United Nations ECE/TRANS/WP.6/2023/7

Economic and Social Council

Distr.: General
3 March 2023

Original: English

Economic Commission for Europe

Inland Transport Committee

Working Party on Transport Statistics

Seventy-fourth session
Geneva, 15–17 May 2023
Item 3 of the provisional agenda
Big Data and innovation in transport statistics production

Using Mobile Phone Data for Transport Statistics

Note by the secretariat

Summary

This document highlights the role of the secretariat in increasing knowledge around mobile phone data (MPD) in the field of transport statistics, as part of international efforts, in particular through the forthcoming handbook on using MPD in transport statistics as part of the UN Committee of Experts on Big Data Task Team on MPD.

I. Background

1. Transport data is crucial to understand how people and goods are moved across various transport modes, and how many people have access to each transport mode. It helps to better plan and organize transport systems, and further feeds in to measuring other areas such as access to job and services, climate and energy goals, gender equality in mobility and road safety.

2. Until now, transport data have come from traditional sources such as travel surveys, administrative sources and direct measurement such as traffic counters. These sources normally have a high degree of accuracy, but their main downsides relate to timeliness, cost and granularity. Mobile phone data (MPD) may play an important role in achieving transport data that is up to date, cost-effective, and of high quality with rich, detailed information. MPD has its own limitations, but also a number of advantages and potentials.

3. There has been an increasing interest in MPD since the recent COVID-19 pandemic. MPD allowed continuous mobility monitoring and analysis during the pandemic, while other traditional data collection has been disrupted due to COVID-19 induced measures and restrictions. Several countries see potentials in using MPD and have already started to use MPD in transport statistics, typically in an experimental or temporary way. Separately from (or in complement to) statistics production, MPD can also be used for transport planning

purposes, for example in ascertaining where transport infrastructure or public transport routes should be planned, and what time of day public transport services would be most effective.

4. In response to increased interest in MPD in general, the UN Committee of Experts on Big Data and Data Science for Official Statistics established a Task Team[1] specifically to look at MPD. The task team has six subgroups, covering displacement and disaster statistics; dynamic population mapping; measuring the information society; migration statistics; tourism statistics; and transport and commuting statistics. The first five subgroups have published methodological guides in 2022, and the aim is that a similar guide or handbook can be published for transport in 2023. The transport handbook will be able to draw on these existing methodological guides to cover issues related to MPD that are not specific to transport; for example, issues related to data access, data processing and ensuring individual confidentiality are typically not domain-specific.

5. Given the experimental nature of much transport statistics production using MPD to date, it would not yet be possible to provide step-by-step guidance on best practices to follow, or to establish internationally accepted methodological guidance, in the handbook. Instead, the handbook aims to describe where MPD could be used in transport statistics production, and to examine existing use cases provided by countries in detail. This would allow countries in the first stages of MPD to learn about processes that other countries have already undertaken.

6. The handbook should hopefully be ready in time for the third quarter of 2023. But this first handbook should not be considered the end product. Updates can be made over time as countries' capacity in using MPD increases and thus the available use cases develop.

II. Structure of proposed manual

7. The proposed manual includes six sections: introduction, transport data basics, MPD basics, advantages of using MPD for transport, the case for official statistics, and detailed case studies. The manual only considers data stored from mobile network operators (MNO data), rather than GPS data, data obtained through an app, or other MPD sources. The manual illustrates where MPD may have uses in improving transport and mobility, urban planning, and monitoring the transport related Sustainable Development Goals.

8. The manual illustrates the types of information included under MPD, such as Call Detail Records (CDR), Passive Signaling Data, Active Signaling Data, the type of event (call, text, or data usage), and the call duration. It notes that MPD is valuable as it allows collection of frequent, granular and quality data, and it enables various analysis, when used in combination with other data sources, such as traffic counters, road network routing, public transport ticketing and travel surveys.

9. The manual contains information on MPD processing, that is necessary to transform raw data into a state that is feasible for analysis, as well as integrating MPD with other data sources. It emphasizes the importance of following data privacy and confidentiality rules. The manual also covers limitations of MPD. MPD can contain inherent biases due to representative issues, and precision could be limited by cell tower density, which is not a major problem for long-distance journeys but can be relevant for shorter walking and cycling journeys. It may also require big infrastructure and long data processing procedures. Another issue is duplications of data as some individuals own more than one sim card.

10. The manual recognizes that MPD is not about to replace traditionally gathered transport statistics in the near future. Some transport statistics sources already have excellent levels of coverage, accuracy, timeliness and affordability, such as vehicle registers and traffic counters, and so MPD may only have a limited role, if at all, in these fields. Other sources, such as travel surveys, have a greater cost and longer timelag, but provide crucial information on trip purpose and traveller characteristics that cannot be ascertained from other sources. Rather, it may be that MPD can complement travel surveys with estimates of trip information

[1] https://unstats.un.org/bigdata/task-teams/mobile-phone/index.cshtml.

for smaller geographical levels, on a more timely basis, or for modes that are not widely integrated into transport statistics such as walking and cycling.

11. The case studies in the manual will include the use of MPD in analysis for public transport passengers in Austria and Netherlands, analysis for road traffic and origin-destination metrices in the United Kingdom, in walking and cycling statistics, and in urban planning in Estonia, Republic of Korea and United Arab Emirates. The next section describes three of these examples.

III. Examples of case studies to be included in the Handbook

United Kingdom

12. Prior to the coronavirus pandemic, the UK Department for Transport (DfT) collected mode specific data on transport users through methods such as surveys, automatic counters, ticket sales, and mobile phone data only occasionally. However, understanding the impact of coronavirus travel restrictions affected every mode of travel and wasn't something that DfT has been faced with before, so a new approach was needed. For this they worked with providers to access anonymised and aggregated mobile phone data and combined this with other government sources (such as the census). The resulting dataset, along with existing modal datasets, formed the basis for the analysis and insights needed to enable effective national mobility monitoring both during and after each lockdown.

13. MPD can provide additional information on travel by time of day, allowing analysis on peak travel times across travel modes. It also provides more detailed information on demographics of transport users and trip chains. Thus, it helps the UK better understand the demographics of users of the rail networks and how it varies by time of day and throughout the week. It has supported national crises with real-time MPD to monitor whether the transport network was overloaded. Going forward, DfT plans to use MPD to better understand users of Electric Vehicles and evaluate policies that facilitate the use of the vehicles.

14. DfT purchases MPD, that is processed, anonymized and aggregated, directly from a mobile network operator. The purchased data must comply with data protection laws and small numbers of trips on a very granular basis cannot be provided. The processed mobile data vary in granularity, but typically are at daily or hourly level and usually cover 2-3 months. However, it can also include real-time data, including information for 5-minute intervals with around 5 minutes' latency over the period of a week.

15. Data quality was ensured by procuring consistently formatted and processed datasets from the same mobile network operator each time. In addition, a new MPD product is compared with the data from existing published and internal statistics. Finally, MPD is not used in isolation but typically is combined with datasets with other statistics such as census data. The overall impression has been that the data largely showed similar trends to other sources, and where differences do exist these have been explainable by the collection methods used and caveats associated with one or more of the sources. However, the MPD were less well suited to measuring smaller trips, for example local walks.

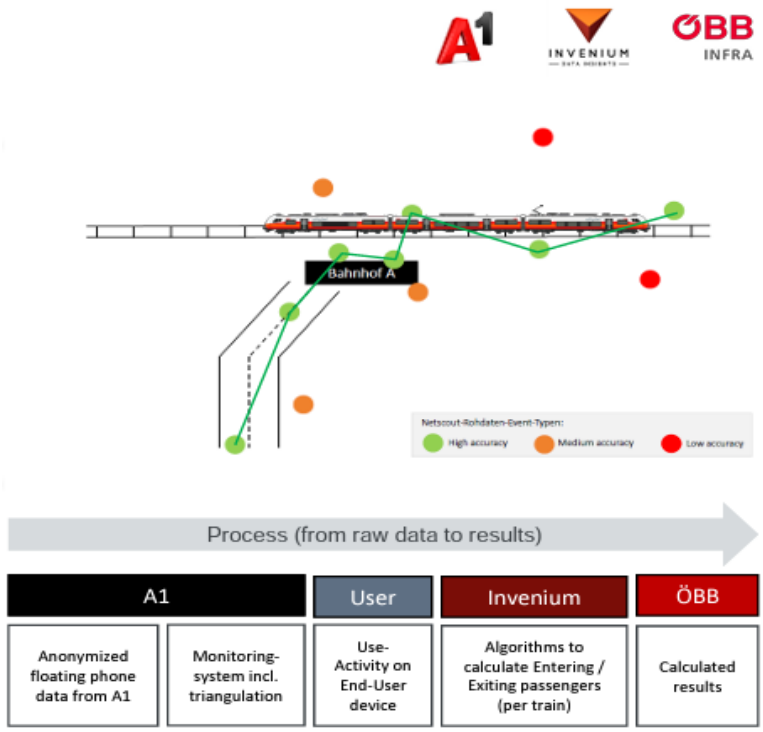
16. The DfT are not using MPD to produce official statistics, but instead it is being used as part of a suite of evidence and analysis to support policy decisions and internal reporting. While it is unlikely to become part of official statistics production, largely due to the costs involved and processes that would need to change, it increasingly frequently is used to support analysis on the evaluation of policies across multiple modes.

Austria

17. Invenium Data Insights GmbH, a private company in Austria, actively uses MPD in transport statistics, planning, operations and regulations. They acknowledge that even though MPD has its own limitations, it has a big advantage as the data is available area-wide and on a daily basis. Invenium has used MPD to analyse daily passenger demand of rail related

transport in Austria. The analysis provided basis for internal planning purposes at Austria Federal Railways. In addition, Invenium also uses or plans to use MPD to analyse the following cases: entering and exiting passengers per station, a station-based Origin-Destination-Matrix, loads of passengers on defined cross sections, analysis of delays and demand peaks.

Figure
Example of how Invenium uses MPD for counting rail passengers



18. Invenium receives anonymized MPD from A1 Telkom Austria, and uses their algorithm platform to estimate the activity of transport users and their transport modes (see the figure above). Invenium also uses actual train timetable data on a daily basis, provided by OBB Infrastruktur, who obtains this from the Advanced Railway Automation Management Information System (ARAMIS). The MPD is managed in compliance with the strict data protection guidelines in the European Union. In addition, micro data is not available to the end users of MPD, and they can access aggregated results only.

19. Invenium implements quality monitoring throughout the entire processing to ensure data quality. In addition, they have implemented around 2000 empirical counting points to validate its calculation models.

United Arab Emirates

20. Roads and Transportation Authority (RTA) of Dubai in the United Arab Emirates uses MPD data to better understand the commuters' travel activity patterns. By analysing the trip origin and destination, they aim to improve the mobility in the city by improving effectiveness and efficiency of the exiting transportation services. MPD allows RTA to plan public transport routes to serve real time demand.

21. MPD is often used along with other data sources (bus data, internet map data, social media data, metro data) to analyse travel demands. RTA notes MPD has its unique advantages as it provides information on detailed demographics of transportation users, geospatial mobility data, and allows examination of travel patterns.

22. RTA receives anonymized MPD data from Locatium, the geospatial data science platform. RTA do not have access to raw data and data is managed following data regulations and laws. To ensure the qualities of MPD data, they compare and crosscheck the data with existing transport statistics.

IV. Future work and conclusions

23. MPD has advantages and can enable in-dept analysis in various fields including transport and urban planning and has potential to be used in official statistics. It would be beneficial to further facilitate sharing the case studies of mobile phone usage across countries and build international standards and rules to improve the comparability of MPD outputs across countries.

24. WP.6 is invited to reflect on the structure of the handbook and provide comments. In particular, the adding of existing use cases for MPD in transport statistics would be particularly welcome, even if these use cases are not yet fully fledged. Further collaboration around this topic is welcome; this could include further demonstrations of MPD use in transport during WP.6 sessions, or also informal webinars according to demand.