PUBLIC TRANSPORTATION MULTIMODALITY IN THE CITY OF LISBON

Carlos Lemonde
INESC-ID and IST, Universidade de Lisboa, Lisbon, Portugal
LNEC, Department of Transport, Lisbon, Portugal
Elisabete Arsenio
LNEC, Department of Transport, Lisbon, Portugal
Rui Henriques
INESC-ID and IST, Universidade de Lisboa, Lisbon, Portugal

ABSTRACT

Mobility in major European capitals is not yet sustainable. The need to respond to the ongoing changes in public transportation demand, operationalize safety norms of social distancing, and reach carbon neutrality are prompting cities to reassess public transport systems. Cross-mode synergies in multimodal transport systems can be explored (including convenience, reliability, cost, speed and predictability) to foment public and active modes of mobility. In this context, multimodal traffic pattern analysis can unravel cross-mode vulnerabilities, a possibility that is finally rising with the sensorization of cities, integration of ticketing systems, and consolidation of traffic data sources and their situational context.

This work introduces a methodology for the analysis of spatiotemporal indices of multimodality against available situational context, aiding specialists to find vulnerabilities on the public transportation network. Traffic generation poles, large-scale public events, and weather records are the considered sources of situational context. We discuss the role of context-aware multimodality indices to understand demand and its emerging changes, assess cross-modal transfers and preferences, and support cross-mode route and schedule planning. This work further discusses the relevance of multimodal pattern discovery to offer data-centric views ensuring: fully transparent decisions to the citizens; and an objective coordination between carriers, municipalities and authorities. Lisbon is further introduced in this work as a reference case study for context-aware multimodal mobility.

1. INTRODUCTION

Worldwide, city municipalities are establishing efforts to collect urban data in order to gain more comprehensive views of the ongoing changes to city traffic dynamics. The undertaken initiatives are particularly relevant in face of pandemic-driven shifts to traffic demand and private car ownership, city reforms, the need to enforce safety norms, and the rising advocacy towards active modes of transportation. In particular, the Lisbon City Council has established relevant initiatives to this end: city traffic sensorization, consolidation of relevant sources of urban data on its Intelligent Management Platform , the integration of automated fare collection systems and tariffs along the public transport system, and entry requirements for carriers operating in the Lisbon metropolitan area. These initiatives offer unique opportunities for multimodal pattern analysis – encompassing road, railway and inlands waterways modes, as well as active modes such as walking and cycling – and cross-carrier coordination. Still, the inherent nature of multimodal traffic data – heterogeneous, massive in size, rich in spatiotemporal dynamics, subjected to variable aspects, and dependent on context factors – together with the increasing disruptive changes in urban traffic poses challenges to data-centric multimodal decisions.

This work proposes a methodology for a preliminary assessment of multimodal synergies in demanding urban areas of a city using multimodality indices estimated from available traffic data. Context-aware corrections are proposed in the presence of information associated with large-scale public events and available weather conditions. These contextual factors can account for meaningful variations to the level and distribution of traffic demand across transportation modes. A spatiotemporal analysis of multimodality indices is conducted for the city of Lisbon using three major modes of public transportation: bus (CARRIS), subway (METRO) and cycling (GIRA). In addition, we further extend this methodology to relate the gathered knowledge against additional sources of situational context, including traffic attraction-generation poles.

This paper aims at bridging the existing gap on the integrative analysis of multimodal traffic data and its situational urban context. A discussion on the relevance of cross-modal pattern analysis for the articulation between operators, and alignment of the public transport supply with the self-actualizing city dynamics are further provided.

This work is anchored in the pioneer research and innovation ILU project (DSAIPA/DS/0111/2018), a project that joins the Lisbon City Council and national research institutes to bridge ongoing research on urban mobility with recent advances on artificial intelligence.

The manuscript is organized as follows. Section 2 provides essential background. Section 3 introduces a novel methodology for the context-aware assessment of multimodality indices from heterogeneous traffic data. Section 4 gathers results of its application on the Lisbon city. Final remarks and major implications are then synthesized.

2. BACKGROUND

Multimodality is generally defined as the use of more than one transport mode to complete a trip within a certain period. In contrast, monomodality commonly refers to the exclusive use of one mode of transport (Nobis, 2007). Multimodality is a subfield of a broader body of research focused on intrapersonal variability of travel behaviour, comprising temporal, spatial, purpose and modal dimensions (Buehler and Hamre, 2016). Nobis (2007) emphasizes the fact that the general definition of multimodality must entail a time frame, possibly going beyond end-to-end trips to further encompass a week period. The longer the period, the higher the probability for multimodal transport use.

2.1. Opportunities and challenges in Lisbon

The Intelligent Management Platform of Lisbon City (PGIL) was publicly launched in 2017 to meet various policy and planning goals at the operational level. In this context, advancing towards sustainable and multimodal mobility in Lisbon is understood as a policy opportunity while it offers multiple research challenges around the interdisciplinary triaxial lens: data science and statistics – urban mobility planning – artificial intelligence. Using the available big data from various heterogeneous sources of traffic data across public transport operators, the following research challenges have been addressed so far:

- processing and consolidation of raw sources of urban data across public transport operators, and their multimodal descriptive and predictive analysis to support mobility planning decisions at the city level
- discovery of emerging spatiotemporal traffic patterns, while accounting for the stochastic nature of mobility dynamics (Neves et al., 2020, 2021).
- inference of dynamic and multimodal origin-destination (OD) matrices sensitive to missing boarding validations and monthly patterns of circulation (Cerqueira et al., 2020)
- incorporation of contextual data (weather, public events, traffic incidents, etc.) to enhance traffic data analysis and its integrative learning potential for future behaviour inferences (Leite et al., 2020; Cerqueira et al., 2020)
- exploratory policy framework of multimodality indices to measure social equity regarding users' multimodality options in the city (Lemonde et al., 2020)
- development of the ILU app, namely, to enable visualization of large-scale and heterogeneous data and to enhance specific data-centric views of the relevant information while accounting for each operator/city specificities and both cross-modal or unimodal user-centric perspectives.

The implementation of this set of urban analytics tools within the PGIL platform, managed by the city Council, is expected to support urban mobility planning, giving priority for public transport options and the integration of active travel modes (walking, shared public bicycles) with bus and/or metro/subway. Moreover, the full scalability and online nature of the devised tools can be enriched by targeting other dimensions of the city dynamics in the post-pandemic era.

3. METHODOLOGY

The assessment of multimodal urban transportation is a key issue in modern transport research, and the proposed solutions must handle the massive, heterogeneous, spatiotemporal, context-dependent and incomplete nature of traffic data. In this section, a sequence of procedures (Figure 1) is proposed for context-aware multimodal analytics.

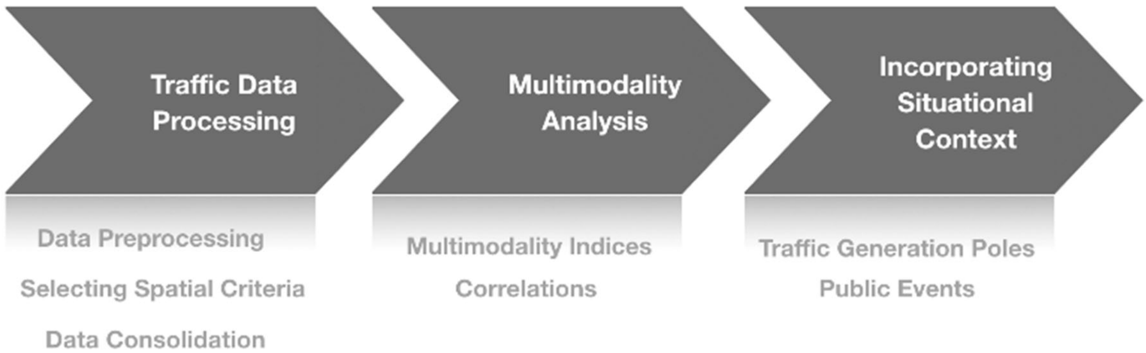

Figure 1: Public transportation multimodality analysis process

3.1. Traffic data processing

Raw traffic data from different sources can come in different formats and structures, being essential to start by preprocessing the collected data, where noise and inconsistencies are removed and high multiplicity of data sources are integrated within a single repository. The consolidation step was accomplished in this work using a multi-dimensional schema to allow for an effective and efficient cross-modal retrieval of trip records from specific users, operators, geographies and time periods. To this end, shared dimensions between sources were identified, including user, carrier, spatial and temporal dimensions.

Given the massive size of urban data, proper indexation of spatial, temporal and modal information was pursued for efficient data retrieval (Mamoulis et al., 2004). Efficient slicing, dicing and drilling query facilities were further made available. Unnecessary memory inefficiencies were further avoided by, for example, decoupling stations, vehicle or card details from the validation records. In addition, data cleaning procedures were applied to ensure the absence of duplicates and gross errors, and estimate missing entries using alight stop inference methods (Cerqueira et al., 2020). Finally, updating routines were applied for the automatic extraction, transformation and storage of continuously arriving trip records.

Given the presence of consolidated trip records, the user can select among pre-established spatial and temporal granularities for the subsequent multimodal traffic data analyzes. In terms of spatial specifications, two main possibilities are made available. One of them is to customly specify the target geographical region of interest (i.e. using a polygon or circular marking). The other is to select predefined regions.

The following zoning maps are considered for the Lisbon Metropolitan Area (Figure 2):

- traffic analysis zones (TAZ): geographical unit used in transportation planning models to assess socio-economic indicators
- administrative zones: coarsest geographical unit for the city, ranging from municipalities to parishes, depending on the geographical organization of the target city; and
- sections: finest geographical unit, comprising small districts and neighbourhoods.

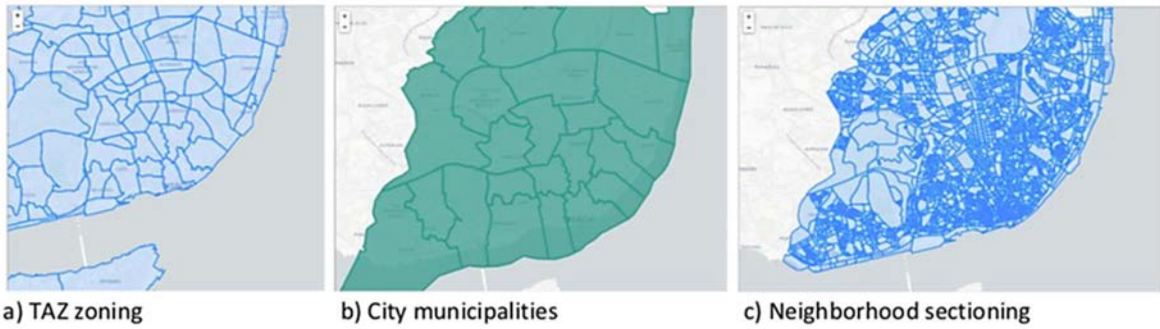

Figure 2: Zoning: geographical decomposition of the Lisbon city at different granularities

Two major types of temporal constraints can be placed. First, calendrical constraints – such as day of the week (e.g. Mondays), weekdays, holidays or on/off-academic period calendars – can be specified to segment the available traffic data. Second, time intervals (e.g. on/off-peak hour intervals) or a fixed time granularity (e.g. 15 minute) can be optionally placed to guide traffic data description and prediction.

Once these constraints are fixed, multi-dimensional querying can be automatically derived to produce the consolidated data. Given consolidated traffic data from cross-mode carriers, data mappings are generally further applied to transform the queried data into more conducive data structures (Mamoulis et al., 2004). Illustrating, spatiotemporal data structures can be mapped into georeferenced multivariate time series structures to facilitate subsequent mining tasks.

These time structures can be aggregated at different granularities and barycenter averaging further applied. Correlation between time series from different modes can also support the understanding of multimodal synergies. Linear correlation coefficients (e.g. Pearson's, Spearman's, Kendall's) and detrended cross-correlation analysis for correlating non-stationary time series can be considered to this end (Podobnik and Stanley, 2008).

3.2. Multimodality Data Analysis

Spatial multimodality indices support the analysis of multimodal transport usage in specific urban areas using inequality measures to assess available and enacted options by citizens

(Lemonde et al., 2020). An inequality measure is a function that describes a distribution of ‘income’ in a way that allows direct and objective comparisons across multiple distributions (Cowell, 2011). Although inequality measures are usually used in socio-economics studies, they can by extended to transportation by reformulating their core properties (Diana and Pirra, 2016):

- Weak Principle of Transfers: considering two travel modes with I and Iδ intensities of use. If the intensity of the most used mode decreases and the other increases in same degree I<2δ, then multimodality increases;
- Scale Independence: if the frequency of use of each mode changes by the same proportion, the multimoda-lity index should remain the same;
- Principle of Population: the multimodality index should remain the same for any replication of the modes with their corresponding intensities of use.

The choice of a suitable index for multimodality analysis will depend on the context of the problem. The Gini coefficient is a summary statistic of the Lorenz curve and is usually used as a measure of inequality in a population. Diana and Pirra (2016) translated the usual formulation of the index to the context of multimodal transportation,

$$Gini = 2/\mathrm{n} \cdot (\textstyle\sum_{i=1}^{n} i \cdot f_i)/(\sum_{i=1}^{n} f_i) - (n+1)/(\mathrm{n}) \tag{1}$$

where f_i is the intensity of use of ith transport mode and n the total number of modes. Another possible measure is the Herfindahl–Hirschman index, a typical measure of market concentration to determine market competitiveness,

$$\mathrm{HH}_m = 1/\mathrm{m} \cdot (\mathrm{n} \cdot \textstyle\sum_{i=1}^{n}(f_i - \bar{f})^2)/((\sum_{i=1}^{n} f_i)^2) + 1) \tag{2}$$

where m corresponds to the effective number of used modes (Diana and Pirra, 2016). Both indices range from zero, corresponding to an equal usage of all modes, up to a maximum of one, which refers to monomodality in the presence of an infinite population of modes.

3.3. Incorporating Situational Context

The analysis of multimodality indices can be enriched with the presence of situational context. A major constituent element of such context is ornament information, specially traffic generation poles (Figure 3), including commercial areas, employment centers such as business parks and enterprises, and collective equipment like hospitals, schools and stadiums, that generate or attract a significant volume of vehicle trips, either from contributors, visitors or providers (IMTT, 2011). The combined analysis of these traffic poles locations against the computed multimodality indices, as well as station-route maps, provides a comprehensive and dynamic way of assessing causal factors pertaining to the spatiotemporal distribution of traffic along the city.

Additionally, the surveyed indices can be revised to further measure how the volume of passengers generated and attracted by nearby poles are being currently satisfied by the co-located modes of public transport.

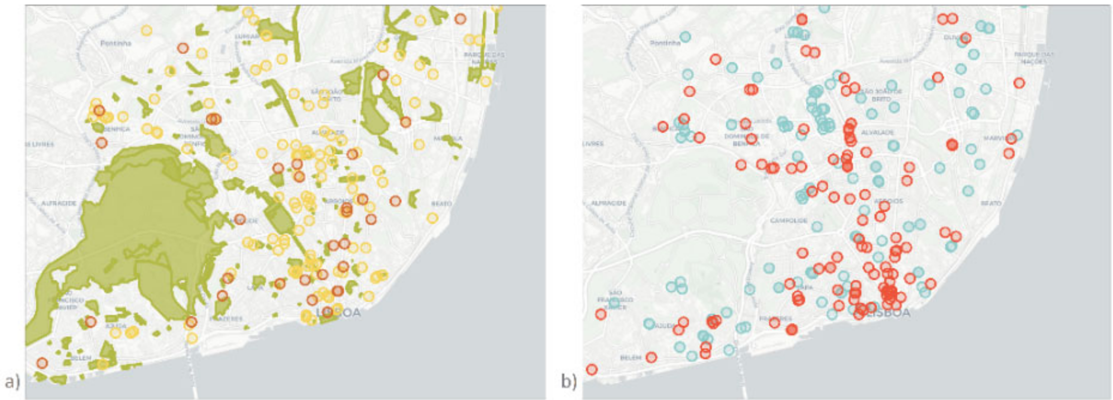

Figure 3: Some traffic generation poles: a) large commercial poles and parks; b) healthcare poles in red and educational poles in blue.

In addition, weather factors and public events can further impact traffic demand and modal choices (e.g. decreased cycling activity under rainy weather conditions). Two major strategies for context-sensitive analysis are suggested to this end. First, data can be segmented according to the available situational context followed by context-specific inference of multimodality patterns (Cerqueira et al., 2020). Second, and in alternative, the context can be directly accommodated in the indices by capturing correlations with the context and using these correlations as correction factors to adjust traffic demand.

4. LISBON'S CASE STUDY

Three public transport modes were considered for this study – bus-and-tram, subway and cycling modes. The bus mode, operated by CARRIS, and the subway mode, operated by METRO, are the two most used public transport modes and their stations offer a good spatial coverage within the Municipality of Lisbon. The stations of GIRA, the biking sharing system, however, can only be found in the center axis of the city and in the neighborhood of Parque das Nações (Figure 4); and the validations during the week are significantly lower than the other modes (Figure 5).

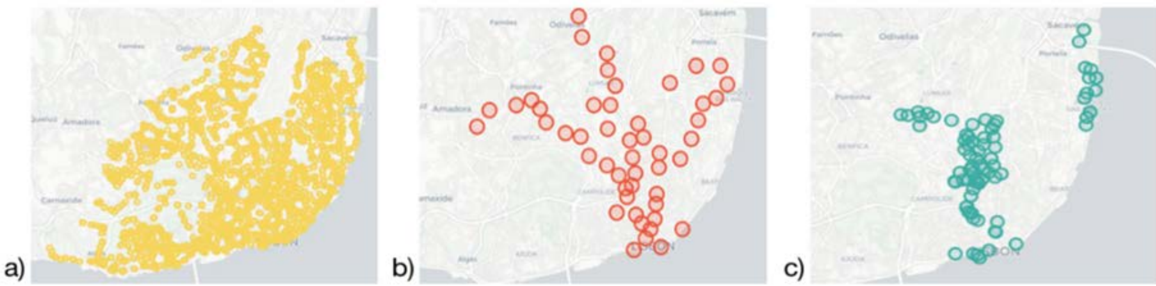

Figure 4: Lisbon's stations location: a) CARRIS, b) METRO, c) GIRA.

Smart card technology was used to gather public transport traffic data. For the bus transportation, as smart card data only monitors entries, estimators are necessary to infer exit validations from vehicle-to-vehicle transfers, daily pendular movements and circulation patterns across days. In this context, multimodal circulation views were considered to capture cross-mode transfers and thus increase the success of the alight stop inference task (Figure 6).

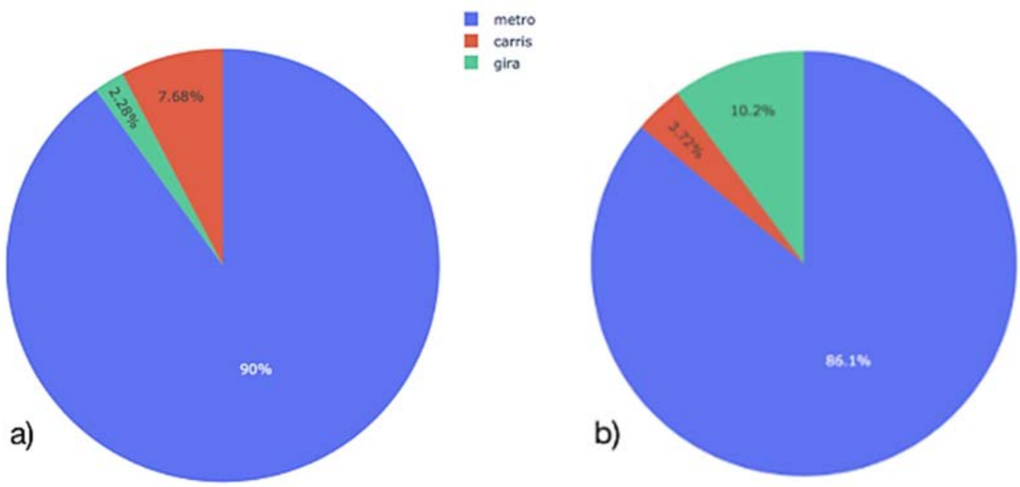

Figure 5: Weekly mode share distribution of TAZ nº 66 (Entrecampos): a) week days, b) weekends.

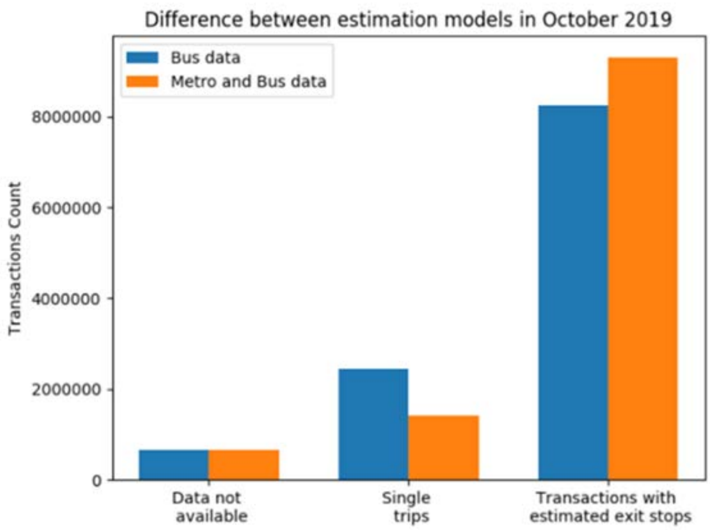

Figure 6: Multimodal alight stop inference success for CARRIS data in October 2019

Traffic Analysis Zones (TAZ) were selected in this study as the spatial granularity criteria (Figure 7). This form of spatial modelling is derived from trip generation densities processed by delineation algorithms that use the peaks of densities as the centre of a zone (Martínez et al., 2009). Subsequently, the public traffic demand is estimated (section 3.1) by retrieving the volume of validations from the stations and stops of the chosen transport modes per zone.

Figure 8 shows the daily volume of validations in TAZ nº66. Figure 9 provides a complementary view of the demand distribution per mode and TAZ.

Figure 7: TAZs of Lisbon Municipality: a) all TAZs, b) TAZs with all modes (subway, bus, bike).

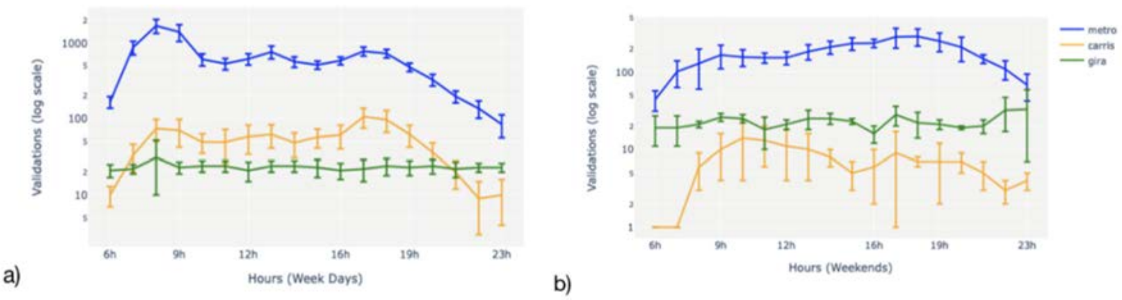

Figure 8: Daily volume and variation of validations in TAZ nº66: a) week days, b) weekends

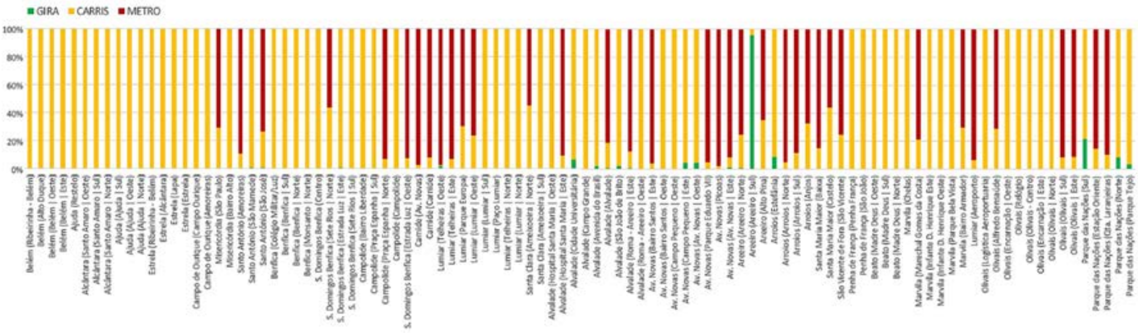

Figure 9: Distribution of demand per transport mode in 2019.

Correction factors corresponding to weather variables (Figure 10) and public events were applied by respectively removing correlation factors on mode-specific demand and replacing demand observed during the spatiotemporal footprint of an event by the average demand on a comparable time period

			station 406	*station 407*	*station 408*	*station 416*	*station 417*	*all*
temperature	check-in	11-13h	0.147	0.178	0.491	0.043	0.05	0.239
		14-16h	0.127	0.255	0.05	0.05	0.088	0.138
	check-out	11-13h	0.112	-0.171	0.273	0.19	-0.057	0.09
		14-16h	0.303	0.082	-0.065	-0.065	0.115	0.167
precipitation	check-in	11-13h	0.124	0.161	0.151	0.251	-0.07	0.161
		14-16h	-0.204	0.017	0.005	-0.163	-0.011	-0.119
	check-out	11-13h	-0.423	-0.146	-0.42	-0.124	-0.237	-0.414
		14-16h	0.146	-0.344	-0.205	-0.267	0.287	-0.068
wind	check-in	11-13h	-0.029	-0.044	-0.033	-0.248	-0.41	-0.288
		14-16h	-0.122	-0.276	-0.116	-0.201	-0.251	-0.268
	check-out	11-13h	-0.417	-0.412	-0.398	-0.147	-0.258	-0.501
		14-16h	-0.14	-0.471	-0.404	-0.332	0.097	-0.337
humidity	check-in	11-13h	0.067	0.278	0.235	-0.112	-0.008	0.111
		14-16h	0.08	0.111	0.027	0.058	0.081	0.1
	check-out	11-13h	-0.107	0.021	0.113	-0.24	-0.09	-0.088
		14-16h	0.244	0.199	-0.159	-0.168	-0.042	0.001

Figure 10: Correction factors produced by the Pearson correlation between weather data and observed check-ins/outs at GIRA's bike stations for 2 hours intervals (7/1/2019 to 28/2/2019).

The public traffic demand from each zone can be correlated with computed spatial multimodality indices, to further detect vulnerabilities in the public transport system of a particular zone. Figure 11 displays four TAZ maps of Lisbon coloured with the values of the Gini index and Herfindahl–Hirschman index respectively at different hours.

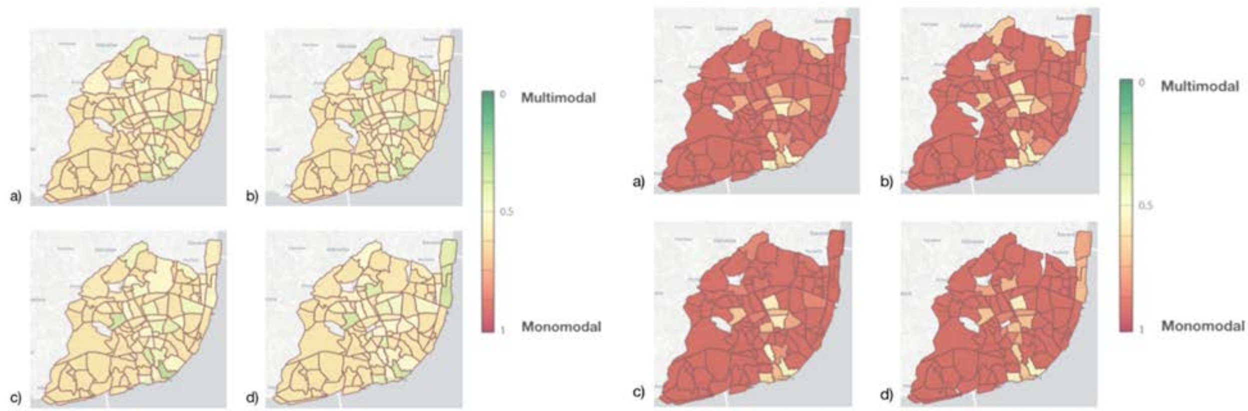

Figure 11: Gini index (left) and HH index (right) on week days: a) 8h, b) 12h, c) 17h, d) 21h.

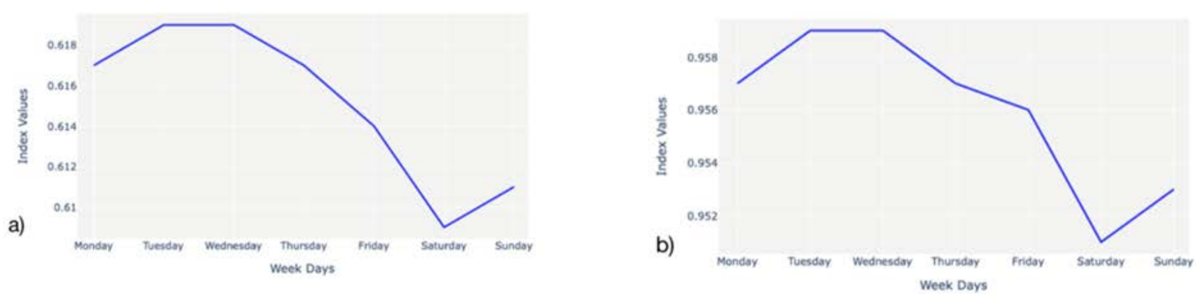

Figure 12: Weekly multimodality index variation: a) Gini, b) HH.

The spatiotemporal analysis of the indices provides an exploratory policy framework to measure social equity regarding the available multimodality options along the zones of an urban region. In the context of the Lisbon case, the TAZs with higher degree of multimodality generally correspond to zones that contain all targeted modes - bus, subway and cycling (Figure 7) - and encompass a large number of traffic generation poles (Figure 3). Considering the Gini index, a few TAZs contain all selected modes yet yield a medium index value (around 0.5). This occurs due to heavy subway usage patterns (scale dependence property). Herfindahl–Hirschman index (HH) results are similar to Gini indices, with the exception that HH is more sensitive to the number of used modes (population property), justifying the red coloring on most TAZs. The variation along a week is coherent for both indices (Figure 12), with a subtle multimodality increase in weekends moved by an increased cycling demand and decreased subway demand.

5. CONCLUDING REMARKS

This work introduced a methodology that offers a solid ground for multimodal traffic data analytics, including a means for the consolidation and efficient retrieval of heterogeneous sources of trip record data; the possibility to estimate missing validations at the entry or exit of stations and vehicles; and the accommodation of correction factors to discount the impact that a given situational context can have on circulation patterns.

We provide preliminary empirical evidence for the relevance of the proposed methodology to aid the calculus of multimodality indices for an initial characterization of mobility restrictions and social equity aspects.

We are currently extending the proposed methodology to accommodate more advanced descriptive and predictive analytics of multimodal traffic, including the discovery of emerging multimodal patterns, inference of origin-destination matrices, and modeling of inter-mode dependencies to assist context-aware predictors of public traffic demand.

These contributions are expected to assist the municipality of Lisbon and comparable cities in moving towards urban mobility plans closely aligned with the real traffic dynamics as objectively given by trip record data, therefore:

- supporting the transparency of urban mobility planning decisions to the citizen
- offering a solid ground for coordination efforts among municipalities and public transport operators
- promoting the continued alignment of the public transport network with the ongoing city transformations, thus ensuring that the public transport system responds to emerging multimodal traffic vulnerabilities, a growing need given the transformations and changing regulations observed in a pandemic context.

ACKNOWLEDGEMENTS

The authors thank CARRIS, METRO and *Câmara Municipal de Lisboa* (*Gabinete de Mobilidade* and *Centro de Operações Integrado*) for the data provision, support and valuable feedback. This work was further supported by *Fundação para a Ciência e Tecnologia* under project ILU (DSAIPA/DS/0111/2018) and INESC-ID pluriannual (UIDB/50021/2020).

REFERENCES

BUEHLER, R. AND HAMRE, A. (2016). An examination of recent trends in multimodal travel behavior among american motorists. International journal of sustainable transportation.

CERQUEIRA, S., ARSENIO, E., AND HENRIQUES, R. (2020). Integrative analysis of traffic and situational context data to support urban mobility planning. In European Transport Conference 2020.

COWELL, F. (2011). Measuring inequality. Oxford University Press.

DIANA, M. AND PIRRA, M. (2016). A comparative assessment of synthetic indices to measure multimodality behaviours. Transportmetrica A: Transport Science, 12(9):771–793.

IMTT (2011). Guia para a elaboração de planos de mobilidade de empresas e pólos. Instituto da Mobilidade e dos Transportes Terrestres (IMTT).

LEITE, I., FINAMORE, A. C., AND HENRIQUES, R. (2020). Context-sensitive modeling of public transport data. In Transport Research Arena 2020.

LEMONDE, C., ARSENIO, E., AND HENRIQUES, R. (2020). Exploring multimodal mobility patterns with big data in the city of lisbon. In European Transport Conference 2020.

MAMOULIS, N., CAO, H., KOLLIOS, G., HADJIELEFTHERIOU, M., TAO, Y., AND CHEUNG, D. W. (2004). Mining, indexing, and querying historical spatiotemporal data. In ACM SIGKDD, pages 236–245.

MARTÍNEZ, L. M., VIEGAS, J. M., AND SILVA, E. A. (2009). A traffic analysis zone definition: a new methodology and algorithm. Transportation, 36(5):581–599.

NEVES, F., FINAMORE, A., AND HENRIQUES, R. (2020). Efficient discovery of emerging patterns in heterogeneous spatiotemporal data from mobile sensors. In EAI Int. Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (Mobiquitous), pages 1–10.

NEVES, F., FINAMORE, A. C., Madeira, S. C., and Henriques, R. (2021). Mining actionable patterns of road mobility from heterogeneous traffic data using biclustering. IEEE Transactions on Intelligent Transportation Systems.

NOBIS, C. (2007). Multimodality: facets and causes of sustainable mobility behavior. Transportation Research Record, 2010(1):35–44.

PODOBNIK, B. AND STANLEY, H. E. (2008). Detrended cross-correlation analysis: a new method for analyzing two nonstationary time series. Physical review letters, 100(8):084102.