Analytics Use Cases for Landside Traffic Optimization in the Catchment Area of the Airport: Case Study of Zagreb Airport

Marin Tica*, Igor Štimac**, Krešimir Vidović***, Saša Vojvodić*** and Irina Stipanović****
*International Zagreb Airport Jsc., Zagreb, Croatia
**Zagreb Airport Ltd., Zagreb, Croatia
***Ericsson Nikola Tesla, Zagreb, Croatia
****Infraplan, Zagreb, Croatia
kresimir.vidovic@ericsson.com

***Abstract* - Airports and their facilities are general attractors that significantly influence the supply and demand of landside traffic in the adjacent and larger catchment areas of the airport. This paper will draft a concept of a decision support solution for strategic planning that will enhance, optimize and significantly improve the management of land transportation-related activities for catchment areas of the airport. Following the principles of FRAME Next ITS architecture creation process, data-driven analytical use cases will be identified and proposed that utilize commonly available data sources in the airport environment (e.g. anonymized telecom network data, public transport data, flight information data, airport parking data, transport on-demand data, etc.) and identify and prioritize real analytical use cases of importance to key stakeholders (e.g. airlines, local community, public transport operators, etc.). Proposed analytical use cases are demonstrated for Franjo Tuđman Airport in Zagreb, Croatia.**

***Keywords*: airport landside traffic, optimization, data analytics**

I. INTRODUCTION

Airports are in general known for having a significant impact on both the environment and public health and are, therefore, subject of improvement so that they can meet nominated Sustainable Development Goals of the United Nations regarding environmental efficiency. Specific actions are required with the purpose of increasing performance from these perspectives, and one of the most significant ones is related to the improvement of ground transport in catchment areas of the airport. The goal is to foster the use of sustainable and green multimodal solutions for all journeys related to the airport (employees, passengers, supporting services, etc.) in order to decrease the direct environmental footprint of the airport in the area of ground transportation. Therefore, this research will propose a decision-making system concept that will utilize available common (traffic counter, airline data, etc.) and advanced data sources (anonymized mobile network data) that will be used as input for defined analytical use cases. Analytical use cases will be identified in order to target the real problems and needs from relevant stakeholders influenced by ground transport.

According to literature review, the existing researches have identified that both conventional and emerging data sources can be used as a base for quality data-driven decision-making in the area of urban mobility management in small/medium-sized cities. Also, a software-based decision-making system for urban mobility management can be used as a tool in Intelligent Transport System (ITS) applications, where it, besides data requirements, usually consists of a data aggregation platform and relevant applications as a key building blocks. Functional requirements on those subsystems are identified especially in priority areas which are related to passenger information, weather and environmental monitoring, traffic management and operations, people's safety in the transport system, public urban transport and electronic transport-related payments. [1]–[3].

Besides the introduction in first chapter, the paper will address the process of identification and core use cases in Chapter 2. Chapter 3 will provide overview of data sources used in this research, Chapter 4 will present the decision support system architecture, Chapter 5 will present the proposed methodology, and Chapter 6 will include conclusion remarks.

II. USE CASES IDENTIFICATION

Since the goal of the research is to develop a decision support solution for strategic planning that will enhance, optimize and significantly improve the management of activities related to land transportation for the catchment area of the airport, relevant stakeholders were involved in order to identify targeted areas of improvements. Since optimization measures will be proposed in the area of Intelligent Transport Systems (ITS), the methodology for use case development is based on the combination of Delphi method and utilization of FrameNext ITS architecture (FRAME). An ITS architecture is a conceptual design that defines the structure and/or behavior of an integrated Intelligent Transport System (ITS). The Delphi method is a process used to arrive at a group opinion or decision by surveying a panel of experts [4]. The FRAME Architecture comprises top level requirements and functionality, or use cases, for almost all ITS applications and services that have been considered for implementation somewhere in the European Union. It is at a "level" such that it can be used as a reference by all ITS architects and is intended to be the foundation for building other types of architecture that will be necessary. It will enable them to guarantee compliance at the interfaces of other systems so that seamless services can be provided to cross-border travelers, and an open European market of compatible components can be established. [5], [6] Therefore, the Delphi methodology has been used to extract knowledge from relevant experts nominated from key stakeholders, and FRAME architecture was used to shape the user requirements in the matter to fit the standardized requirements. During a series of workshops, a number of stakeholders (12) were interviewed so as to identify key issues of landside transportation in the catchment areas of the airport. Relevant stakeholders involved representatives of neighboring cities, representatives of the airport and connected companies, representatives of public transport companies operating in the area, and they were nominated in order to fullfill the FRAME next principles, according to which, all relevant stakeholders' groups have to be represented (stakeholders that want ITS, use ITS, rule ITS, make ITS and service providers). The entire process consisted of five workshops, which resulted in the identification of analytical use cases proposed for land transport optimization in narrower airport area. Following the identification of analytical use cases that will be addressed in this research (user aspirations), research team has proposed a research methodology, required data sources and infrastructure, and has consolidated them in a consistent manner that is suitable for the next stage in the process (user needs). The consolidated requirements where then presented to stakeholders, and after a common agreement has been achieved, use cases have been prioritized and completed. These use cases are set up as key requirements for the decision-making system. The use cases are the following:

The first use case is public transport optimization for airport users (passengers) with the goal of optimization of existing public transport lines or introduction of new public transport lines in a narrower catchment area of the airport (Zagreb, Velika Gorica) serving airport passengers.

The second use case is public transport optimization for airport users (employees) with the goal of optimization of existing public transport lines or introduction of new public transport lines in a narrower airport catchment area (City of Zagreb, City of Velika Gorica) serving airport employees.

The third use case is related to airport strategic planning and includes an analysis of catchment areas of the airport based on the analysis of users migrations where either trip origins and/or destinations are at the airport, with the goal of identification of real catchment areas, enhancement of sustainable transport connections for these areas, analysis of potential areas which are not, but are expected to be, served by the analyzed airport.

The fourth use case is related to airline strategic planning (analysis of airline users). It consists of an analysis of airline users based on the information of the initial country of origin/destination in order to identify a passenger's country of origin based on telecom data and to identify countries that have a significant travel demand, but have not initially been served by airline companies.

The fifth use case is based on the analysis of migration and retention habits of passengers that include analysis of migration and retention habits of passengers, identification of passenger type based on its migration patterns (tourists, business users) and optimization of service/offering for targeted users.

The sixth use case is based on the analysis of transport demand prediction based on the information from airport arrivals/departures and expected utilization of public transport services serving the airport area, including providers of on-demand transport. Based on historical data, migration patterns of airport passengers following the arrival of an aircraft (e.g. the number of passengers leaving airport with a provider of on-demand transport or public transport), the demand will be predicted for a specific transport mode/service (e.g. on-demand transport, rent-a-car, etc.)

The next step in the research is to propose a system architecture and data sources that will enable realization of identified use cases.

III. IDENTIFIED DATA AND CORRESPODING DATA SOURCES USED FOR DECISION-MAKING

The system relies on a number of data sources used as input data for decision-making. These data sources include transport data, public transport data, anonymized mobile network data, airline data and other data. Since anonymized mobile operator data have a great potential and are still not widely used, most of this chapter will be devoted to them.

A. Transport data

Traffic data includes data on traffic flow parameters as well as data on traffic situations. The Datex standard is used to exchange traffic data between traffic control, information centers, traffic operators and other stakeholders. Datex allows the collection of data on traffic accidents, current roadwork and other traffic-related events. The data is presented in XML format and modeled with a UML diagram. The structure of the message includes a publication from the table of measurement locations and the publication of measured data. The data source includes the following data: vehicle type and vehicle flow rate. Individual Vehicle Data includes the data on the arrival and exit road lane, vehicle speed, direction. Moreover, situation data is available as well. Situation data includes information on incidents, like improper turn, a vehicle went through red light, a vehicle moving in the wrong way, a vehicle improperly stopped or parked, and other events, such as accidents. In this research, vehicle flow data will be mostly utilized.

B. *Public transport data*

Public transport data is exchanged via NeTEx/GTFS or Siri (excluded from this research due to the unavailability of data in this format). NeTEx provides a means to exchange data for passenger information, such as stops, routes, timetables and fares, among different computer systems, together with related operational data. It can be used to collect and integrate data from many different stakeholders, and to reintegrate it as it evolves through successive versions. NeTEx provides a breakdown of transport modes by types: scheduled (air, rail including high-speed rail, conventional rail, light rail, long-distance coach, maritime including ferry, metro, tram, bus, trolleybus), demand-responsive (shuttle bus, shuttle ferry, taxi, car-sharing, car-pooling, car-hire, bike-sharing, bike-hire) and personal (car, motorcycle, bicycle.) In this research, data is obtained using NeTEx protocol from National Access Point for Multimodal data and the following is used: information on public transport lines, stops and timetables.

C. *Anonymized mobile network data*

Commonly used traffic sensors within the road network are very precise in performing traditional measurements (traffic flow characteristic on a road section). However, since the requirements from the road authorities and operators might surpass that level of data or be oriented on those road sections where the sensor infrastructure does not exist, alternative methods have been studied for years [7]–[11]. One of the most researched data sources, with the highest potential, is mobile network operator data, with an increased rate of popularity as a result of increased penetration of mobile phone users (market penetration in developed countries is more than 100%), and a result of introduction of new generations of mobile networks (3G, 4G, 5G). Every new generation brings significant improvements in telecommunication service and, as a result, a more useful accompanying databases generated because of network operations. Furthermore, every new generation brings improvement in user positioning as well, and since it is still not as precise as Global Navigation Satellite System (GNSS) systems, this data can still be very useful in transport engineering.

The most commonly used data sources from mobile network data include Call/Charging Data Records (CDR) data, mobile network signalization data logs (CTUM, CTR, etc.), mobile network subscriber database data, and performance counters data.

Call/Charging Data Records (CDR) data represents a dataset that is collected because of user telecommunication activities, since every time a user performs a telecommunication activity, the activity is logged in this database for future charging purposes. Mobile network signalization data consists in majority of cell trace data, where communication between mobile network terminal and corresponding antenna is logged. An advantage of these data sources is that the logs are populated even if mobile phone is not being actively used (in call, data usage, etc.). These two data sources are user-based, and required processes, organization measures and anonymization techniques have to be applied which ensure that data is handled in accordance with the General Data Protection Regulation (EU) 2016/679 (GDPR), which regulates data protection and privacy in the European Union (EU) and the European Economic Area (EEA). The use of this data enables anonymized reconstruction of user migrations, and therefore user trips with additional parameters (origins, destinations, duration, length, speed) can be extracted. If fused with anonymized user database data (CRM – Customer Relationship Management), additional data can be applied to extracted trip database, and this data can include information on user sex and age range. That enables various kinds of analytics for identification of habits: identification of travel patterns for age group and sex, identification of average speed, identification of the type of users which obey speed limits, etc. As opposed to previously mentioned data sources, performance counters data are not user-oriented; rather, they are location-oriented. LTE/E-UTRA performance management (PM) counters are aggregated over 15-minute intervals. These counters are inherently privacy-preserving, since they can represent the number of terminals being served by a corresponding cell and can be used as virtual traffic counters.

Telecom data sources can be utilized for various transport-related analyses. Signaling data, CDR data and subscriber data can be used for both strategic and operational purposes for the identification of traffic flow characteristics (traffic volume, speed), for driving distraction use cases (analysis of mobile phone usage by drivers), for the identification of transport demand (origin and destination data). Performance counters data can be utilized for highly efficient operational virtual traffic counters, for supplementing locations where traditional traffic counters do not exist.

In this research, this data source will be used to identify user types and migration patterns (more details in the Methodology chapter).

D. *Airport data*

Airport data contains information on aircraft arrivals and departures. The source of data is the airport management system. This data source provides information on real-time aircraft arrivals and departures, including information on air carrier data, scheduled time data of arrival and departure, estimated time data, stopover data, country data, airport data, flight code – airline code and flight number, gate data/exit data, aircraft stand data and status data (arrived/departed). In this research, all this data will be used in order to match landside transport demand and its correlation with aircraft arrivals and departures.

E. *Parking data*

Parking data contains information on the occupancy and utilization of the airport parking areas. The data source is the parking management system and due to the lack of standardization in this area, the data is not standardized and is available in a proprietary format. The data contains the following information: entry and exit time, user type (for taxi vehicles), entry and exit point, etc. This data source will be used to identify specific types of airport users (e.g. taxi) to validate transport mode data obtained from anonymized mobile network data.

F. Other

Other data sources might include general and well-known data, such as statistical data on the number of employees at the airport, census data on user migrations, data from transport models, etc. This data will be used to validate results of analyses from different data sources (like anonymized mobile network data).

IV. Decision Support System Architecture

The proposed decision support system architecture is presented on Figure 1. It consists of a four-level architecture. The first layer consists of available and required data sources. These data sources are integrated with the second layer—the data aggregation platform using a set of application programming interfaces. Where applicable, standardized protocols should be used to ensure compatibility of the solution. Therefore, transport and public transport data will be supported by using standardized protocols (Datex, Netex, GTFS, SIRI). Other data sources are not covered with standardization, so anonymized telecom data, parking data, airport data and other data sources will be integrated using proprietary data formats.

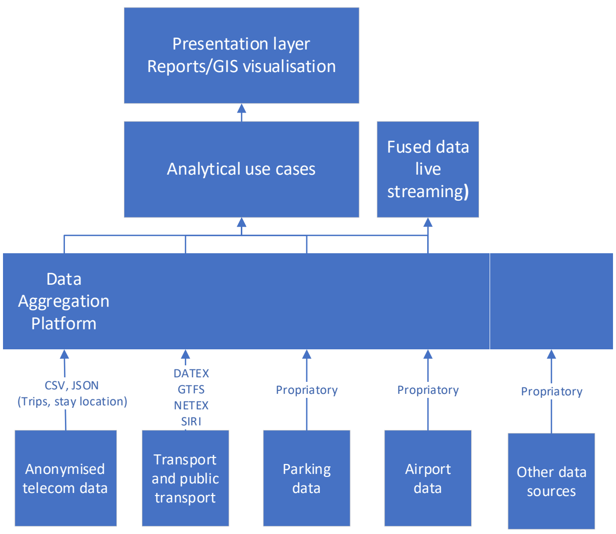

Figure 1. System architecture

The data is then stored in a data aggregation platform. It includes a three-layer architecture, where the southern interface is used for acquisition of data (input interface), the middle layer is used for operations with data, and the third layer—the northern interface—is used for publication of data toward users/applications (output interface).[12]

In the application layer, applications will be used to perform analytical use cases identified in this project. Besides an analytical engine, an interface will be ensured that enables live streaming of fused data entering into the platform. Above these layers, data presentation layers will be established that will enable generation of reports and visual presentation of data on an interactive web GIS tool.

V. Proposed Methodology for Identified Analytical Use Cases

Following the identification of available data sources and the decision support system architecture, a methodology has to be proposed in order to fulfill the use case requirements using available data sources. The methodology is presented on Fig. 2.

Anonymized telecom data is used for identification of all users stay locations, which are then transformed into user migrations (trips). The output is a table of stay locations per anonymized user with the information on time spent in a specific geographical area, together with a list of trips of all users that will include information on the trip start and end location, time of the beginning and end of a trip and information about distance between the starting and the last point of the trip. Based on this data, the first step includes identification of different user types. The following categories are identified: airport employee, airport user—passenger— arriving by aircraft airport user—passenger—departing by aircraft airport user (greeter/byer), airport user taxi driver, person passing near the airport—local residents—and commuters. All these user types can be identified based on their characteristic behavior following criteria such as: identification of Home/Work/Other sector, retention time at airport sector, trip characteristics, user type (domestic or foreign), time of the day, etc.

For example, a candidate for an airport employee might be a person whose estimated home sector is not the airport but spends significant time in their day at the airport sector. An airport employee should have a pair of trips originating and departing the airport. These trips should occur during airport working hours. On the other hand, a taxi driver should have a significant number of trips originating and departing at the airport and retention time at the airport should not be significant. A passenger arriving at the airport should have first occurrence at the airport sector, passenger leaving the airport should have last occurrence at the airport sector, etc. After identification, all these user types can be validated using various data sources, as presented on Fig. 2. Following this step, all trips in the trip database can be assigned to the appropriate user type, and therefore, for each use case only trips relevant to a specific user type can be considered.

As previously mentioned, based on trip characteristics (origin, destination, duration, speed and distance), transport mode can be identified. For the purpose of this research, a methodology has been applied that has differentiated two most common transport modes used to arrive to the airport—personal vehicles and public transport. The proposed methodology is based on a set of rules and application of Bayes statistical modelling, as presented in [13]. Following this step, all trips that include identified user types are then assigned with an appropriate transport mode. Transport mode data can be validated using supporting datasets from the airport parking, and data from traffic counters on road network in the airport area.

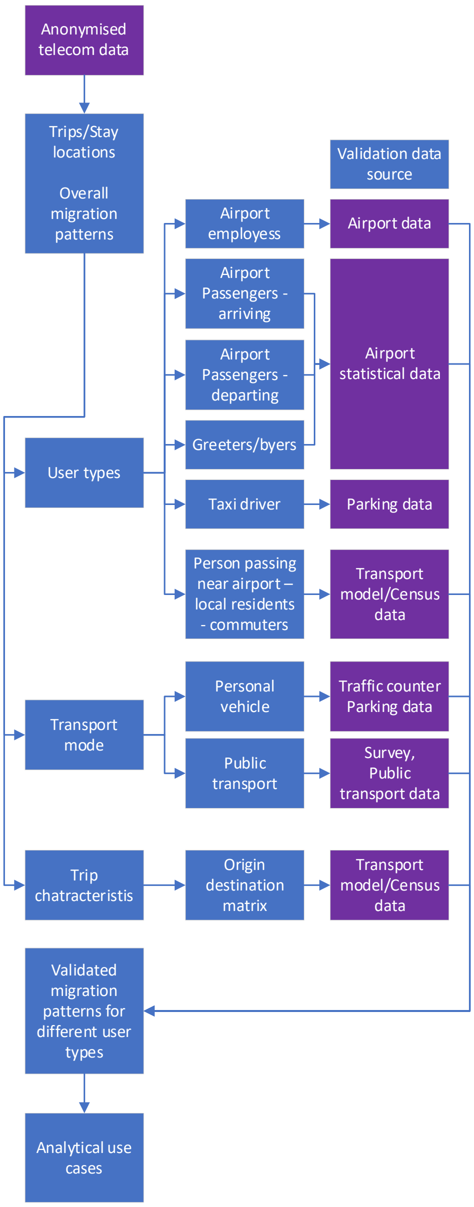

Figure 2. Methodology

At the end of this process, validated migration patterns have been generated and stored into an appropriate database, and can, therefore, serve previously identified use cases. Validation data sources have been proposed and considered as a ground truth data for each data source and each measured/identified parameter. At the end of analytical process, the analytical results will be validated using ground truth data in order to identify results inaccuracy.

VI. Conclusion

In order to meet health and environmental goals for the improvement of airport-related ground transportation activities, a decision support system concept has been introduced in this paper. Relevant stakeholders have identified most common ground transport-related issues, and analytical use cases have been proposed in order to quantify these values and become the base for optimization activities. Available relevant data sources have been utilized to identify target airport users (employees, passengers, supporting services operator, etc.) and to identify what their real travel needs and migration patterns are and what real airport catchment area is. These objectively measured values represent a real snapshot of airport ground transportation reach using origin destination matrixes for these users, together with the transport mode that they use when performing these travels. The proposed methodology can be used for other analytical use cases in various research domains besides transport (tourism, economy…) that are based on spatio temporal analysis of population migration. Next steps of the research will focus on the bare results of this proposed concept and will focus on optimization measures that will be introduced in order to fulfill the sustainability related goals.

Acknowledgment

This paper presents the results of activities performed within the OLGA project. OLGA project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement № 101036871.

References

[1] Medved, D., D. Blažinić, V. Galijan, & N. Antolović, Evolution of data sources for integrated data-driven urban mobility management, Transp. Res. Procedia, vol. 64, pp. 68–75, 2022.

[2] Draganić, M. et al., Requirements on applications within the decision-making system for urban mobility management, Transp. Res. Procedia, vol. 64, pp. 60–67, 2022.

[3] Radović, R., I. Marasović, V. Čačković, D. Pleština, D. Keresteny, & Z. Anić, The concept of a data aggregation platform in the function of a decision-making system for urban mobility management, Transp. Res. Procedia, vol. 64, pp. 53–59, 2022.

[4] Pérez, V. L. & R. Schüler, The Delphi Method as a tool for information requirements specification, Inf. Manag., vol. 5, no. 3, pp. 157–167, 1982.

[5] Bossom, R. & P. Jesty, Extend FRAMEwork architecture for

cooperative systems - D15 – FRAME Architecture – Part 1: Overview, 2011.

[6] Jesty, P. H. & R. A. P. Bossom, Using the FRAME Architecture for planning integrated Intelligent Transport Systems, 2011 IEEE Forum Integr. Sustain. Transp. Syst. FISTS 2011, pp. 370–375, 2011.

[7] Galloni, A., B. Horváth, & T. Horváth, Real-time Monitoring of Hungarian Highway Traffic from Cell Phone Network Data, in ITAT 2018 Proceedings, 2018, vol. 2203, pp. 108–115.

[8] Kos, G., P. Brlek, & K. Vidovic, SHARED SPACE CONCEPT IN LOCAL COMMUNITIES: CASE STUDY, in 8TH INTERNATIONAL CONFERENCE ON ROAD SAFETY IN LOCAL COMMUNITY, 2013, pp. 1–8.

[9] Goulding, J., Best Practices and Methodology for OD Matrix Creation from CDR Data, NLAB, Univ. Nottingham, vol. 44, no. 0, pp. 1–37, 2018.

[10] Furletti, B. et al., Use of mobile phone data to estimate mobility flows. Measuring urban population and inter-city mobility using big data in an integrated approach, Sis, pp. 1–10, 2014.

[11] Yaghoubi, F., A. Catovic, A. Gusmao, J. Pieczkowski, & P. Boros, Traffic Flow Estimation using Machine Learning and 4G/5G Radio Frequency Counters, vol. 2022-June, no. 1. Association for Computing Machinery, 2022.

[12] Radović, R., I. Marasović, V. Čačković, D. Pleština, D. Keresteny, & Z. Anić, The concept of a data aggregation platform in the function of a decision-making system for urban mobility management, in Transportation Research Procedia, 2022, vol. 64, pp. 53–59.

[13] Vidović, K., P. Čolić, S. Vojvodić, & A. Blavicki, Methodology for public transport mode detection using telecom big data sets: case study in Croatia, Transp. Res. Procedia, vol. 64, pp. 76–83, 2022.