Observing Road Freight Traffic from Mobile Network Signalling Data While Respecting Privacy and Business Confidentiality

Rémy Scholler, Oumaïma Alaoui-Ismaïli, Jean-François Couchot, Eric Ballot, Denis Renaud

Observing road freight traffic from mobile network signalling data while respecting privacy and business confidentiality

Rémy Scholler[1,2], Oumaïma Alaoui-Ismaïli[1], Jean-François Couchot[2], Eric Ballot[3], and Denis Renaud[1]

[1] Orange Labs, 92320 Châtillon, France
[2] Femto-ST Institute, DISC Department, UMR 6174 CNRS, University of Bourgogne, Franche-Comté, Besançon, 25000, France
[3] Mines ParisTech, CGS—Centre de Gestion Scientifique, 75272 Paris Cedex 06, France

Abstract. Nowadays, there is no tool that provides a global, permanent and "real time" view of road freight transport flows. However, this type of mapping is already available for air and sea traffic and could be useful to transport companies, e.g., setting up logistics hubs in strategic locations, and to public authorities, e.g., quickly knowing the impact of regulations, the contribution to congestion, or the impact of emissions. This kind of tool could obviously make information about road freight traffic more accessible, and allow for the consolidation of flows at both the interurban and urban levels to help decarbonize freight transport and logistics. The main contribution of this paper aims to provide a design sketch of an observatory of road freight transport flows based on signalling data from mobile network, which is accurate enough for that type of study and which does not require any supplementary installation of application on mobile devices. This kind of observatory is therefore related to the concept of Physical Internet through its objectives. This observatory will have to ensure privacy and business confidentiality by respecting the constraints set by the General Data Protection Regulation (GDPR) and the ePrivacy directive, i.e., a short-term anonymization in the French case. Thus, the second contribution of this paper is a literature review on the methods that could be useful to solve these questions.

Keywords: Privacy · Anonymization · Business Confidentiality · Mobile Network Data · Road Freight Transport

1 Introduction

Roads are currently the most common freight transport mode in France. Indeed, in 2020, the ton.kilometer share of road transport represented 89.1% of land transport excluding pipelines, while rail transport represented 9.0% and inland waterway transport 1.9% [11]. However, there is less public data and global

models for road freight transport compared to other sectors (air or maritime transport). This limits the action of public actors, potentially represents a limit for road freight transport actors and restricts academic research to a fragmented vision.

This work aims to provide a design sketch of a dynamic observatory of road freight transport in France. The goal of such an observatory is to obtain a near-real-time inventory of freight transport flows across the country and an end-to-end vision of these flows. This observatory will initially focus on interurban transport, and then on urban transport, as most urban deliveries are made by light commercial vehicles, which are harder to identify because they can be related to a wider variety of behaviours and uses (hence the less strict regulations for the latter). It is based on signalling data from Orange's mobile network collected by observation probes on a regular basis. The signalling data of a device, also called mobile trace, is a time-stamped sequence of events (usually calls, SMS, data connection, network re-selection) occurring on a network antenna whose position is known. It is basically trajectory micro-data as defined in [17], i.e., information about single individuals that describe their spatiotemporal trajectories. These trajectories are sequences of geographical positions of the monitored individuals over time except that here we do not have access to precise positions but to network cells positions, i.e., the positions of the network cells (defined by an antenna of the network and its coverage area) which the user connects to through time.

Signalling data have already proven their great potential in different fields [6], especially in the study of human mobility [28] and road traffic [7]. In particular, some works using signalling data have tried to classify vehicle types [20], to study congestion and traffic states [19, 15] or even to estimate vehicle speeds [10]. However, these kind of studies are often very local, e.g., they focus on a highway segment or on one city center and to the best of our knowledge there is no work targeting a business sector and the vehicle uses related to this sector as we are trying to do with the road freight transport sector.

The observatory sketched here aims to improve the accessibility and quality of information on transport flows across all the country. In the long run, it could help to promote the best behaviours (reduction of empty trips, pooling, logistics networks interconnection) and thus increase the efficiency and sustainability of freight transport. These objectives match those of the Physical Internet concept [3], which is, in logistics, an open global logistics system founded on physical, digital, and operational interconnectivity, through encapsulation, interfaces and protocols, intended to replace current logistical models. The aim of the sketched observatory is not to model [25] or simulate freight transport on a national scale [12], but to observe it, by collecting a large amount of quality data and generate statistical indicators while respecting privacy and business confidentiality. This objective can be divided into two parts:

- Provide global statistics at the finest possible spatial-temporal granularity while respecting the anonymization constraints imposed by the GDPR, the ePrivacy directive and their respective versions in France, Spain and Belgium

to ensure privacy and business confidentiality. This is the main objective and would constitute a dynamic real-time observatory of road freight transport flows nationwide.
- Provide more specific statistics with consent of the interested industrial actors. This would allow specific marketing analyses in comparison with the global data of the observatory (carbon footprint estimation, estimated market share, etc.) and would lead to partnership optimizations between economic actors.

The main objective of the observatory of freight transport flows is to show counting of trucks on various origin-destinations, zones of interest or road segments, augmented with various statistical indicators such as counts or density in a logistic zone or on a specific road, pollution estimations, or some points of interests. To achieve this goal an interesting way is to compute origin-destination matrices [8] for road freight traffic that ensure privacy, using probabilistic data structures for example [1]. In order to produce useful estimates for the transport sector, it is necessary to propose a methodology for classifying objects and their behaviours. We have to propose an algorithm that can both predict the class (e.g., truck, Light Commercial Vehicle) of a new object while assigning it to a group of objects of the same class with similar spatial-temporal behaviours [26]. This methodology would be based on supervised and unsupervised incremental clustering techniques [4].

Particular attention will be paid to compliance with the requirements of the European GDPR and the new ePrivacy directive [16]. Especially in France, attention will be paid to compliance with the "Loi Informatique et Libertés" and the "Code des Postes et Télécommunications Electroniques". In particular, all the algorithms must be applicable to a history of events whose retention period is limited by a legal constraint. The legal constraints due to the GDPR and more specifically by the ePrivacy directive in Europe could be different from one country to another but in France, mobile signalling events can be processed only if an irreversible short-term anonymization is carried out. The short-term anonymization terminology refers to a French legal constraint imposed by the "Commission Nationale de l'Informatique et des Libertés" (CNIL). In our case, we have a "short" time to do all the necessary treatments to go from raw signalling data to the publication of aggregate statistics that respects privacy and business confindentiality, and then delete raw data (which is personal). This "short" time is a result of a negotiation with the CNIL, and for works about human mobility that could be useful for cities, public organizations they usually give a time of 15 minutes. Moreover, in the case of signalling data, when users appears in the dataset we only have access to a "short" history of their personal data (usually around 15 minutes too). This short-term anonymization constraint does not exists in other European countries as Spain and Belgium, so when we will reach our goal in France it will be easier to adapt our work to other countries.

This rest of this paper is organized as follows. Section 2 describes some related work concerning privacy and business confidentiality in our case, section 3

presents our methodology to solve the problem, and section 4 draws conclusions and outline areas of future work.

2 State of the Art

Signalling data concerns all types of connected objects such as connected watches or industrial production equipment but we do not use signalling data from drivers' devices or transport management systems (TMS) data. In fact, we use in-vehicle IoT modems (2G/3G/4G) which are devices that receive wireless data from remote sensors and forward these data to a different communications format. However, using signalling data corresponding to road freight vehicles remains a privacy issue, even if this is not the data from the vehicle driver's phone. For example, if an attacker knows the location of a freight vehicle, he knows indirectly the location of its driver. The common practice of pseudonymization approach which consists of removing identifiers and replacing them with dummy identifiers is not sufficient. More precisely it has been shown that a small number of locations can be used to identify individuals with a high probability [14]. A difference can be made between two ideas of privacy protection: the protection of business confidentiality and the protection of personal information. To understand the problem of business confidentiality it is important to know that the trips and routes of freight vehicles are essential in the business model of freight companies. For example, well known homogeneity issues can occur if multiple trucks of the same company are counted on a road segment. As another example, we could consider a warehouse, and the freight flows flowing in. If an attacker deduces all the business partners of this warehouse thanks to the origins of these freight flows, there is a confidentiality issue. All the statistics published from signalling data have to be anonymized [13] and in France the methods used for creating the observatory of freight flows have to satisfy the short-term anonymization constraint defined in Section 1, be applied in real time and be adapted to streaming data.

We are close to a situation of privacy-preserving data publishing (PPDP) of trajectory micro-data databases which recommends that databases should be transformed prior to publication in potentially hostile environments, so as to grant that the published data remains useful while individual privacy is preserved [17]. Therefore literature in this domain is a good start for our work. Fiore et al. in [17] explains that in the case of trajectory micro-data publishing, databases of millions of records are mined offline, and the challenge is ensuring that their circulation does not pose a threat to user privacy, but retains data utility. In the case of Location Based Systems (LBS), single (geo-referenced and time-stamped) queries generated by mobile devices must be processed in real-time, and the objective is location privacy, i.e., ensuring that such a process preserves users' privacy by preventing the service provider from locating users. This difference leads to very diverse attacker models and anonymization techniques for the two scenarios. Indeed, Xiao and Xiong [29] and Bindschaedler et al. [5] have shown that individual spatiotemporal points anonymized via solutions for

location privacy are still vulnerable to attacks when their time-ordered sequence is considered, i.e., when they are treated as a spatiotemporal trajectory. Fiore et al. propose in [17] the first survey that provides a literature overview that comprehensively addresses trajectory micro-data privacy. They explored the attacks against trajectory micro-data that allow re-identifying users, the anonymization of trajectory micro-data, i.e., the counter measures against privacy threats, and discuss open issues and research opportunities.

However, we may not try to publish anonymized trajectory databases but databases containing origin-destinations, statistical indicators and aggregate statistics about the original trajectory micro-data databases instead. Moreover, our techniques must satisfy a short-term anonymization constraint and will have to be applied in real-time to streaming data. Therefore, the main problem we are trying to address is the privacy preserving data publication of aggregate statistics from trajectory micro-data taking in account spatial and temporal dimensions, adapted to the case of streaming data (in real time) and respecting short-term anonymization constraints. We could add to this problem statement that the trajectories should be constrained by a road network.

To the best of our knowledge, the two main methods that could be adapted to achieve our goal are differentially private synthetic trajectory datasets preserving the statistics of originals datasets [23, 18] and differential privacy methods adapted to streaming data [21, 9, 27]. In what follows, we develop some ideas used to generate some differentially private trajectory datasets.

As explained in [17], differential privacy can be ensured by a different process where some representation of the original trajectory micro-data is randomized so as to meet differential privacy constraints, and synthetic trajectories are derived from such representations. Then, databases of synthetic trajectories can be distributed with strong privacy guarantees. The two main approaches here are representing trajectory datasets as trees [18] or as probability distributions [23].

The first idea is to model the original database as a prefix tree, i.e., a hierarchical structure where trajectories are grouped based on matching location subsequences whose length grows with tree depth. A privacy-preserving version of the prefix tree is then obtained by considering multiple levels of spatial generalization, and adding noise to the nodes. Following an iterative process nodes are created for all locations at the highest level of generalization, as children of each leaf from the previous iteration. Then, Laplacian noise is added to the count of trajectories associated to each generalized node at the current prefix tree layer. Finally, nodes with a noisy count below a tunable threshold are not expanded further, while nodes with noisy counts above threshold generate children nodes for all locations at the following level of generalization. The process is repeated from the second step above until a user-defined tree height is reached with Laplacian noises set so that the total privacy budget is equally divided across all tree and nodes. The tree is then pruned so that only nodes at the lowest level of generalization are preserved. The noisy counts associated to such nodes are made consistent across levels, ensuring that the count of each node is not less than the sum of counts of its children nodes. Finally, the synthetic trajectories are

generated by visiting the resulting prefix tree. He et al. in [18] demonstrate that the approaches above work well with coarse trajectories defined on small location domains, but fail to scale to realistic database with large geographical span. Therefore, the authors propose to generate multiple prefix trees, each referring to a different spatial resolution. Each transition in a trajectory contributes to one specific tree, based on the travelled distance (i.e., low-resolution trees for long distances, and high-resolution trees for short distances). This results in multiple trees with a very small branching factor each, and in a significant reduction of the overall number of counts maintained. Then, the usual procedure of adding Laplace noise to counts, pruning the prefix trees, and extracting the synthetic trajectories is followed. In this last step, the authors also adopt an original sampling technique that allow preserving the correct directionality in the output trajectories. The proposed solution, named Differentially Private Trajectories (DPT), is evaluated with both real and synthetic datasets that are queried for distributions of diameters and trips, and for frequent sequential patterns.

The second idea is to create a differentially private synthetic trajectory generator that does not rely on a tree model of the original trajectory micro-data. Instead, DP-WHERE [23] performs the following steps: derives a number of distributions that describe different statistical features of the movements in the original trajectory database, such as the spatial distribution of home and work locations, or the number of spatiotemporal points in a trajectory; adds Laplacian noise to such distributions; extracts realizations from the noisy distributions to generate synthetic trajectories. The synthetic movement data produced by DP-WHERE is proven to preserve population density distributions over time, as well as daily ranges of commutes in the reference area.

Orange already uses methods that satisfy a short term anonymization constraint to calculate aggregate statistics of mobility from signalling data. These methods are mainly based on structures of probabilistic sets and k-anonymity [24], and satisfy a short term anonymization constraint at every steps of the process (e.g., during the creation of probabilistic sets that respects k-anonymity and when publishing aggregated statistics). The limits of these methods could be usual attacks to break k-anonymity by using extended knowledge or by combining some of the probabilistic sets created. However, the data used to create statistics is often not precise spatial-temporal location but is blurred in time (due to short-term anonymization constraint) and space (a certain area size) so even if an attack is successful, the utility of this data that leaks seems not very high. Moreover, raw data and created data are secure at every step of the process. It could be very useful to dig into these methods and try to create probabilistic sets such as in [1, 2] in order to calculate aggregate statistics with union, intersection of these sets, or tests of membership in them.

3 Methodology

In the following section, we sketch an approach to create an observatory of road freight transport flows which respects privacy and business confidentiality. First,

we describe globally the areas of interest, then we present the different phases and technical steps of our project work, and finally we discuss how we intent to solve the short-term anonymization problem in our case.

3.1 Areas of interest

We study in parallel the different areas listed below.

The quality of the raw material, i.e., the IoT modems traces: spatio-temporal uncertainties, sample characteristics, perspectives of evolution, pros and cons compared to other data sources.

The legal framework of signalling data processing and its evolution: GDPR, ePrivacy directive, and their respective versions in France, then in Spain and Belgium. The differences in practice depending on the context, research or operational. The feasibility of real-time processing for short-term anonymization.

Modelling steps needed to increase the data source's utility: increase in spatiotemporal accuracy, dynamic correction of sample biases, data science compatible with "On Line" and "Off Line" processing.

Software processing tools: finding a suitable software stack (big data and data science), defining target infrastructures to provide the service.

Business Exploration: be able to make demos, prototypes to target B2B, B2G or B2R, carry out discussions on various business plans.

3.2 Phases and technical steps

There are two main phases of technical work which are presented below.

Research phase: This is the design of the processing of pseudonymised data awaiting for anonymization. The aim is to define the main operations that will serve as specifications for the transition to operations. The use of pseudonymised data is allowed only in the research phase, because this method of privacy protection is weaker than the anonymization process (irreversible by definition).

Operational transition phase: This is the design of the real-time data processing for short-time anonymization. Here, we seek to satisfy the specifications of the first phase under a constraint of short time anonymization.

Our methodology for estimate and visualise freight traffic is summarised in Figure 1. Then, we detail the six different steps.

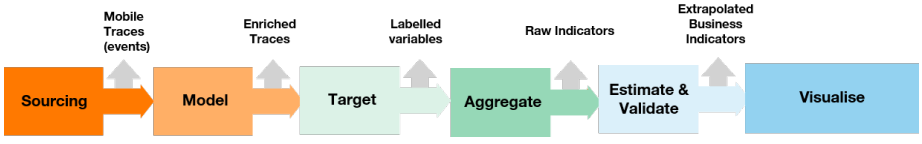

Fig. 1. *Technical steps of the sketched approach.*

Source: Design of a dataset capture and processing architecture. Adapt this architecture for the operational transition phase. The output of this part is a dataset containing IoT modems traces (events).

Model: Modelling on longitudinal data with the aim of reducing spatial uncertainties, estimating mobility conditions, and creating more reliable trajectories from mobile network data. The work carried out in this step uses simulations of signal propagation and probabilistic mapping of devices' pick-up by network antennas. GPS tracks from fleets of consenting partners can be used as validation data. We then obtain enriched IoT modems traces (more precise position, more precise speed, in particular). This part has been the main area of interest for this first year of work, and two patents have been filed. However, we do not detail this work here because it is not related to privacy and business confidentiality.

Target: Creation of trips, mapping, prediction of vehicles types and identification of behaviours and uses. In this part, we can use supervised incremental clustering methods. We thus obtain groups of traces associated with an origin, a destination, places and times of breaks, a use, a type of vehicle, and possibly other behavioural variables. We can obtain business variables (which will characterise uses) and useful variables for the estimation of indicators by group (for example, travel speed or intensity of signalling on the network). We could also use GPS data from willing partners to validate this part (trip creation, spatiotemporal mapping). This part has been the second mail area of interest for this first year of work, and we are now able to detect some typical behaviors for vehicles and we will link these behaviors with vehicle types (heavy trucks, light freight vehicles) thanks to labelled data from partners.

Aggregate: Use clustering methods to generate new groups of traces that satisfy anonymization constraints. Compute indicators corresponding to the different aggregates, maximise the aggregates' meaning, and construct specific aggregates for subsequent correction. This aggregation allows us to obtain raw indicators,, i.e., statistics representative of certain types of vehicles and uses.

Estimate and Validate: Creation of reference spatio-temporal data, development of a model for the transition from the sample studied to the total population, errors estimations and validation, feedback on the parameters of the "Aggregate" and "Target" parts. The additional data useful for the transition from the sample studied to the total population, for the validation and for the errors estimation would be, for example, ticketing data or usual traffic measurement data [22] as magnetic counting loops. Through statistical indicators and a database of aggregates associated with origins, destinations, locations and break times, uses, type of vehicles, etc, we obtain an estimate of the overall freight traffic across the country.

Visualise: Show data and indicators obtained in the previous part, carry out a state of the art on the specific data visualisation of moving connected object flows and find a suitable solution for our visualisation question. Then make adjustments in contact with potential users (e.g. economic actors, cities, freight transport companies) in order to make sure that the tool developed is useful for

them. Finally, find additional data sources that may be useful for the analysis and the calculation of new indicators such as pollution or congestion.

3.3 Short-term anonymization considerations

The database obtained at the end of the "Estimate and Validate" part and the method to create it must respect privacy and business confidentiality. It is possible to integrate anonymization constraints into several steps, e.g., adding noise in an origin-destination matrix (with uses and other behavioural variables) created in the "Target" part, adding noise in the clusters of uses in the "Target" or "Aggregate" part, adding noise within each aggregate in the "Aggregate" part, adding or leaving noise in the enriched IoT modems traces resulting from the "Model" part. Privacy constraints are easier to apply in the "Aggregate" part. We could use aggregates based on k-anonymity, methods based on differential privacy or any other method based on those cited in Section 2 such as [23, 18] because our datasets are quite similar to the ones used in these approaches. The main adaptations to these two methods could be to take in account the temporal dimension in trees or distribution approaches and to constrain all trajectories in a road network. However, because of the research experience at Orange in probabilistic sets that respect privacy, we will probably concentrate on approaches mixing probabilistic sets and differential privacy, as in [1, 2]. The kind of attacker we will consider is one with nearly unlimited computing power, and we want to prevent reconstruction attacks, i.e., if an attacker have some side knowledge (for example he knows that a user is in the raw data and have some of his locations), can he reconstruct his trajectory with the aggregate statistics that we publish? Concerning the evaluation of our approach, the private counts of trucks should be "not too far" from the real counts, at every time slot (temporal granularity still to be defined), and we have multiple choices of metrics to measure a distance between those counts and define an appropriate threshold. Unfortunately, we cannot make clear the time we can retain data legally because it depends on the CNIL decision. However, we can consider the time granted usually for this type of applications, which is around 15 minutes.

4 Summary and Future Research

In this paper we proposed an approach to create an observatory of road freight transport flows based on traces on the cellular network of IoT modems in vehicles. This observatory has to respect anonymization constraints to protect privacy and business confidentiality. The work is divided in two parts: a research phase and an operational phase. We proposed a methodology in six steps (Source, Model, Target, Aggregate, Estimate and Validate, Visualise) to create such an observatory during the research phase. For the moment, we have made good progress in speed, direction and mobility state estimations from signalling data (in the "Model" step), in identifying typical behaviours of vehicles in our datasets (in the "Target" step), in identifying and studying various data sources

that could be useful to correct and validate our estimations on freight traffic (in the "Estimate and Validate" step), in visualisation of moving objects' flows (not associated to freight transport for the moment) and in the calculation of some basic indicators about these flows.

This work will continue during at least two years at Orange Labs and in the future we will continue to explore in depth the "Target", "Aggregate" which are essential to obtain a functional observatory. It is also essential to implement an approach that will be compliant with the anonymization constraints we need to respect in order to ensure privacy and business confidentiality. Anonymization is one of the main focuses of our work for the next two years and we already participated in the DARC hackathon at the workshop APVP 2021, where we had to protect and attack trajectory datasets, in order to improve our knowledge in this domain.

Acknowledgements

This work is carried out at Orange Labs, in collaboration with the Internet Physics Chair (Mines ParisTech - PSL University). This work is supported by the Île de France region as part of the "Territorial Support" call for interest (launched in 2020) which aimed at promoting collaboration between local authorities and logistics professionals, to develop a virtuous logistics system that will enhance the attractiveness of the Ile-de-France region and reduce environmental pollution. The project was selected in the "new methods for collecting and processing logistics data for companies and local authorities" category. This work is (partially) supported by the EIPHI Graduate School (contract ANR-17-EURE-0002).

References

1. Mohammad Alaggan, Sébastien Gambs, Stan Matwin, and Mohammed Tuhin. Sanitization of Call Detail Records via Differentially-Private Bloom Filters. In Pierangela Samarati, editor, *29th IFIP Annual Conference on Data and Applications Security and Privacy (DBSEC)*, volume LNCS-9149 of *Data and Applications Security and Privacy XXIX*, pages 223–230, Fairfax, VA, United States, July 2015. Springer International Publishing. Part 5: Privacy and Trust.
2. Mohammad Alaggan, Sébastien Gambs, and Anne-Marie Kermarrec. Blip: Non-interactive differentially-private similarity computation on bloom filters. 10 2012.
3. Eric Ballot, Benoit Montreuil, and Russell Meller. *The Physical Internet*. 09 2014.
4. Panthadeep Bhattacharjee and Pinaki Mitra. Bisdbx: towards batch-incremental clustering for dynamic datasets using SNN-DBSCAN. *Pattern Anal. Appl.*, 23(2):975–1009, 2020.
5. Vincent Bindschaedler and R. Shokri. Synthesizing plausible privacy-preserving location traces. *2016 IEEE Symposium on Security and Privacy (SP)*, pages 546–563, 2016.
6. Vincent D. Blondel, Adeline Decuyper, and Gautier Krings. A survey of results on mobile phone datasets analysis. *EPJ Data Sci.*, 4(1):10, 2015.

7. Noelia Caceres, Johan Wideberg, and Francisco Benitez. Review of traffic data estimations extracted from cellular networks. *Intelligent Transport Systems, IET*, 2:179 – 192, 10 2008.
8. Francesco Calabrese, Giusy Di Lorenzo, Liang Liu, and Carlo Ratti. Estimating origin-destination flows using mobile phone location data. *IEEE Pervasive Comput.*, 10(4):36–44, 2011.
9. Yang Cao and Masatoshi Yoshikawa. Differentially private real-time data release over infinite trajectory streams. 2:68–73, 2015.
10. Chi-Hua Chen. A cell probe-based method for vehicle speed estimation. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E103.A:265–267, 01 2020.
11. DataLab. *Chiffres clés du transport*, 2020.
12. Gerard de Jong, L.A. Tavasszy, John Bates, Stein Grønland, Stefan Huber, Oskar Kleven, Peter Lange, Ole Ottemöller, and Nora Schmorak. The issues in modelling freight transport at the national level. *Case Studies on Transport Policy*, 4, 08 2015.
13. Yves-Alexandre de Montjoye, Sébastien Gambs, Vincent Blondel, Geoffrey Canright, Nicolas Cordes, Sébastien Deletaille, Kenth Engø-Monsen, Manuel García-Herranz, Jake Kendall, Cameron Kerry, Gautier Krings, Emmanuel Letouzé, Miguel Luengo-Oroz, Nuria Oliver, Luc Rocher, Alex Rutherford, Zbigniew Smoreda, Jessica Steele, Erik Wetter, and Linus Bengtsson. On the privacy-conscientious use of mobile phone data. *Scientific Data*, 5:180286, 12 2018.
14. Yves-Alexandre de Montjoye, Cesar Hidalgo, Michel Verleysen, and Vincent Blondel. Unique in the crowd: The privacy bounds of human mobility. *Scientific reports*, 3:1376, 03 2013.
15. Thierry Derrmann, Raphael Frank, Francesco Viti, and Thomas Engel. Estimating urban road traffic states using mobile network signaling data. pages 1–7, 2017.
16. Council EU. *Proposal for a regulation of the European parliament and of the council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)*, 2021.
17. Marco Fiore, Panagiota Katsikouli, Elli Zavou, Mathieu Cunche, Françoise Fessant, Dominique le hello, Ulrich Aïvodji, Baptiste Olivier, Tony Quertier, and Razvan Stanica. Privacy in trajectory micro-data publishing: A survey. 08 2020.
18. Xi He, Graham Cormode, Ashwin Machanavajjhala, Cecilia M. Procopiuc, and Divesh Srivastava. Dpt: Differentially private trajectory synthesis using hierarchical reference systems. *Proc. VLDB Endow.*, 8(11):1154–1165, July 2015.
19. Andreas Janecek, Danilo Valerio, Karin Anna Hummel, Fabio Ricciato, and Helmut Hlavacs. The cellular network as a sensor: From mobile phone data to real-time road traffic monitoring. *IEEE Transactions on Intelligent Transportation Systems*, 16(5):2551–2572, 2015.
20. Qiang Ji, Beihong Jin, Yanling Cui, and Fusang Zhang. Using mobile signaling data to classify vehicles on highways in real time. pages 174–179, 2017.
21. Georgios Kellaris, Stavros Papadopoulos, Xiaokui Xiao, and Dimitris Papadias. Differentially private event sequences over infinite streams. *Proc. VLDB Endow.*, 7(12):1155–1166, August 2014.
22. Guillaume Leduc. Road traffic data: Collection methods and applications. 01 2008.
23. Darakhshan Mir, Sibren Isaacman, Ramon Caceres, Margaret Martonosi, and Rebecca Wright. Dp-where: Differentially private modeling of human mobility. *Proceedings - 2013 IEEE International Conference on Big Data, Big Data 2013*, pages 580–588, 10 2013.

24. Latanya Sweeney. k-anonymity: A model for protecting privacy. *Int. J. Uncertain. Fuzziness Knowl. Based Syst.*, 10(5):557–570, 2002.
25. Florence Toilier, Mathieu Gardrat, Jean-Louis Routhier, and Alain Bonnafous. Freight transport modelling in urban areas: The french case of the freturb model. *Case Studies on Transport Policy*, 6, 09 2018.
26. Huayong Wang, Francesco Calabrese, Giusy Di Lorenzo, and Carlo Ratti. Transportation mode inference from anonymized and aggregated mobile phone call detail records. In *13th International IEEE Conference on Intelligent Transportation Systems, Funchal, Madeira, Portugal, 19-22 September 2010*, pages 318–323. IEEE, 2010.
27. Shuo Wang, Richard Sinnott, and Surya Nepal. Privacy-protected statistics publication over social media user trajectory streams. *Future Generation Computer Systems*, 87, 08 2017.
28. Zhenzhen Wang, Sylvia He, and Yee Leung. Applying mobile phone data to travel behaviour research: A literature review. *Travel Behaviour and Society*, 11, 03 2017.
29. Yonghui Xiao and Li Xiong. Protecting locations with differential privacy under temporal correlations. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, CCS '15, page 1298–1309, New York, NY, USA, 2015. Association for Computing Machinery.