Analysis and Traffic Management in Smart Cities

D Namiot¹, O Pokusaev² and A Chekmarev²

¹Lomonosov Moscow State University, 119991, Leninskie Gory Moscow, Russia
²RUT (MIIT), 127994, 9b9 Obrazcova Street, Moscow, Russia

E-mail: dnamiot@gmail.com

Abstract. This article examines the issues of traffic management in the city. It is well known that cars are a significant contributor to urban air pollution. Accordingly, the management of transport in the city is one of the imperative tasks in terms of supporting the environment and ensuring urban development. Such management is impossible without collecting information on traffic flows in the city. It is logical to assume that there should be a single data source (information store) that contains information about all shipments. Urban governance, of course, involves not only collecting and analyzing data, but also managing this process. This article is devoted to the discussion of such systems.

1. Introduction

The main challenge for atmospheric air pollution in all megalopolises of the world, and Moscow is no exception, is road transport. This is a challenge of the present moment. 90% of all harmful emissions into the atmosphere are carried out by cars [1]. In general, transport accounts for 26% of CO₂ emissions worldwide. And the main (81%) is the contribution of road transport. The academic literature notes that car emissions are a source of more air pollution than any other human activity [2].

Accordingly, the management of transport in the city is one of the imperative tasks in terms of supporting the environment and ensuring urban development. As an example of transport management, which directly affects the environment, one can point out, for example, numerous works on how the joint movement of cars in a convoy (platooning) affects the environment. Such convoys came from military practice, where they consisted of two or more trucks that moved a short distance from each other, communicating wirelessly (communication between vehicles - V2V) [3]. The driver of the first car (leading) controls the direction and speed of the cars following him. The first attempts to build such systems were based on adaptive cruise control (ACC) systems. Now more complex systems are used, based on the constant exchange of information both between the cars themselves and with the back office. In the literature, it is noted that in such convoys for trucks, only due to the lower air resistance, fuel savings can reach 15%. Consequently, emissions are also reduced [4].

What, in fact, is the basis of the listed development? Naturally, this is information about the movements themselves. This is what is, first of all, necessary for building a convoy. Transport management is, of course, the management of its individual units. It is for them that it is necessary to build routes, identify "hot" areas in the city with traffic congestion, look for trends in the use of various types of transport, etc.

This, in turn, requires travel information about the city. In a modern city, there are many traditional means of transportation (public transport, private cars, taxis, city railways, etc.), new ones appear (car-
sharing, electric scooters, etc.). In the future, drones and air taxis will also belong to urban transport. This means that the city must collect information about such trips. About all trips, and not only in a historical context, but also in real time. The information should be depersonalized in terms of users (drivers, passengers), but it should include the type of vehicle, the start (end) time of the trip, and the start (end) point of the route.

Of course, all the information about routes is collected by transport companies. When ordering a taxi, the route is fixed. The car-sharing company knows where the user got into the car and where he left it. For renting a bicycle (scooter), it is known where the user took it and where he left it. All this information is needed by such companies directly to carry out their main activities. First of all, based on such data, users are billed. Secondly, this, of course, is needed to manage the fleet of own funds - where at any given time are the fixed assets (property) of the company.

It turns out that in order to optimize the use of transport at the city level, the same information is needed, but at the city level. This information should come from different companies, each of which can (and does) use their own IT systems. Accordingly, the first task is to develop standards for the provision of such data.

The rest of the paper is organized as follows. In Section 2, we discuss the Open Mobility Foundation project. In Section 3, we describe digital twins for cities. And Section 4 is devoted to the conclusion.

2. Open mobility foundation

In this regard, we would like to focus on the open standard for information on mobility - MDS, proposed by the Open Mobility Foundation [5]. Launched in 2019, the Open Mobility Foundation describes itself as a "public-private forum" designed to help local governments gain control of their roads from private mobile companies using big data and open source. A central part of the OMF’s mission is to manage the new mobility data standard. This MDS standard was proposed by the Los Angeles Department of Transportation. MDS is now receiving extensive real-time information on the status of dockless scooters and shared bikes. Many other cities in the United States (Miami, Seattle, Portland, San Francisco, Austin, Minneapolis, and others), as well as some cities in Europe, have adopted this standard.

The information that will be received in real time will allow the city to implement the so-called digital twin for urban mobility. This can be seen as an extension of the well-known simulation systems. Simulation systems (discrete-event, continuous) have been widely used since the 60s of the last century. They already have a fairly rich history of application, including transport systems modeling. Naturally, any modeling starts with the available data, and any transport modeling system must make some assumptions about its incoming flows. Strictly speaking, modeling consists in evaluating the performance of the system with different initial data.

A digital twin, according to the classical definition, is a digital copy of a living or artificial physical object [6]. The term digital twin refers to a digital copy of potential and real physical assets (physical twin), processes, people, places, systems, and devices that can be used for a variety of purposes. Digital twins are designed to facilitate the control, understanding, and optimization of the functions of all physical assets, ensuring the smooth transfer of data between the physical and virtual world [7].

Accordingly, the digital twin, by definition, includes a simulation model. Hence, it is natural to assume that the tools for developing simulation models can be used to create digital twins. In fact, the digital twin can be thought of as transferring (supplying) real data to a simulation model [8]. But there are also differences.

All digital twins potentially deal with heterogeneous and multidimensional data coming from different sources. Simulation systems most often adopt simplified data models. For example, for a variety of discrete-event modeling models, it is typical to model the occurrence of events in accordance with some statistical distribution. For digital twins, simplifications are not possible (otherwise it will not be a digital twin). This model can receive directly from physical objects, while the data can be both static and dynamic, the data can be generated by the model itself, or come from
various services that process both measurement data and external (in relation to the physical object) data. Note also that data can also result from data fusion. Naturally, in this case, this data does not come by itself, and the digital twin must be characterized by support for multiple connections. This is necessary to receive data from a physical object, as well as from external services. These connections are shown in Figure 1.

![Figure 1. Connections in a digital twin [9].](image-url)

The main difference between the digital twin and the simulation model can be formulated as follows. Modeling is the analysis of “what will be if ...”, while the digital twin is “what is happening now and what will be if ...”. The latter use is the key to urban governance.

According to the definition (vision) of the Open Mobility Foundation "In the future, each city should manage its own digital twin, which will provide the basic data on which mobility services depend." And further, as the design guiding principles, it is indicated that the work of OMF will be “based on the digital twin” model, which determines that the municipalities own and control the final digital data model for urban mobility.” Having a virtual copy of real mobility flows - for scooters and motorcycles, cars, self-driving cars, and drones - will allow local authorities to both track the movements of individual vehicles and, to some extent, control their movement.

Naturally, data privacy issues immediately arose along the way. As mentioned above, all transport companies are required to collect and analyze this kind of data. But they do it for their own purposes. These goals may not be the same as those of the city. Private companies may have their own view of what is (and most importantly - how) traffic management is carried out.

The Los Angeles project was based on an understanding of what the city could become with self-driving cars. For example, they can easily block streets or even city districts. It was from this that the idea of constant information about our location arose: “our job is to move people and goods as quickly and safely as possible, but we can only do this if we have a complete picture of what and where is on our streets.”

It should be noted that Moscow is no exception in this regard. Car-sharing cars can be parked in courtyards, completely excluding parking for local residents. Numerous private taxi companies may choose to park their cars on quiet streets in dormitory areas without any other use. Excessive motor rallies of residents' cars in search of parking is a direct deterioration of the environment.

Private companies in the United States (Uber and Lyft) prevented cities from accessing vehicle locations, numbers, time stamps, and routes. Local governments are keen to use this data to improve traffic flows and, for example, regulate car rentals. These are the two issues that most affect traffic congestion and public transport use. In contrast, search companies prefer to set their own terms. One
of the main objections is that it is possible to reconstruct the movements of specific individuals from anonymized data [10]. This is due to the fact that anonymous data is compared with other available sources of information.

But it is also important to note something else. The standard also includes traffic control. That is, in fact, instructing (directing) private carriers. A mixed reaction can be expected here. Informing carriers about street closures is something they will welcome, and street closures and parking restrictions are unlikely to generate enthusiasm.

3. On digital twins

The earliest proof of concept for digital twin technology came in 2014. Researchers at the University of Washington announced that they have created "a self-organizing and scalable multi-camera tracking system that tracks people through cameras," combining Google Street View with a suite of smart surveillance cameras trained on city streets. They demonstrated the ability to build a near real-time, fluid visualization of mobility flows displayed in a panoramic photograph of the Google world.

Machine learning software enabled seamless technology splicing [11].

The urban digital twin in the German city of Herrenberg is different from other modeling-based research in smart cities. The twin enables the linking and combining of various urban data, including social data collected from citizens [12].

![Figure 2. Replica from Side Walks [13].](image)

The Replica project from Sidewalk Labs (Google) uses data collected from smartphones to represent real city flows (Fig. 2). The Android platform, controlled by Google, is the most popular (widespread) mobile platform, giving the company real-time measurement tools [13].

Overall, the current tracking capabilities in cities inevitably raise privacy questions. But this is, to some extent, an inevitability that is hard to deny. In Moscow, the city authorities are collecting (with the help of telecommunications operators) a fairly large set of data on the movements of mobile subscribers in the region. This information is now used for movement analysis and planning of transport systems, which was used in many of our works [14, 15]. But technically, this kind of data can be used in real time.

However, to date, none of the approaches used have reached the level of direct oversight (governance) provided for in the OMF charter. This oversight takes the current capabilities of MDS and its various communications programs to a much higher level. The traffic planner of the centralized control system will be alerted to changing conditions in real time as real-time data streams continually redraw their digital territory. Built-in artificial intelligence will inform the engineer about the best decision to make. OMF offers several examples of how this ultra-modern Smart City will run traffic
simulation [11]. Currently, at least in the US, private carriers provide limited data to city services. For example, Uber Movement (Fig. 3).

![Uber Movement](image_url)

**Figure 3.** Uber movement [17].

Uber's R&D department even has an interesting trajectory visualization system - Kepler.gl. The problem is that city platforms want more data, consistently from different carriers, and most importantly, predictably. The sources mentioned above, for example, report the following data issues from Uber:

- It does not cover all pairs of route source and destination points for each time slot. There are gaps in it, since sometimes there are simply not enough trips on a given route.
- It does not provide aggregated data for a specific date and time range in a downloadable format. Average values are provided for specific intervals only.
- The dataset is aggregated by districts. It does not have the location (longitude / latitude) of the start and end points of the trip.
- And of course it only covers some selected cities

Accordingly, such data can be used for analysis, but are not suitable for operational management. And, of course, cities must be able to exchange data. The presentation format cannot be private.

Of course, any city is looking for ways to increase profitability in such data. In particular, dynamic pricing for the same parking lots, depending on the workload and demand for specific sites, dynamic determination of fees for entering certain areas at the calculated time, etc. One of the widely discussed ideas, for example, dynamic pricing of parking at curbs [16]. And, naturally, by driving transport, cities expect to seriously reduce emissions.

As the virtual living equivalent of many city systems, the digital twin allows the management organization (Department of Transportation) to model possible strategies to plan and fix problems before and as they arise, and to implement a solution that has been virtually tested in many simulated scenarios to minimize risk. Considered, for example, the completion of a concert at the stadium (for Moscow it could be Luzhniki or VTB Arena). These are tens of thousands of visitors. What would be the best temporary one-way street configuration in the event area, and how long should the temporary configuration remain in effect? City planners and operators can use dashboards, which provide access to different planes of the physical and virtual worlds to get the information they need to make effective decisions.
4. Conclusion
It seems to us that the Open Mobility Foundation's approach can at least serve as a base and a prototype for similar solutions in Moscow and other Russian cities. The problems facing the city authorities in terms of transport are exactly the same. Private carriers in cities also do not inform urban systems about their transport. Regarding the Replica project, it can be pointed out that the city authorities in Moscow collect a fairly large amount of data on the geography of mobile users' movements in the region. This is a good basis for creating a digital twin.

5. References
[1] Kupriyanovsky V et al 2016 On Internet of Digital Railway International Journal of Open Information Technologies 4(12) pp 53-68
[2] McBain Bonnie et al 2018 Reducing the ecological footprint of urban cars International Journal of Sustainable Transportation 12(2) pp 117-127
[3] Kupriyanovsky V et al 2017 On intelligent mobility in the digital economy International Journal of Open Information Technologies 5(2) pp 46-63
[4] Pajak M & Cyplik P 2020 Truck platooning in the context of sustainable development’s targets defined in European Union’s strategies LogForum 16(2) pp 311-321
[5] Open Mobility Foundation https://www.openmobilityfoundation.org/
[6] El Saddik Abdulmotaleb 2018 Digital twins: The convergence of multimedia technologies IEEE multimedia 25(2) pp 87-92
[7] Khajavi, Siavash H et al 2019 Digital twin: vision, benefits, boundaries, and creation for buildings IEEE Access 7 pp 147406-147419
[8] Namiot Dmitry et al 2021 Digital twins and discrete-event simulation systems International Journal of Open Information Technologies 9(2) pp 70-75
[9] Tao, Fei et al 2018 Digital twin driven prognostics and health management for complex equipment Circuits 67(1) pp 169-172
[10] Li S, Tian H, Shen H & Sang Y 2021 Privacy-Preserving Trajectory Data Publishing by Dynamic Anonymization with Bounded Distortion ISPRS Int. J. Geo-Inf 10 pp 78
[11] Bloomberg https://www.bloomberg.com/news/articles/2019-07-19/why-cities-want-digital-twins-to-manage-traffic
[12] Dembski, Fabian et al 2020 Urban digital twins for smart cities and citizens: the case study of Herrenberg, Germany Sustainability 12(6) pp 1-17
[13] Goodman, Ellen P and Julia Powles 2019 Urbanism under google: Lessons from sidewalk Toronto Fordham L. Rev. 88 pp 457-498
[14] Bulygin M & Namiot D 2021 Anomaly Detection Method For Aggregated Cellular Operator Data In 2021 28th Conference of Open Innovations Association (FRUCT) pp 42-48
[15] Misharin A et al 2018 On Passenger Flow Estimation for new Urban Railways In IOP Conference Series: Earth and Environmental Science Vol 177 1 p 012012
[16] Wan Y, Zhou J, He W & Ma C 2020 Modeling the Curb Parking Price in Urban Center District of China Using TSM-RAM Approach Journal of Advanced Transportation Article ID 4905059
[17] Uber movement https://towardsdatascience.com/putting-uber-movement-data-into-action-machine-learning-approach-71a4b9ed0acd