT Smart Survey

Mårten Rignell, Skånetrafiken, The Minimum Viable Device Project

Angelika Schulz, Institute of Transport Research (Germany), http://www.dlr.de/vf/ – International monitoring

Ivana Semanjski, UGent (Ghent University), www.ugent.be, Route-Coach/M-App

Marcelo Simas, WeStat, https://www.westat.com/ – DailyTravel

Amritpal Singh, Viscando, www.viscando.com – OTUS3D

Remko Smit, Ministry of Infrastructure and the Environment (The Netherlands) https://www.rijkswaterstaat.nl/ – International monitoring

Yusak Susilo, KTH – on research into the TSs of the future and the MEILI tool

Clas Rydergren, Linköping University, www.liu.se – International monitoring

Fredrik Vastad, Facility Labs, http://facilitylabs.com/ – Facility Labs

Reference group:

Abboud Ado, Transport Analysis

Karin Björklind, City of Gothenburg

Leonid Engelsson, Swedish Transport Administration/KTH

Per Eriksson, Swedish Transport Administration

Jenny Eriksson, Swedish National Road and Transport Research Institute (VTI)

Andreas Holmström, Transport Analysis

Eva Lindborg, Transport Analysis

Simon Moritz, Ericsson

Gunnar Ohlin, Västra Götaland Region/Lindholmen

Sara Rogerson, VTI

Ary Silvano, VTI

Kimiko Sörensen, Samtrafiken

Martin Ullberg, Sweco

Charlotte Wahl, Sweco

Per-Åke Vikman, Transport Analysis

Mats Wiklund, Transport Analysis

Reference list

Abbasi, A., Rashidi, T.H., Maghrebi, M., Waller, S.T. Utilising location based social media in travel survey methods: Bringing twitter data into the play (2015)

Alexander, L., Jiang, S., Murga, M., González, M.C. Origin-destination trips by purpose and time of day inferred from mobile phone data (2015)

Anda, C., Erath, A., Fourie, P.J. Transport modelling in the age of big data (2017). Article in Press

Asakura, Y (Asakura, Yasuo); Hato, E (Hato, Eiji); Maruyama, T (Maruyama, Takuya). Behavioural Data Collection Using Mobile Phones. In: Rasouli, S; Timmermans, H Origin-Destination Data Collection Technology. In Mobile Technologies for Activity Travel Data Collection and Analysis. Advances in Data Mining and Database Management (ADMDM) Book Series

Assemi, B., Safi, H., Mesbah, M., Ferreira, L. Developing and Validating a Statistical Model for Travel Mode Identification on Smartphones (2016)

Berger, M and platzer M (2015). Field evaluation of the smartphone-based travel behaviour data collection app “SmartMo”. Transportation Research Procedia 11 (2015) 263 – 279

Bliemer, M. C.J., van Amelsfort, D. H. (2010) Rewarding instead of charging road users: a model case study investigating effects on traffic conditions, European Transport \ Trasporti Europei, 44, pp. 23-40

Bricka, SG (Bricka, Stacey G.); Simek, CL (Simek, Christopher L.); Baker, R (Baker, Richard Trey); Wood, N (Wood, Nicholas) (2014). In: Rasouli, S; Timmermans, H Origin-Destination Data Collection Technology. In Mobile Technologies for Activity Travel Data Collection and Analysis. Advances in Data Mining and Database Management (ADMDM) Book Series

Carrel, A; Lau, PSC; Mishalani, RG; Sengupta, R; Walker, JL (2015). Quantifying transit travel experiences from the users' perspective with high-resolution smartphone and vehicle location data: Methodologies, validation, and example analyses. TRANSPORTATION RESEARCH PART C: EMERGING TECHNOLOGIES; SEP, 2015, 58

Cesario, E., Iannazzo, A.R., Marozzo, F., Morello, F., Riotta, G., Spada, A., Talia, D., Trunfio, P. Analyzing social media data to discover mobility patterns at EXPO 2015: Methodology and results (2016)

Chen N C (2017). Urban data mining: social media data analysis as a complementary tool for urban design. MIT Thesis (Department of Architecture).

Chen, C., Ma, J., Susilo, Y., Liu, Y., Wang, M. The promises of big data and small data for travel behavior (aka human mobility) analysis (2016)

Chiang, MF (Chiang, MengFen); Chen, CC (Chen, ChienCheng); Peng, WC (Peng, WenChih); Yu, PS (Yu, Philip S.) (2014). Mining Mobility Evolution from Check-in Datasets. IEEE International Conference on Mobile Data Management 2014

Chua, A (Chua, Alvin); Servillo, L (Servillo, Loris); Marcheggiani, E (Marcheggiani, Ernesto); Moere, AV (Moere, Andrew Vande) (2016). Mapping Cilento: Using geotagged social media data to characterize tourist flows in southern Italy. Tourism Management 57

Cottrill, CD (Cottrill, Caitlin D.); Pereira, FC (Pereira, Francisco Camara); Zhao, F (Zhao, Fang); Dias, IF (Dias, Ines Ferreira); Lim, HB (Lim, Hock Beng); BenAkiva, ME (BenAkiva, Moshe E.); Zegras, PC (Zegras, P. Christopher) (2013). Future Mobility Survey Experience in Developing a Smartphone-Based Travel Survey in Singapore. TRANSPORTATION RESEARCH RECORD 2354

Davidsson P, Hajinasab B, Holmgren J, Jevinger Å, Persson J A (2017). The Fourth Wave of Digitalization and Public Transport: Opportunities and Challenges. Sustainability 2016, 8, 1248; doi:10.3390/su8121248

De Gennaro, M (De Gennaro, Michele); Paffumi, E (Paffumi, Elena); Martini, G (Martini, Giorgi. Big Data for Supporting LowCarbon Road Transport Policies in Europe: Applications, Challenges and Opportunities. BIG DATA RESEARCH 6

Di Lorenzo, G., Sbodio, M., Calabrese, F., Berlingerio, M., Pinelli, F., Nair, R. AllAboard: Visual exploration of cellphone mobility data to optimise public transport (2016)

Diao, M., Zhu, Y., Ferreira, J., Jr., Ratti, C. Inferring individual daily activities from mobile phone traces: A Boston example (2016)

Eriksson, J., Rogerson, S., Henriksson, P., Silvano, A.P., Sörensen, G., Nilsson, A., Wahl, C., Ullberg, M. & Adell, E. (2017). Intressent- och behovsanalys för resvaneundersökningar. Resultat från intervjuer och enkätundersökning. VTI-rapport 939-2017 (delrapport från AP 2 i Transport Analysis projekt kring att utveckla nya lösningar för resvaneundersökningar). [Stakeholder requirements from travel surveys. Results from interviews and questionnaire study. VTI Report 939-2017 (Subreport from AP @ in Transport Analysis project concerning the development of new solutions for travel surveys)]

Gould, J (2013). Cell Phone Enabled Travel Surveys: The Medium Moves the Message. In Zmud, J (Zmud, J); LeeGosselin, M (LeeGosselin, M); Munizaga, M (Munizaga, M); Carrasco, JA (Carrasco, JA) (2013): transport survey methods: best practice for decision making

Hasan S, Gonzalez M (2013). Spatiotemporal Patterns of Urban Human Mobility. Journal of Statistical Physics 151(1-2)

Holleczek, Thomas; Yu, Liang; Lee, Joseph K.; Senn, Oliver; Kloeckl, Kristian; Ratti, Carlo; Jaillet, Patrick (2013). Digital breadcrumbs: Detecting urban mobility patterns and transport mode choices from cellphone networks. Working paper

Inbakaran, C., Kroen, A. Travel surveys Review of international survey methods (2014)

Jiang S (2016). Mining big data for sustainable urban future. MIT PhD thesis

Kim, Y (Kim, Youngsung); Pereira, FC (Pereira, Francisco C.); Zhao, F (Zhao, Fang); Ghorpade, A (Ghorpade, Ajinkya); Zegras, PC (Zegras, P. Christopher); BenAkiva, M (BenAkiva, Moshe) (2014). Activity recognition for a Smartphone based travel survey based on cross-user history data. 2014 22ND INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION (ICPR)

Lu, S., Fang, Z., Zhang, X., Shaw, S.L., Yin, L., Zhao, Z., Yang, X. Understanding the representativeness of mobile phone location data in characterizing human mobility indicators (2017)

Lu, SW (Lu, Shiwei); Fang, ZX (Fang, Zhixiang); Zhang, XR (Zhang, Xirui); Shaw, SL (Shaw, ShihLung); Yin, L (Yin, Ling); Zhao, ZY (Zhao, Zhiyuan); Yang, XP(Yang, Xiping) (2017). Understanding the Representativeness of Mobile Phone Location Data in Characterizing Human Mobility Indicators. ISPRS INTERNATIONAL JOURNAL OF GEOINFORMATION 6(1)

Mearian, L (2016). Ford, MIT use Bostonians' cellphone location data for traffic planning. MIT

Nassir N, Hickman M, Ma ZL (2015). Activity detection and transfer identification for public transit fare card data. TRANSPORTATION; JUL, 2015, 42(4)

Nassir, N; Hickman, M;Ma, ZL (2015). Activity detection and transfer identification for public transit fare card data. TRANSPORTATION; JUL, 2015, 42 4

Pastinen, V., Keskikiikonen, H, Vallenius, J., Kivari, M. & Lehto, H: Paikannusmenetelmät osana henkilöliikennetutkimusta. Liikennevirasto, liikenne ja maankäyttöosasto. Helsinki 2017. Liikenneviraston tutkimuksia ja selvityksiä 18/2017. (Positioning methods as a part of personal traffic surveys. Swedish Transport Administration, land and ground use. Helsinki 2017. Swedish Transport Administration surveys and studies 18/2017. ISSN-L 1798-6656, ISSN 1798-6664, ISBN 978-952-317-381-1

Petrunoff NA; Xu H, Rissel C, Wen LM, van der Ploeg HP (2013). Measuring workplace travel behaviour: validity and reliability of survey questions. Journal Of Environmental And Public Health [J Environ Public Health] 2013; Vol. 2013

Poonawala, H., Kolar, V., Blandin, S., Wynter, L., Sahu, S. Singapore in motion: Insights on public transport service level through farecard and mobile data analytics (2016)

Rashidi, Taha H.; Abbasi, Alireza ; Maghrebi, Mojtaba ; Hasan, Samiul; Waller, Travis S. Exploring the capacity of social media data for modelling travel behaviour: Opportunities and challenges. Transportation Research: Part C. Feb2017, Vol. 75

Richard A. Beckery Ramón Cáceresy Karrie J. Hansony Sibren Isaacman_Ji Meng Lohy Margaret Martonosi_ James Rowlandy Simon Urbaneky Alexander Varshavskyy Chris Volinsky (2013). Human Mobility Characterization from Cellular Network Data. Communications of the ACM. Jan2013, Vol. 56, Issue 1

Richard J. Lee; Ipek N. Serner; James A. Mullins (2016). An evaluation of emerging data collection technologies for travel demand modeling: from research to practice. Transportation Letters 8(4)

Safi H, Assemi B, Mesbeh M, Ferreira L (2017). An empirical comparison of four technology-mediated travel survey methods. Journal of traffic and transportation engineering. Article in press

Safi, H., Assemi, B., Mesbah, M., Ferreira, L., Hickman, M. Design and implementation of a smartphone based travel survey (2015)

Safi, H., Mesbah, M., Ferreira, L. ATLAS project Developing a mobile-based travel survey (2013)

SalasOlmedo, María Henar; Quezada, Carolina Rojas (2016). Mapping mobility patterns to public spaces in a medium-sized city using geolocated tweets. Working paper

Shafique, M.A., Hato, E. Travel mode detection with varying smartphone data collection frequencies (2016)

Shen, Li, Stopher, Peter R. (2013). A process for trip purpose imputation from Global Positioning System data. Transportation Research Part C November 2013(36)

Shi X, Lin H (2014). The analysis of bus commuters' travel characteristics using smart card data: the case of Shenzhen, China. Transportation Research Board 93rd Annual Meeting

SilaNowicka, K (SilaNowicka, Katarzyna); Vandrol, J (Vandrol, Jan); Oshan, T (Oshan, Taylor); Long, JA (Long, Jed A.); Demsar, U (Demsar, Urska); Fotheringham, AS (Fotheringham, A. Stewart) (2015). Analysis of human mobility patterns from GPS trajectories and contextual information. International Journal of Geographic Information Science 30(5)

SiłaNowicka, K; Vandrol, J; Oshan, T; Long, J A; Demšar, U; Fotheringham, A S (2016). Analysis of human mobility patterns from GPS trajectories and contextual information. International Journal of Geographical Information Science. May 2016, Vol. 30, Issue 5

Sobolevsky, S., Sitko, I., Des Combes, R.T., Hawelka, B., Arias, J.M., Ratti, C. Money on the move: Big data of bank card transactions as the new proxy for human mobility patterns and regional delineation. The case of residents and foreign visitors in Spain (2014)

Soltani, K., Soliman, A., Padmanabhan, A., Wang, S. UrbanFlow: Largescale framework to integrate social media and authoritative landuse maps (2016)

Statistics Sweden. 2016. Kvalitet för den officiella statistiken – en handbok. [Quality of the Official Statistics – A Handbook]

SWECO, 2013. New input data sources for traffic forecasts

Tosi D, Marzorati S (2016). Big Data from Cellular Networks: Real Mobility Scenarios for future Smart Cities. PROCEEDINGS 2016 IEEE second international conference on big data computing service and applications

Wang Z, He S, Leung Y (2016). Applying mobile phone data to travel behaviour research: A literature review. Travel Behaviour and Society July 2016 (HKSTS)

Wang, Q., Taylor, J.E. Massive online geosocial networking platforms and urban human mobility patterns: A process map for data collection (2014)

Wang, Q., Taylor, J.E. Process Map for Urban-Human Mobility and Civil Infrastructure Data Collection Using Geosocial Networking Platforms (2016)

Wang, X., Wong, A.K.S., Kong, Y. Mobility tracking using GPS, WiFi and Cell ID (2012)

Wenjie Wu ;Jianghao Wang; Tianshi Dai (2016). The Geography of Cultural Ties and Human Mobility: Big Data in Urban Contexts. Annals of the American Association of Geographers. 2016, Vol. 106, Issue 3

Vlassenroot, S., Gillis, D., Bellens, R., Gautama, S. The use of smartphone applications in the collection of travel behaviour data (2015)

Wolf, J (Wolf, Jean); Wilhelm, J (Wilhelm, Jeremy); Casas, J (Casas, Jesse); Sen, S (Sen, Sudeshna) (2013). A Case Study: Multiple Data Collection Methods and the NY/NJ/CT Regional Travel Survey. In Zmud, J et al: transport survey methods: best practice for decision making

Yang Xu, Shih-Lung Shaw, Ziliang Zhao, Ling Yin, Feng Lu, Jie Chen, Zhixiang Fang, and Qingquan Li (2016). Another Tale of Two Cities: Understanding Human Activity Space Using Actively Tracked Cellphone Location Data. Annals of the American Association of Geographers Vol. 106, Issue 2, 2016