Cerqueira *et al.*
European Transport Research Review (2025) 17:30
https://doi.org/10.1186/s12544-025-00720-1

European Transport Research Review

ORIGINAL PAPER Open Access

Data analytics to advance the inference of origin–destination in public transport systems: tracing network vulnerabilities and age-sensitive trip purposes

Sofia Cerqueira[1,2*], Elisabete Arsenio[1], José Barateiro[3] and Rui Henriques[2]

Abstract

Knowing the passengers' final destinations, underlying motifs, and commuting habits is critical to optimise public transportation systems, guide urban planning and contribute to a more sustainable urban mobility. In entry-only Automated Fare Collection systems, the body of literature has focused on the spatial dimension by estimating alighting stops, overlooking the inference of robust alighting times. Moreover, discriminating between transfers and activities is pivotal for determining their ultimate destinations. However, current methods often struggle to adapt to the stochastic nature of passenger behaviour, further disregarding the multiplicity of routes and stops to access specific facilities and individual motivations. Further research is required to address an effective spatio-temporal and contextual inference in both challenges. With the above concerns in mind, this research uses data analytics to propose an enhanced methodology for the inference of OD matrices, with the final goal of providing a comprehensive view of OD mobility patterns across distinct age-sensitive profiles—youth, adults, and older adults. Our methodological framework integrates the following approaches: (i) alighting stop-and-time inference, (ii) ensembled model for transfer classification, (iii) indicators retrieved from statistical analysis of network vulnerabilities (e.g., number of transfers, walkability needs), frequent destinations and their underlying putative motifs against the city amenities and others points-of-interest. The reliability of alighting data (timestamp and location) inference is improved by integrating OpenStreetMap data and the past boarding data from bus and railway systems. Considering Lisbon as the target study case, we apply the methodology over smart card data collected both from metro and bus systems. A comparative analysis with state-of-the-art methods revealed that the enhanced framework for alighting and OD inference led to longer journey times for trips. Furthermore, throughout the day, the older adult group experiences longer transfer times on average compared to both the children and young adult segment and the adult segment.

Keywords Alighting data inference, Transfer classification, Origin–Destination inference, Spatio-temporal data analytics

*Correspondence:
Sofia Cerqueira
scerqueira@lnec.pt
[1] LNEC, Lisbon, Portugal
[2] INESC-ID and IST, Universidade de Lisboa, Lisbon, Portugal
[3] FCT and NOVA-LINCS, University of Algarve, Faro, Portugal

1 Introduction

Public Transport PT and city decision-makers seek to improve people's mobility by obtaining reliable insights into passengers' trip dynamics, especially end-to-end travel needs. For this purpose, Origin–Destination (OD) matrices are pivotal in transportation planning, enabling a comprehensive understanding of travel patterns and progress towards sustainable mobility planning. Yet, this goal is hindered by several challenges, including the existence of stochastic travel behaviour and, in some cases, the incomplete information of entry-only Automatic Fare Collection (AFC) systems, which only record the boardings. Considering that trip data is complete, a critical task for a successful end-to-end OD inference is the discrimination of transfers (intermediary stops) from activities (end stops) between trips. As a guiding assumption for this task, most of the literature uses a predefined time threshold to discriminate stop motifs, instead of finding more personalized mechanisms for dynamically identifying transfers with greater efficacy. Most studies conventionally apply a fixed threshold of around 30 min, assuming that exceeding this limit indicates an activity rather than a transfer [1–4]. However, relying on a static threshold disregards the stochastic nature of user behaviour—for example, elderly individuals may take longer to walk between stops or stations. In addition, variations in transport services in different areas, such as low service availability, traffic congestion, traffic disruptions caused by major events or construction works can further affect transfer times. These assumptions can result in inaccurate estimates of transfer detection. Previous authors emphasize the importance of adapting OD inference to account for factors such as urban demography, network characteristics, and passenger mobility features within cities [5, 6]. To this end, further research is required to devise alternative methodologies that dynamically and effectively capture the stochastic behaviour of the urban transport service. Previous authors emphasize the importance of adapting OD inference to account for factors such as urban demography, network characteristics, and passenger mobility features within cities [5, 6]. To this end, further research is required to devise alternative methodologies that dynamically and effectively capture the stochastic behaviour of the urban transport service.

To address the above concerns, this research proposes a robust framework to improve the inference of OD trips, yielding three major contributions. Firstly, enhanced alighting stop-and-time inference by integrating data from OpenStreetMap to refine the inter-transaction distances, and further using timestamps from boarding data for robust alighting time inference. Transaction term refers to a trip and inter-transaction distance refers to walking distance between consecutive trips. Secondly, a baseline sample with inter-interactions is classified into transfers or activities based on state-of-the-art principles. Thirdly, with robust classification of transfers and activities, origin–destination (OD) data inference is accomplished by training and testing an ensemble classifier on this baseline sample. The trained model will then classify inter-transactions that did not fit into the established state-of-the-art principles. Particularly, this approach is sensitive to relevant temporal and spatial patterns by integrating passenger-specific, spatial and temporal (network) features. Moreover, a principled understanding of passenger behaviours and mobility patterns within a given city transportation system is achieved by transforming the inferred OD data into complementary statistics (e.g., walkability needs, transfer times, multimodal trips).

The application of this methodology to the city of Lisbon offers a solid means to study mobility patterns by user segments (including age groups) in other potential urban contexts. Indeed, the proposed methodology is applicable to a wide range of transportation systems, including public, private, and multimodal networks, which record user trips or partial logs. Regardless of the specific characteristics of the system—such as those based on Automated Fare Collection (AFC), Wi-Fi, or other technologies—the methodology is compatible with systems that record only entries, only exits, or both. It can also be applied to networks with varying numbers of operators and modes of transport, ranging from single-operator systems to complex multi-operator networks. Furthermore, the methodology does not rely on real-time General Transit Feed Specification (GTFS) data. In scenarios where precise vehicle location data is unavailable, it incorporates principles for estimating timings, ensuring both its robustness and adaptability. In the case of Lisbon, this methodology enabled the: (i) characterisation of frequent destination clusters, analysis of the potential travel motives, average time spent in the cluster, and other relevant statistics; (ii) quantification of the degree of mode connectivity (number of transfers) and distribution of transport modes (multimodality) to these main destinations; and (iii) detection of vulnerabilities on the transport network by identifying OD clusters with a high average number of transfers. By studying the OD behaviours of these segments, we aim to recognize their unique needs, thereby promoting social equity and enhancing inclusivity in public transport.

The remainder of this paper is structured as follows: Sect. 2 provides a review of previous related work in OD inference methods. Section 3 introduces the proposed methodology for OD estimation, including the principles of enhancing the reliability of inter-transaction distances and alighting time data. In Sect. 5, the improved OD data is used to provide insights into the mobility dynamics of elderly, as well as children and young adults, as detailed. Section 6 concludes by summarising key findings, practical recommendations and suggests directions for future research.

2 Literature review

Origin–Destination trip inference (OD) in the context of public transport has become crucial for understanding passenger travel behaviour and optimising public transport services. Indeed, effective OD trip inference has significant implications for real-world applications such as city planning, traffic management, and route planning [7, 8]. By shedding light on the travel habits and preferences of passengers, public operators and city councils can tailor public transport to meet changing needs for improved passenger experience. Typically, contributions to OD inference involve two essential stages: (i) imputation of alighting data, and (ii) inference of end-to-end travel flows.

2.1 Alighting inference models and methods

In presence of entry-only AFC systems, the literature review shows that the trip-chaining algorithm (principle-based model) is still the most used approach to infer alighting locations from smart card data [1, 4, 6, 9–12]. The trip-chaining algorithm aims to choose an alighting stop where the distance between it and the boarding stop of the following transaction is minimal and falls below a threshold. This limit is often called the inter-transaction threshold (ITT). Building upon this heuristic, different principles have been proposed to represent various urban mobility scenarios in bus public transportation. The prevailing principles employed in the literature encompass the following: (i) the passenger will alight at the closest to the next boarding point, whose distance is not more than a certain ITT, (ii) at the last trip of the day, the passenger will alight close to the first boarding of the day, whose distance must be below than a certain ITT. The predominant ITTs on the body of the literature range between 600 to 1000 m [1, 11, 12]. Ultimately, other principles were suggested by other authors to attend to the stochasticity of passengers' behaviour and multimodal transport choices. More recently, approaches such as neural networks or probabilistic models are either used to aid the inference after the trip-chaining algorithm or solely [13–18]. Prior work, from the authors Cerqueira et al. [19], provides a critical examination of the existing principles used in trip-chaining methodologies and suggests a three-stage alighting model that takes into account not only the best practices, but also user-centric pattern mining stances to improve inference. The work gathers state-of-the-art principles for trip-chaining from key studies, including Munizaga et al. [20], Barry et al. [9], Alger et al. [21], Nassir et al. [11], Nunes et al. [22], Faroqui et al. [23], and Trepanier et al. [4]. The impact of these principles is individually assessed within trip-chaining models. To address the limitations in estimating travel behaviour using traditional trip-chaining algorithms, two additional algorithms are introduced. The first is a clustering-based algorithm that identifies key spatial–temporal features (e.g., home, school, workplaces) and estimates unobserved trips. The second is a frequent pattern-based algorithm used as the third phase of the estimation process. As an unsupervised learning model, the adopted principles and methods are evaluated using confidence and distance-based metrics. The overall model achieves a 90.54% estimation rate, with an average transfer distance of 105 m. To further validate the approach, the model was tested using only bus data (unimodal) instead of both metro and bus data (multimodal). As expected, the estimation accuracy of the unimodal model was slightly lower at 83%. However, a 73.32 percentage point match was observed between the two models, demonstrating the model's learning capability through clustering and pattern analysis. This previous contribution, whose methodology and essential findings are detailed in the appendix, provides a solid foundation for the subsequent gathering of OD movements.

2.2 OD inference models and methods

The subsequent stage, OD inference, involves passenger trajectory by determining the starting and ending points of their journeys through inference of stop motifs between transactions (transfer or activity). The majority of studies propose an ITT based on time to differentiate transfers from activities [1–3, 11, 24–28]. Notwithstanding the topicality of time thresholds, other heuristics have also been proposed:

H1 Inter-Transaction Time: If the time elapsed between the end of one trip and the boarding of the next trip exceeds a predefined threshold, it is assumed that the user was performing an activity. Otherwise, it is considered a transfer.

Using the H1 principle, Seaborn et al. [3] recognized and evaluated comprehensive multimodal OD trips using fare payment data from Oyster smart cards within the London transit system. To achieve this objective, distinct time thresholds were applied to three various transfer scenarios: transfers from underground to buses (15–25 min),transfers from buses to the underground (30–50 min); and transfers between buses (40–60 min). The author's findings indicate that using the midpoint or lower values of these ranges is effective in accurately identifying journeys, leading to 20% of OD trips involving two transactions, while 5% encompass three or more transactions. Alsger et al. [1] further studied the effect on OD matrices by using different allowable transfer time thresholds (ranging from 30 to 60 min and even 90 min) for the validation of principle H1. The study's findings show that OD matrices with a 30-min allowable transfer time exhibit slightly better accuracy compared to those with 60 and 90-min thresholds. Also, Hamedmoghadam et al. [2] introduce a method for deriving optimal ITT. Inter-transaction times were classified into transfers and activities using a statistical approach. Optimal thresholds are found for weekdays and weekends, 47 min and 52 min respectively. Zhao et al. [29] focus solely on identifying subway-to-bus transfers in Nanjing, China by employing association rules and cluster analysis. The study indicates that transfers take a median time of less than 20 min.

H2 Same-Route: Regardless of the inter-transaction time, if two consecutive transactions occur within the metro system or on the same bus route, it is assumed that an activity occurs between the two transactions.

H3 Time-Based Activity Detection for Missing Alighting Information: For a transaction without alighting point information, if the time elapsed between the boarding of this transaction and the consecutive boarding exceeds a certain threshold, it is assumed that an activity occurred in between.

H4 Transfer Opportunity: If the user uses more than a certain number of transfer opportunities (e.g., $Opp > 1$) between the end of one trip and the boarding of the next, it is assumed that the user is performing an activity. Otherwise, it is considered a transfer.

Munizaga et al. [27] use a 30-min threshold for applying the heuristic H1 in a multimodal network. The work also introduces the principle described in H2, which states that, regardless of the time interval, consecutive transactions on the same route and in the same direction indicate that the passenger has reached a destination after the first transaction. This assumption was later adopted by other researchers, such as Gordon et al. [24] and Yap et al. [28]. Furthermore, Munizaga et al. [27] apply H3 in cases where a transaction lacks alighting information, suggesting that a time gap of over two hours before the next transaction implies an activity occurred within that time frame. While this approach provides a practical solution, the assumption is not explicitly validated in their study. In fact, the user could be transferring to other transport modes that are not explicitly present in the dataset.

Nassir et al. [11] suggested that any activity undertaken by an individual would require a minimum of 30 min. They also assumed that passengers would not wait more than 90 min for a transfer. Within the 30- to 90-min time frame, the authors evaluated whether passengers boarded the first available bus. To do so, they applied H4, which states that if a passenger boards the first available bus, it is considered a transfer,otherwise, it indicates an activity. This assumption was later adopted by researchers such as Gordon et al. [24], Yap et al. [28], and Kumar et al. [26]. However, Hussain et al. [5] suggest that relying on this principle could lead to underestimating transfers when the boarding is denied and transit disruption. Moreover, many studies fail to accurately estimate alighting times by relying on expected service schedules, which often do not account for real-world conditions, such as delays caused by traffic congestion. Without precise timing data for bus services, the calculation of transfer opportunities may be unreliable. As evidence of the time discrepancies, our research shows statistically significant disparities between the actual boarding and the expected schedule.

H5 Journey Ratio (distance-based criterion): For a journey involving more than one trip, if the ratio $f_d = \frac{d_{on-route}}{d_{euclidean}}$ exceeds a certain threshold, it indicates that an activity is occurring. Here, $d_{on-route}$ represents the total travelled distance along the route, and $d_{euclidean}$ is the straight-line distance between the journey's origin and destination.

Gordon et al., [24] propose a new principle to calibrate the OD estimation. The work uses the principle H5 that states that when the ratio between the travelled distance along the journey and the Euclidean distance between the journey's origin and destination exceeds a certain threshold, an activity may have occurred.

H6 Relative Time Ratio: If the ratio between the inter-transaction time (time between trips) and the total travel time of the journey exceeds a specified threshold, it indicates that an activity is occurring. Otherwise, it is classified as a transfer.

H7 Excess Journey Time (time-based difference): Considering a journey involving more than one trip, if the difference between the actual travel time and the fastest possible travel time exceeds a specific threshold, it is assumed an activity is taking place. Otherwise, it is classified as a transfer.

H8 Journey Ratio (Time-based criterion): If the ratio of the difference between the actual travel time and the fastest possible travel time to the total travel time exceeds a specified threshold, it is assumed an activity is taking place. Otherwise, it is classified as a transfer.

Nassir et al. [30] proposed a set of criteria for identifying short transfers, defined as those taking less than 60 min, with longer gaps considered indicative of an activity. The model incorporates several calibration criteria—time gap, off-optimality, gap ratio, off-optimality ratio, and circuity factors. These criteria align with the principles described in H1 and H5 to H8. In particular, off-optimality (principle H5) measures the difference between the actual journey travel time and the fastest possible travel time. This metric is especially suited for cases where the network offers limited direct transit paths between origin–destination (OD) pairs or relies on central transfer hubs. However, these assumptions may not perform well in other network topologies. As highlighted by Hussain et al. [5], this estimation model is specific to a particular area and requires recalibration before being applied to a different region. Moreover, since most of the criteria are time-based, transit disruptions like congestion, mislead the transfer inference.

Overall, previous literature has made considerable progress in defining optimal thresholds and criteria for identifying transfers in various scenarios. While reviewing these principles, a few limitations and areas for improvement have been identified, highlighting opportunities for further refinement. As mentioned previously, the stochastic behaviour of the transport network introduces variability and unpredictability, like diversity in the bus services during the day and week, transit disruptions caused by major events or road disruptions, and user profile itself. Therefore, the attempt to define a single threshold that satisfies the criterion may be overly simplistic and insufficient to account for the complexities of real-world transit systems. Comprehensive review studies, including Hussain et al. [5] and Mohammed and Oke [6], suggest more comprehensive studies are required to address this variability in the OD inference.

To address this, the proposed approach encompasses an ensemble classifier to distinguish the inter-transaction time as transfers and activity stops. We use features related to the inter-transaction time as inputs for our classifier, including user-specific and network-related characteristics such as route, stops, distances between trips, and the type of card associated. The classifier learns to identify patterns in network usage and user behaviour for transfer identification. Subsequently, a sigmoid function adjusts the classifier's probability weights based on the time between transactions. This mechanism aims to refine our predictions by leveraging the detailed trip data, with the goal of uncovering the nuanced demographic and mobility patterns of passengers within the public transport network.

2.3 Segmented OD analysis

Academia, policymakers, and public transport operators are committed to improving public transport systems to better serve the diverse needs of the population, particularly those who rely heavily on these services [31–33]. The elderly individuals, in particular, have become a focal point in the literature addressing social equity in transport planning, as this age segment often faces specific challenges. Quantitative analysis plays a crucial role in evaluating the age-friendliness of public transport systems. The literature suggests various approaches such as assessments of walkability, network analysis, service quality, and multimodality [31, 32, 34, 35].

For instance, the World Health Organization's report on Measuring the Age-Friendliness of Cities provides detailed guidelines for evaluating how well public transport serves the elderly group[33]. Key indicators include proximity to public transportation, the availability of door-to-door services, access to key destinations, such as community centres and healthcare facilities. Bokolo et al. [31] conducted a literature review on improving friendly, inclusive, and safe mobility for senior citizens. The study highlights the importance of ensuring that elderly individuals feel confident about the availability of seamless intermodal or multimodal transport connections. It also advocates for the adoption of emerging technologies, such as AI-based machine learning and data mining, to enhance mobility, walkability, and wayfinding for senior citizens. Similarly, Fatima et al. [32] reviewed travel patterns, mode preferences, and accessibility indices relevant to the accessibility of the elderly segment. The study highlighted there is a gap in research concerning accessibility indices tailored for the elderly segment due to limited data availability.

Measurable targets, continuous monitoring, and the use of these indicators are essential for evaluating whether transport systems meet the needs of diverse populations, particularly those who depend on public transportation, as highlighted by Arsenio et al. [35]. Origin–destination trip data is particularly valuable for this purpose as it provides detailed insights into how (mode), when (time analysis), and where (destinations) people travel. Effective age-sensitive OD trip analysis not only enhances the understanding of travel patterns among different age groups but also informs the development of tailored transportation policies and infrastructure improvements to better serve the elderly and other vulnerable populations.

3 Methodology

This section describes the main steps of the proposed methodology for enhanced origin–destination (OD) trip inference and description. As a definition, an OD trip is the result of a trajectory compression of trip segments, comprising only the starting (origin) and destination points. A trip segment is denoted by a tuple containing boarding and alighting data (such as route, direction, stop and datetime). As most bus public transport systems lack information regarding alighting locations and times, the infernce of this information is a general critical step for the subsequent analysis of transfers and robust OD inference. In this context, the proposed methodology is composed of the following major steps: (i) advances the approaches from previous work for alighting stop and timestamp estimation, (ii) proposes a novel transfer classification approach for solid origin–destination inference, (iii) validates the efficacy of the proposed methodology in diverse transfer inference scenarios. These steps are essential for constructing a comprehensive view of mobility patterns across a multimodal network, with a particular focus on demographic segments like youth (children and young adults), adults, and the older adults, identified through the type of transport card used. A deeper exploration of specific groups persists as a significant knowledge gap, especially when it comes to understanding the mobility patterns and needs of the elderly population. Therefore, this methodology seeks to bridge this gap via segment-aware OD models to explore how demographic factors influence mobility behaviour, such as transfer behaviour and multimodal choices.

3.1 Alighting stop inference

Cerqueira et al. [19] comprehensively outlines the framework for alighting stop inference, which includes three key steps: the trip-chaining algorithm principle-guided approach, passenger-specific alighting inference based on clustering patterns, and frequent pattern mining (details in appendix). Most existing methods rely on Euclidean distance calculations, often underestimating actual walking distances. To address this limitation, we extend the approach by Cerqueira et al. [19] by incorporating a more accurate network-based distance to assess the proximity between stops. We leverage the OpenStreetMap pedestrian network along with the Python library osmnx to compute distances, considering real walking pathways.

Notwithstanding, the absence of complete and up-to-date data in certain regions may undermine the accuracy and generalizability of the analysis across different geographic areas.

Given the resource-intensive nature of these computations, we use PostGIS, an open-source relational database with robust geospatial query capabilities, to first identify stops within a 1000-m radius. PostGIS facilitates efficient storage, indexing, and querying of geospatial data. Subsequently, OpenStreetMap is employed to ascertain the precise network walking distances between these identified stops.

3.2 Alighting timestamp inference

The topic of defining principles for estimating alighting times receives less attention in the literature compared to the estimation of stops locations. Contributions in this domain generally rely on expected schedules provided by transportation operators, known as the General Transit Feed Specification (GTFS) [11, 13, 26]. However, observing the boarding time data reveals an inconsistency with the timestamps in the GTFS timetables. Recently, transit agencies stream real-time GTFS data (GTFS-RT), which encompasses vehicle positions and timings. As the endpoint for accessing real-time GTFS data may not be available, we propose an algorithmic step to infer vehicle timestamps at each stop along a service based on available smart card data. Accurate alighting time estimation becomes crucial, especially when mobility pattern analysis depends on precise temporal intervals.

To tackle the above issue, this study proposes a methodological framework for the inference of alighting times using historical boarding data. Across a route service, three scenarios require distinct inference methods. For illustrative purposes, Fig. 1 shows potential scenarios for alighting time inference within a bus service. To clarify, a bus service is defined as a route undertaken once a day between the route terminals. The red line represents the route, the green dots denote stops where at least one passenger boarded, and the blue dots indicate stops where no passengers boarded. Thus, by using the stop labels depicted in Fig. 1, three potential inference approaches are described:

1. *Best Scenario:* When a transaction is recorded at a stop s_i, the boarding timestamp is reused as alighting timestamp for passengers' trips alighting at the same stop and service instance. Considering Fig. 1, this applies to stops labelled as s_2, s_4,s_5, and s_7.
2. *Relative Inference:* In this scenario, the determination of the alighting timestamp of stop s_i is contingent upon the boarding timestamps at stops s_k and s_w, where $k < i < w$. This applies to stops labelled as s_3 and s_6 in Fig. 1. This inference is defined by:

$$s_{iTime}(s_k, s_w) = s_{iTime} + \Delta(s_{kTime}, s_{wTime}) * AvgW(s_k, s_i), \tag{1}$$

where s_k, s_i and s_w respectively denote the stops, while $k < i < w$, the term $\Delta(s_{kTime}, s_{wTime})$ connotes the temporal interval between the alighting times of S_k and S_k, and

$$AvgW(s_k, s_i) = \frac{\sum_{j=1}^{n} weight_j(s_k, s_i)}{n}, \tag{2}$$

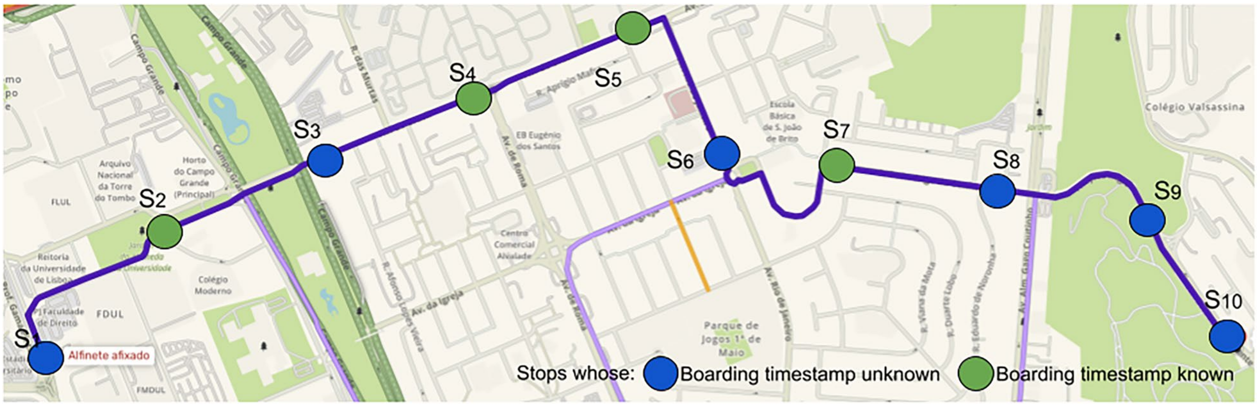


Fig. 1 Illustration of a Bus Route Service with Assigned Boarding Timestamps

where $AvgW(s_k, s_i)$ ranges between 0 and 1 and it is the average weight of the all eligible bus services (n) traversing the (s_k, s_i) segment considering the total time between s_k and s_w

3. *Projection*: This solution is employed when the target alighting stop lacks adjacent stops with identified boarding timestamps. This scenario often occurs at the stops close to the route terminal, which is the case of the stops labelled as s_8,s_9, and s_{10} in Fig. 1. Considering s_w alighting timestamp inference, boarding times at stops s_k , s_i and s_w (where $k < i < w$) from same-route, same-direction services on the same day of the week are gathered. Using this data, then

$$s_{wTime}(s_k, s_i) = s_{iTime} + \frac{\Delta(s_{kTime}, s_{iTimes})}{AvgW(s_k, s_i)} - \Delta(s_{kTime}, s_{iTimes}), \tag{3}$$

where s_k, s_i and s_w respectively denote the stops, while $k < i < w$. Stop s_1 in Fig. 1 does not fit any previous scenario due to the absence of boardings and alightings.

3.3 Origin–destination inference for public transport trips

Passenger OD trip inference is achieved by compressing trajectory data to retain only origin and destination information. This requires classifying inter-transaction events into one of two categories: activity (indicating the end of the OD trip) or transfer time.

The proposed approach consists of three major phases: (i) establishing a baseline sample based on state-of-the-art principles; (ii) training a classifier using the baseline sample; (iii) applying the trained classifier to the remaining data. As previously mentioned, the first phase adopts widely accepted principles for OD estimation (mentioned in literature review section), which are:

- H1. *Inter-Transaction Time*: If the time elapsed between the end of one trip and the boarding of the next trip exceeds a predefined threshold, it is assumed that the user was performing an activity. Otherwise, it is considered a transfer. The consensus asserts that below 30 min is unequivocally considered a transfer.
- H2. *Same-Route*: Regardless of the inter-transaction time, if two consecutive transactions occur within the metro system or on the same bus route, it is assumed that an activity occurs between the two transactions.
- H3. *Time-Based Activity Detection for Missing Alighting Information*: For a transaction without alighting point information, if the time elapsed between the boarding of this transaction and the consecutive boarding exceeds 3 h, it is assumed that an activity occurred in between.
- H4. *Transfer Opportunity*: If, between the end of one trip and the boarding of the next, the user had more than a certain number of transfer opportunities (e.g., Opp > 1), it is assumed that the user was performing an activity. Otherwise, it is considered a transfer.
- *H10. Upper Threshold (extension of principle H1):* After 90 min between transactions, the inter-transaction time is considered to indicate that an activity was performed.

There is a consensus in the literature that inter-transaction times below 30 min are widely accepted as transfers. However, for ITT durations exceeding this threshold, transfer definitions vary across studies. For example, Hamedmoghadam et al. [2] proposed different thresholds for weekdays and weekends, setting them at 47 min and 52 min, respectively. Other studies have also explored varying thresholds (Alsger et al., 2016) [2, 3, 11, 24–28].

Considering that various factors influence transfer time, the proposed approach uses an ensemble classifier, which is trained and tested using a baseline sample (transactions that meet principles H1–H5). This classifier will assign classifications to data that do not meet the previously described principles (H1–H5). The baseline dataset consists of inter-transaction segments from passengers, enriched with trip features and variables derived from feature engineering. These features are: (i) boarding stop, (ii) alighting stop, (iii) route, (iv) weekday, (v) hour, (vi) user card type, (vii) inter-transaction distance, (viii) boarding stop in next transaction, and (ix) route of the next transaction. These segments are labelled as **0** (activity) or **1** (transfer). The validation dataset, on the other hand, includes data that do not meet the H1-H5 criteria. Overall, approximately 91% of the data is used for training and testing the model, while the remaining 9% is reserved for inferring the missing data.

The proposed approach employs a weighted classification model that combines probabilities from an XGBClassifier and a sigmoid function, as shown in the Eq. (4) and (5).

$$P_{combined}(T) = P_{XGBClassifier}(T) + P_{Sigmoid}(T) \tag{4}$$

where, $P_{XGBClassifier}$ is the probability output from the XGBClassifier and $P_{Sigmoid}(T)$ is the probability calculated using the sigmoid function applied to the inter-transaction time after transaction T. $P_{combined}(T)$ is the final probability of time of transaction T is considered a transfer.

$$P_{sigmoid}(T) = \frac{1}{1 + e^{(-T_{ITT}+4)}} \tag{5}$$

where, T_{ITT} is the input value of inter-transaction time taken after transaction T. The inter-transaction time is scaled to fall within a range of 0 to 6 to ensure numerical stability and consistency in the sigmoid calculation. Any values of inter-transaction time exceeding 90 minutes are capped at 90 minutes, ensuring that extreme outliers do not overly influence the resulting probability. The addition of $\Delta = 4$ in the exponent effectively shifts the curve, adjusting the probability output for specific inter-transaction time values.

The classifier focuses on the main features of transactions, excluding transfer time, while the sigmoid function provides a probability based on the transfer time. Separating transfer time from the classifier ensures that the classification process becomes less reliant on this feature, which could otherwise be misleading. This is particularly important because the dataset is imbalanced, with class distributions varying significantly across different ranges of the transfer time feature.

XGBClassifier was selected from empirical evidence and its known ability to handle complex, non-linear data patterns and effectively learn from imbalanced data. The model leverages gradient boosting, which optimally combines the predictions of "weak learners", typically decision trees learnt from subsamples and feature subspaces. XGBoost computes class probabilities via a weighted sum of the outputs of individual decision trees (log-odds). This iterative process ensures that the predicted probabilities are well-calibrated and aligned with the true likelihoods of each class.

OD Aggregation: Following the classification of transfers or activities, trajectory compression is achieved by consolidating travel segments into a unified OD trip, describing only origin information (route features, stop location, and timestamp) and destination details (route features, stop location, and timestamp). For exploration analysis, relevant statistics are computed for each OD trip, including: (i) number of transfers, (ii) stay time at the destination, (iii) transfer time, and (iv) travel time.

Validation: In scenarios lacking validation data, researchers often compare their findings with manual traffic counts or surveys or provide comprehensive sensitivity analysis. In our case, the validation process involves performance analysis conducted on the learning data, divided into training (70%) and testing (30%) subsets for this end. The model's capability to generalize is evaluated using metrics such as recall, accuracy, precision, and the F1 score to assess the model's effectiveness. For data that lacks ground truth labels, exploratory analyses are conducted to identify potential patterns and insights, ensuring a robust and thorough evaluation of the model's behavior.

4 Lisbon as a case study

Smart card data from the public bus and metro (subway) transportation systems in Lisbon city serve both as validation resources for this study and for analysing urban mobility patterns. In total, 44 million transactions were collected from October 2019, with 11 million belonging to the public bus operator Carris and 33 million from the public metro operator METRO.

One of the challenges in this urban mobility study involves understanding the relationship between passenger destinations and the urban characteristics of their surroundings. In this regard, data sourced from the Lisbon Data Portal and OpenStreetMap were used

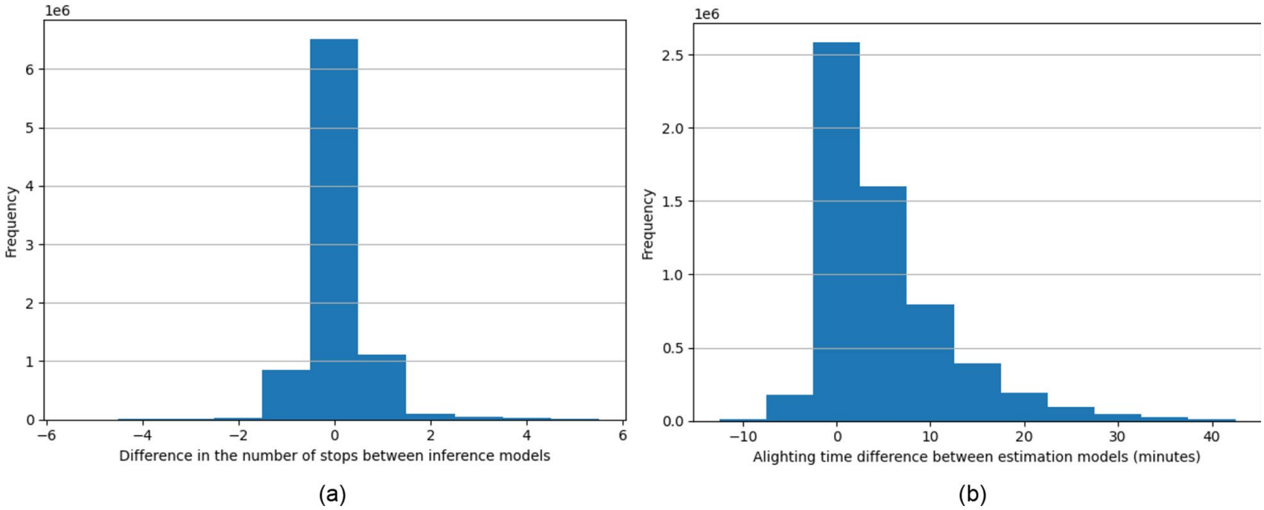


Fig. 2 **a** Distribution of alighting stops, showing the variation in stop count between the conventional and new model; **b** Distribution of alighting stops, presenting the time difference between the introduced alighting timestamp inference method and the conventional approach

to identify commercial spaces, leisure areas, educational institutions, healthcare facilities, citizen service points, sports facilities, and arts venues. These spaces were identified within a 300-m radius of each bus stop or metro station.

A secondary challenge involves the consolidation of the multimodal network, including the segmentation of bus stops, stations, and metro stations into cohesive clusters. To tackle this task, the spatial K-means clustering is implemented, serving as the means to establish these clusters. Given the Lisbon scenario, characterized by approximately 2000 stops/stations, the optimal number of clusters was determined to be 1000, resulting in clusters containing 1 to 3 bus and/or metro stops each. This zoning of Lisbon into these clusters aggregates stops in the same location (including stops in different traffic directions, and redundant stops to prevent overcrowding), therefore facilitating both comprehensive and detailed analyses of passenger destinations, OD trips, and other mobility patterns.

5 Results

5.1 Validation results

5.1.1 Analysis of alighting data inference

Contrasting with the usual Euclidean distance method, this study introduces an enhanced approach for determining walking distances, to support the alighting inference. By leveraging walking distances sourced from OpenStreetMap data, disparities in computed distances have been revealed compared to the traditional method. Figure 2a depicts the number difference of stops, upstream or downstream, between the two methods. The estimated alighting stops for approximately 26% of the transactions (2.75 million transactions) differ by one or two stops before or after from those indicated by the traditional Euclidean distance method. It's worth noting that the Euclidean distance represents the shortest path and not always capture the true walkability needs within urban areas [34]. The walkability is better represented by the pedestrian network from Open Street Maps.

Furthermore, Fig. 2b shows the distribution of time differences between the introduced alighting timestamp inference approach and the conventional method based on expected timetables. The results demonstrate that employing the enhanced approach leads to longer travel times, extending between 1 to 40 min longer. The time estimates give an accurate portrayal of real-world conditions, i.e. bus delays caused by several reasons like traffic congestion, occurrence of attractor events, disruptions in road like accidents. A realistic depiction of alighting times not only enhances transfer identification but can also be used for increasing the reliability of ongoing public transit schedules in the absence of real-time GTFS data.

Both findings (Fig. 2a, b) indicate the importance of establishing more precise and contextual data to produce outcomes that are more in line with real-world scenarios (e.g., traffic congestion).

5.1.2 Analysis of transfer data inference

In this section, we validate the performance of our inference model designed for transfer classification. The validation process involved training and testing. The model was trained and tested on a baseline sample to assess its effectiveness. This approach helps in understanding how well the model generalizes to unseen data. We evaluated

Table 1 Feature importance analysis for the transfer classification model

Feature	Boarding Station	Alighting Station	Route	Card title	Inter-transaction distance	Boarding Station in the next transaction	Route in the next transaction	Weekday	Hour
Gain—Model I	3.52	4.18	3.53	2.90	4.18	4.50	3.18	1.92	2.76
Gain—Model II	6.18	7.53	5.15	5.39	335.48	9.68	4.99	2.09	4.35

Table 2 Performance metrics for transfer classification models I and II

Metrics	Accuracy (%)	Precision (%)	Validation recall (%)	F1 Score (%)	True positives (%)	True negatives (%)	False positives (%)	False negative (%)
Baseline Model	86.29	85.79	86.26	85.37	13.34	72.4	10.41	3.20
Model I	95.96	96.12	95.96	96.01	24.89	71.5	3.05	0.54
Model II	98.08	98.17	98.08	98.11	16.89	81.19	1.53	0.39

the model along two major steps: (i) feature importance analysis of the XGBoost classifier, and (ii) performance analysis using accuracy, precision, recall, and F1-score.

The baseline model, referred to as Baseline Model uses solely the XGBoost classifier without the application of a sigmoid function to translate inter-transaction time. This approach allows for an assessment of the model's generalizability using only network and user features. Later, for a more detailed analysis, the study was segmented by transaction types. Model I focuses on bus transactions that are followed either by another bus or a metro transaction. This model is critical as it specifically evaluates the model's effectiveness in predicting transfers from bus to other transport modes. Model II, on the other hand, encompasses all types of transactions where the subsequent transaction could be either a metro or a bus ride. This categorization helps in achieving clearer interpretations of the results, enabling a focused analysis on the model's capability to predict transfer behaviours across different transaction scenarios.

The XGBoost classifier offers two in-built metrics for feature importance—gain and weight. Gain is considered as the primary criteria as it specifically assesses a feature's impact on improving the model's performance—features with higher gain values contribute most significantly to reducing the model's loss by providing splits that better separate the data. Table 1 indicates that all features show considerable impact; however, the model heavily depends on information about where individuals board and alight, which is crucial in both versions of the model. The significance of both current and future routes, as well as boarding and alighting stations, underscores the model's ability to grasp the complex transfer behaviours within a transit network. This suggests that certain stations or routes serve as critical transfer points. At the Model II, inter-transaction distance shows a significant impact on the model's decisions. The underlying reason for this phenomenon is that the model's decisions are heavily weighted by the substantial volume of subway-to-subway transactions, which typically involve an inter-transaction distance of zero. Furthermore, as the type of card used can indicate user demographics (such as monthly cards for seniors), this feature contributes to the model's predictions by capturing travel habits that vary across different user profiles (e.g., older adults, children and young adults, and other groups).

Considering both the training and testing phases, Table 2 provides an in-depth evaluation of the classifiers' performance, illustrating their capabilities via multiple scoring metrics for Model I and Model II. This table also explicitly details the percentages of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) relative to the test sample size. Despite dealing with an imbalanced dataset dominated by the negative class (activity class), both models demonstrate significant performance. However, Model II shows superior results across nearly all metrics, indicating an advancement over Model I.

Figure 3 provides the distribution of TN, TP, FN, FP for each sample (considering Model II), categorized by mode shift: (i) subway to subway, (ii) subway to bus, (iii) bus to subway and, (iv) bus to bus. When the mode shift involves Subway to Subway or Bus to Subway, the model exhibits a pronounced capability in accurately predicting the True Negatives (activities) and True Positives (transfer) based on inter-transaction time, respectively. Notably, the majority of cases in the 'Subway to Subway' category are classified as activities, whereas 'Bus to Subway' transactions are predominantly identified as transfers.

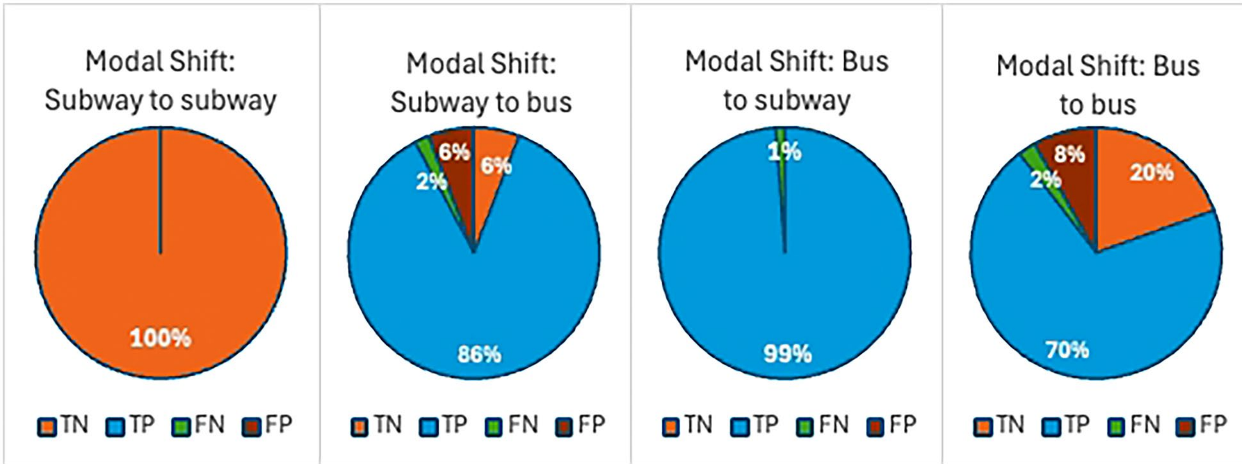


Fig. 3 Distribution of True Negatives (TN), True Positives (TP), False Negatives (FN), and False Positives (FP) for each sample (considering Model II), categorized by mode shift (Subway to subway, subway to bus, bus to subway and bus to bus

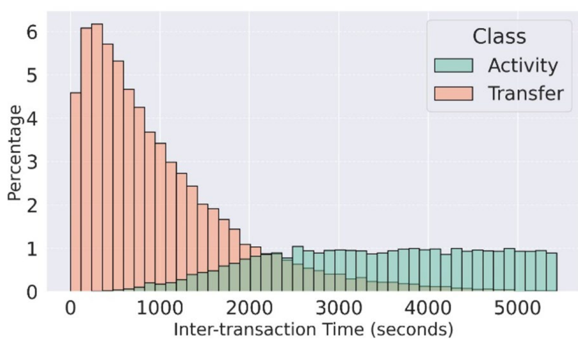


Fig. 4 Transfer and activity classification based on inter-transaction time: results for the train-test sample (left histogram) and for the sample without ground truth (right histogram)

Meanwhile, for Subway to Bus and Bus to Bus shifts, where transfers are notably frequent, the model demonstrates its capability by correctly identifying these transfers with TP rates of 86.3% and 70%, respectively. However, in these scenarios, there is also a more noticeable presence of FN and FP, which suggests that the model faces greater challenges in accurately distinguishing between transfers and regular activities for these particular mode transitions.

Figure 4 illustrates the distribution of inter-transaction segments with respective transfer times for two separate categories: “Transfer” and “Activity”. The histogram bars reflect the percentage distribution considering the total sample size. This sample comprises both the baseline data and transactions where classification was imputed due to the absence of prior classification. As expected, the frequency of transfers is notably higher for shorter intervals, diminishing as the transfer time lengthens. On the other hand, the number of transactions labelled as “Activity” tends to increase with longer inter-transaction durations. While the conventional method employs a threshold rule that restricts transfers to durations below 30 min, the new approach uncovers longer transfers that had been previously disregarded. This dataset will serve as the basis for an exploratory analysis examining different user segments, including the elderly, youth (children and young adults), and adults.

5.2 Profile sensitive results

The proposed methodology is herein validated and applied to acquire a comprehensive understanding of public mobility dynamics within a capital city, ensuring sensitivity to different passenger segments, encompassing adults (24–65 years old), older adults (age > 65), and children and young adults (4–23 years old). The age of each segment is determined based on the card type used. Specifically, for each segment, the following outcomes are presented: (i) characterisation of prominent destination clusters and potential travel motivations; (ii) quantification of the degree of connectivity and transport mode distribution to these primary destinations; (iii) profiling vulnerabilities by highlighting origin–destination trips between clusters characterised by high transfer values, along with other pertinent statistics. To accomplish these objectives, we first provide an exploration and validation analysis is the estimated data by: (i) showing the significant impact of the enhanced framework for location and timestamp inference and (ii) providing a comprehensive analysis of the outcomes from the OD inference model.

Fig. 5 Segment analysis. **a** The distribution of average transfer durations throughout the day in 15-min intervals. **b** The distribution of average post-arrival stays durations throughout the day in 15-min intervals

5.2.1 Analysis of OD trips inference

This section provides an exploration of OD trips, accompanied by pertinent statistics such as the count of transfers, transfer times, and stay durations at the destination. This aims to comprehensively assess the robustness of the proposed approach for the OD trip inference model.

Table 1 describes varying OD travel behaviours among different age groups. The results show that the proportions of trips with multiple transfers range from 20 to 26%, with the older adult segment showing a slightly higher percentage of trips with multiple transfers compared to the other groups. The second row of each segment indicates the subway usage within the N-transfer category. Children and young adult segment and adult segment show a strong preference for zero-transfer trips and increasingly rely on the subway as the number of transfers grows (ranging from 67 to 73% of multimodal OD trips with transfers). Concerning the older adult segment, although subway usage also increases with the number of transfers, the overall percentage of subway usage is significantly lower for the older adult segment, starting at just 22% for zero-transfer trips and rising to 33% for two transfers. The older adult segment tends to rely more on bus-only trips, especially when transfers are required. Several factors may contribute to this preference, including a lack of familiarity with the subway system, the complexity of transfers within subway stations, accessibility challenges, or even the convenience of using bus transport for small tasks that require short transfer times. Additionally, the process of purchasing tickets for the subway, which is often limited to machines, can be inconvenient or unfamiliar to some. These findings highlight opportunities for transit operators to further investigate and better address the needs of older adult passengers.

Figure 5a displays the distribution of average of accumulated transfer durations throughout the day in 20-min intervals after the arrival at the destination. The findings show an equal tendency across all demographics, with the older adult segment showing a higher transfer duration compared to the other groups. The extended transfer times observed among the older adult segment can be attributed to several factors, including their reduced mobility and the increased number of transfers they tend to make, which collectively contribute to a higher overall transfer time. On the other hand, during morning hours, the children and young adults segment experiences the shortest transfer times, primarily because of their higher levels of mobility and lower number of transfers. Figure 5a shows that the average range for accumulated transfer times falls between 6 and 14 min.

Figure 5b shows the distribution of average stay time at the destination durations distributed in the day in 20-min intervals. Again, we observe that the older adult segment has a different pattern from the children and young adult segment and the adult segment. The analysis of children and young adult segment indicates that between 7 AM and 12 AM, the average stay time ranges from 7.5 h to 13h, which aligns with the school hours. After 3 PM, the stay time extends until around 6 PM, likely indicating

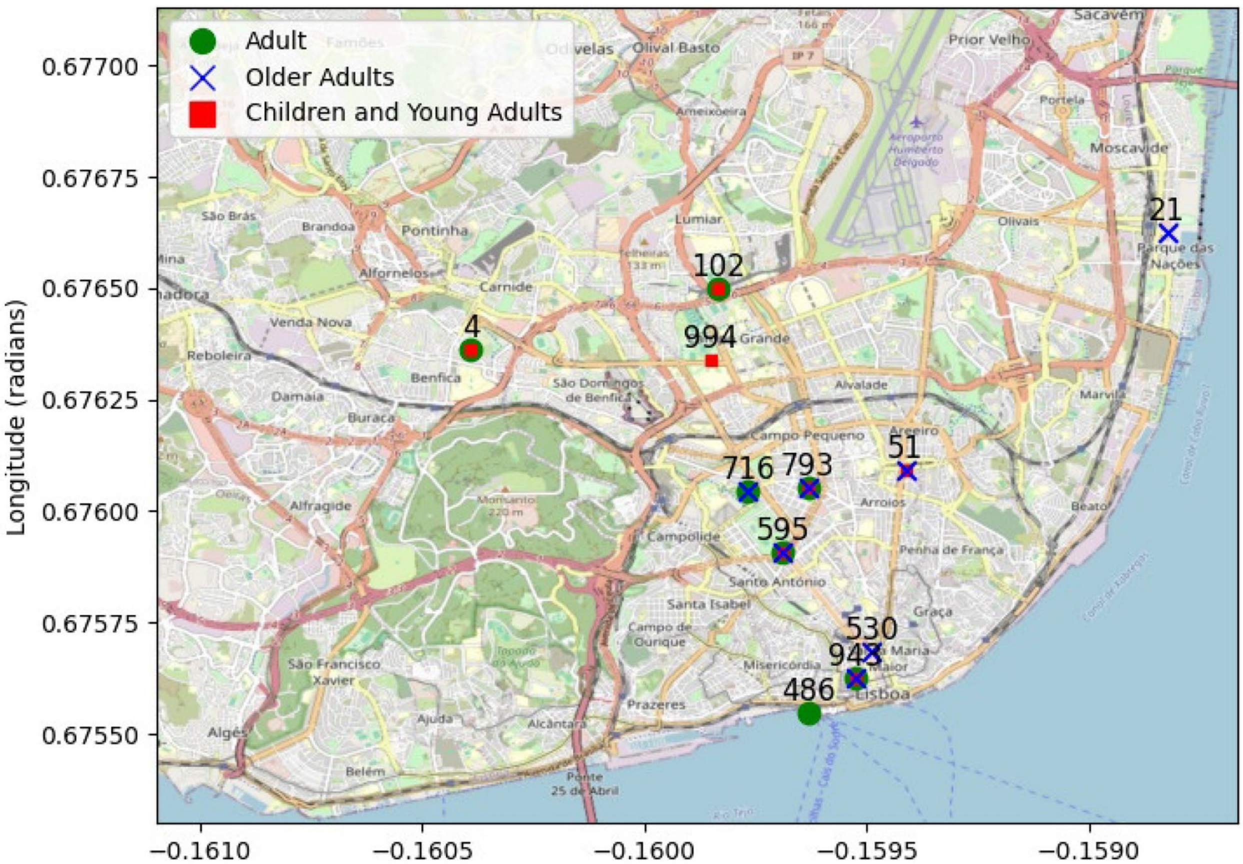


Fig. 6 Spatial distribution of the highest volume stop destinations in Lisbon for the three segments: Children and Young Adults (square markers), Adults (circle markers), and Older Adults (X markers)

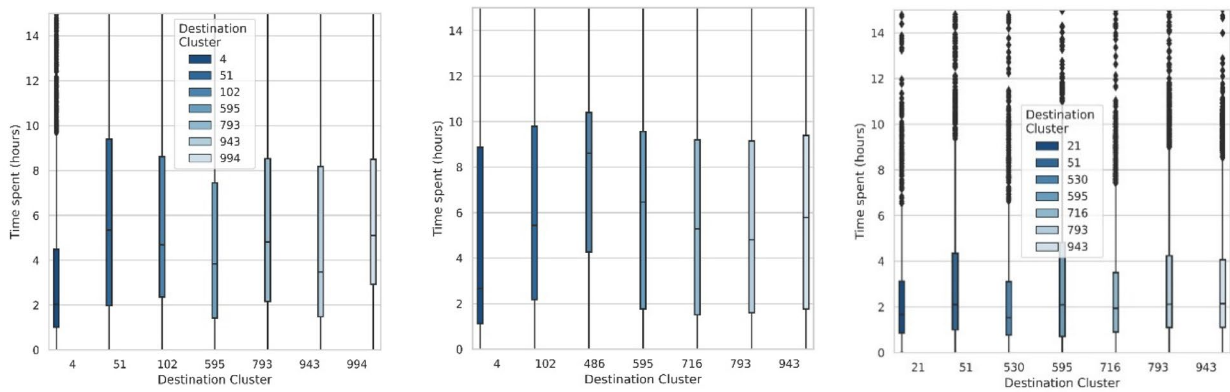
that these destinations predominantly correspond to passengers' residences. Yet, older adult segment exhibit two peaks in this metric throughout the day. The first occurs around noon, suggesting that older adults tend to accomplish their tasks in the morning and then they return home. The second peak emerges after 5 PM. Here, the stay times are considerably higher than in other segments, likely due to fewer daily trips over the course of the week, skewing the distribution and its average.

5.2.2 Analysis of high-frequent destination clusters

This section provides an exploratory analysis of the small areas (clusters) that users select as their destinations. To achieve this, K-means clustering was applied to the bus and subway network, grouping nearby stops into clusters. Each cluster consists of two or three stops, effectively dividing the city into smaller, localized areas. This method allows for a more granular analysis of user destinations by associating them with specific, manageable zones within the network.

The seven most highly frequent destinations for each segment are depicted in Fig. 6. A more in-depth analysis of these clusters reveals that each cluster contains at least one station that significantly contributes to its status as a high-volume destination. Other metrics associated with these clusters demonstrate consistent outcomes across all segments. The average transfer rate ranges from 0.15 to 0.26, signifying a notably low level. The mean travel time averages around 20 min, while multimodality rates fall within the range of 85% to 94%. Furthermore, when transfers do occur within these clusters, they generally take around 15 min.

Figure 7 shows the arrival stay time distribution while Fig. 8 shows the distribution of potential trip purposes among all segments in the main destination clusters. Regarding the post-arrival time depicted in the boxplots, the median duration stands at approximately 4 h for children and young adult segment (except in cluster 4), ranging from 5 to 9 h for adults (except in cluster 4), and around 2 h for the older adult segment. Notably, cluster 4 emerges as a prevailing pattern across two segments, characterised by a consistent median time of 2 h. This scenario can potentially be attributed to activities related to the unique big attractor-generator in the area—a large



a) Arrival stay time distribution for the children and young adult segment.

b) Arrival stay time distribution for the adult segment.

c) Arrival stay time distribution for the older adult segment.

Fig. 7 Spent time in the top frequent destinations

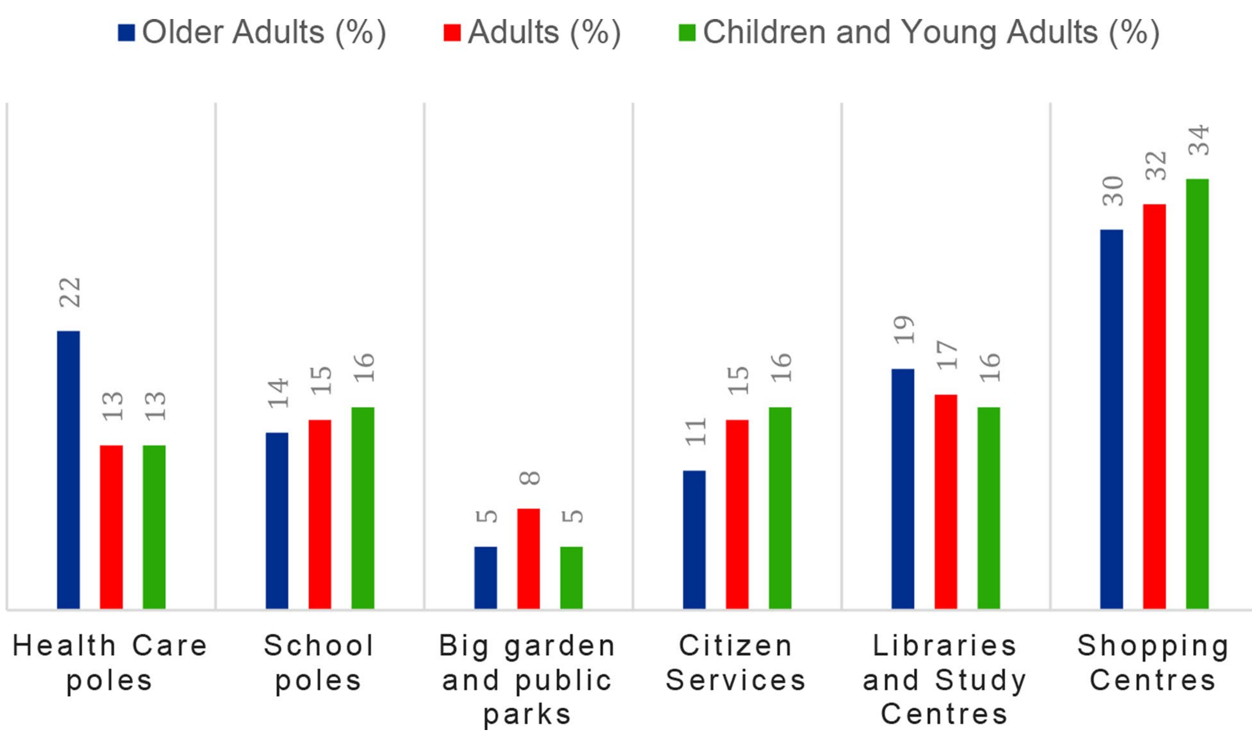


Fig. 8 Potential trip purposes among all segments in the main destination clusters

shopping mall. Notably, in this specific cluster, older adult segment spends less time at their destinations compared to other segments.

Figure 8 depicts the distribution of potential trip purposes within the respective highly frequented destination clusters for each segment. The pie charts highlight that school-related destinations are slightly more prevalent in the children and young adult segment compared to the clusters from other age segments. Health-related destinations, such as primary healthcare centres and hospitals, exhibit greater frequency in clusters associated with the older adult segment. Certain categories consistently predominate across all segments due to their widespread presence throughout the city, thereby posing challenges to trip purpose identification.

Table 3 Percentage of OD trips per segment with N-transfers equal to 0–3 and percentage of trips with at least one trip ad subway

% OD trip with N-transfers	N = 0	N = 1	N = 2	N = 3
Children and Young Adults (4 < age < = 23 years)	82%	16%	1.4%	0.6%
% Subway		Multimodal trips		
	58%	70%	73%	72%
Adults (23 < age < = 65)	81%	16%	1%	1%
% Subway		Multimodal trips		
	61%	70%	67%	67%
Older Adults (age > 65)	74%	21%	4%	1%
% Subway		Multimodal trips		
	22%	30%	33%	31%

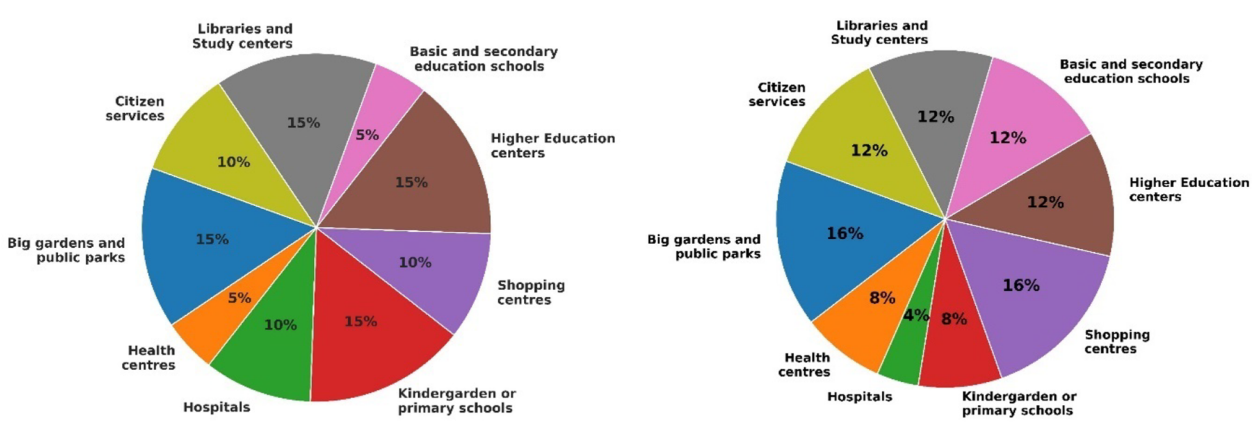
5.2.3 Analysis of vulnerabilities in the origin destination clusters

This section aims to explore and analyse end-to-end trips to destinations associated specific vulnerabilities for the older adults and the children and young adult passenger segments (Tables 2 and 3). "Vulnerabilities" are defined as destinations where at least 80% of trips in the origin–destination sample require one or more transfers to reach them. Given the unique needs of the older adults and the children and young adult passengers, it is crucial for public transit operations to provide optimized, accessible routes that minimize transfers and offer convenient, direct travel options. This need becomes especially important when the destination serves essential purposes, such as schools for children and young adult passengers or hospitals for older adult passengers, as highlighted in Fig. 9. This analysis aims to equip public transit operators with insights to enhance the equity and accessibility of their services.

Tables 4 and 5 provide an analysis on trips between OD clusters in the children and young segment and older adult segment. In both tables, the mean number of transfers exceeds 0.8, indicating that passengers consistently require to board at least two public transports during their journeys.

Table 4 shows the metrics related to the OD cluster in the segment of children and young people. As we observe in Fig. 9a, these destinations are predominantly associated with education poles (50%). The mean transfer times and travel times vary across clusters, but overall, the average end-to-end trip duration ranges from 40 min to 1 h, without including the waiting time prior to the start of the journey. Given that a substantial portion of OD trips involve bus and metro services, minimising bus travel duration can yield notable benefits. For example, introducing solutions such as increased service availability during specific hours or implementing more direct routes can significantly improve travel experiences for students.

Table 5 displays a statistical analysis of OD clusters within the older adult segment, where the mean transfer value exceeds 0.8. The OD clusters collectively indicate that the combined travel and transfer time spans between 16 and 45 min. Specifically, in one of the clusters that has a health centre within the area of the destination cluster, the travel time more transfer time is around 45. Comparing travel purposes, shopping centres is more prominent within the older adult segment, while education centers are more evident in the children and young adult segment. As indicated by prior findings, the elderly population exhibits a lesser inclination to use metro services,



a) Children and young adult segment.

b) Older adult segment

Fig. 9 Distribution of potential trip purposes among all segments in the destination clusters with mean value transfer of transfers higher than 0.8

Table 4 Statistical analysis of OD clusters in the children and young segment, where the mean number of transfers exceeds 0.8

Main origin cluster	Destination cluster	Number of OD trips	OD trips with transfers (%)	Mean transfer time (minutes)	Mean stay time (hours)	Multimodality percentage in OD with transfers	Mean travel time	Education centres close to the destination cluster
<102>	<233>	110	100%	7 min	6 h	99%	41 min	Yes
<102>	<979>	162	100%	10 min	8 h	100%	42 min	Yes
<352>	<979>	108	100%	9 min	9 h20 min	99%	60 min	Yes
<881>	<898>	104	100%	6 min	10 h	75%	24 min	Yes
<881>	<979>	104	100%	9 min	8 h	99%	35 min	Yes

Table 5 Statistical analysis of OD clusters in the elder segment, where the mean number of transfers exceeds 0.8

Main origin cluster	Destination cluster	Number of OD trips	OD trips with transfers (%)	Mean transfer time (minutes)	Mean stay time (hours)	Multimodality percentage in OD with transfers	Mean travel time	Health centres or hospitals close to the destination cluster
<191>	<4>	31	100%	11 min	1h40min	100%	18 min	No
<191>	<943>	30	100%	6 min	2h50min	100%	22 min	Yes (hospital)
<220>	<943>	41	80%	5 min	2h30min	0 (only bus) %	32 min	Yes (hospital)
<486>	<595>	38	100%	6 min	8 h	100%	21 min	No
<408>	<695>	34	100%	7 min	2 h	0% (only bus)	9 min	No
<895>	<300>	30	100%	12 min	7h30min	100%	33 min	Yes (health centre)

particularly when transfers are involved, especially multimodal transfers between metro and bus services. This tendency highlights the distinct mobility preferences and priorities of the older adult segment.

6 Conclusions

This study provides two major contributions: (i) an enhanced approach for the OD trip inference, and (ii) a comprehensive analysis of urban mobility patterns for different user segments through the principled instantiation of the proposed approach in Lisbon city's public transport system. Notable insights into passenger behaviours, travel patterns, and transportation system performance are gathered by the advances in data enrichment offered by the proposed methodologies.

The proposed OD inference process receives smart card data from public transportation as input, where boarding and alighting data must be present or estimated. Previous work on alighting stop estimation is advanced by leveraging OpenStreetMap data for distance analysis and (historical) boarding data to refine alighting time estimates in the absence of real-time scheduling GTFS data. We observed that approximately 26% of stops deviate by one or two stops from traditional estimators, as well as significant time differences (1–40 min) on identical travels due to real-world issues like traffic and events. These differences highlight the importance of this step to guide the subsequent origin–destination trip mapping.

The first main contribution of this work involves a novel OD trip inference algorithm based on a semi-supervised learning framework for transfer and activity identification within trip segments. Initially, a baseline sample of user trips is established, where transfers and activities are identified using state-of-the-art principles. Subsequently, an ensemble classifier leverages this sample to categorize remaining inter-trip segments into transfers or activities. The ensemble classification model integrates user profile, card type, and network characteristics such as route, stops, inter-transaction distances, alongside a sigmoid function that adjusts the classifier's probability weights based on the time spent between transactions. This approach minimizes the dependency on transfer time in classification, addressing the challenges posed by the dataset's imbalance and variable class distributions. The approach achieved approximately 98% across all testing metrics: accuracy, precision, recall, and F1 Score. A focused analysis on these results involved examining each type of mode shift within the inter-transaction segment for the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The results underscore the model's efficacy in predicting subway-to-subway and bus-to-subway transitions, with the predominant categories being activity and transfer, respectively. For other transition scenarios where both classes are evident, particularly the transfer class, the model exhibits promising predictive performance, even though a non-negligible number of false negatives (FN) and false positives (FP) remain.

As a second major contribution, our analysis provides a detailed examination of mobility patterns across various user segments, revealing distinct mobility behaviours among youths (children and young adult segment), adults, and seniors. This highlights the necessity of customizing transportation planning for different age demographics. For instance, the findings indicate that seniors are more likely to engage in transfers, especially bus-to-bus. Additionally, seniors tend to utilize the subway less frequently compared to other age groups. The identification of significant destination clusters and their associations with key facilities, such as education centres and healthcare facilities, adds an essential layer of understanding to urban mobility dynamics. Based on these results, key actionable recommendations for public transport operators can be outlined as follows:

- Route optimization to minimize transfer needs: this includes the design and promotion of more inclusive routes that directly connect key locations frequented by the elderly or children and young adult segment, including essential activities and services such as schools and hospitals. This approach reduces the need for transfers, enhancing accessibility;
- Schedule coordination for enhanced transfer experience: our research findings indicate that the elderly segment undertake more transfers during the day. The observed behaviors can be used to guide the coordination between bus and subway schedules, thereby reducing waiting times and facilitating smoother transfers on journeys frequented by the elderly;
- Segment-specific feedback given the possible reluctance among the elderly to use specific services, such as subway transport. Authorities should conduct qualitative interviews or surveys to gather further insights, potentially addressing issues such as accessibility, travel experience, or the affordability and convenience of the fare system.

Finally, limits to the undertaken validation are acknowledged due to the lack of ground-truth samples (e.g., annotated OD samples from volunteer trips). Future research aims to address this by employing additional statistical metrics to evaluate the efficacy of transfer vs. activity prediction task in the absence of ground truth. Moreover, given the diversity of public transport modes in Lisbon (encompassing cycling and train modes in addition to subway and bus), forthcoming studies will explore their significance in shaping origin–destination travel patterns. The findings from this work further motivate complementary research ends, specifically: (i) test hypotheses for the observed behavioural differences between the studied segments, (ii) assess latent factors along trips such as station accessibility, fare structures, and other potential barriers, and (iii) include further demographic and contextual factors to better inquire trip motifs.

Appendix

This section presents key results from prior research [19] for improving the process of estimating alighting points in the same case study. The developed algorithm is briefly outlined, followed by a discussion of the main findings and their implications for improving accuracy in alighting estimations.

The proposed approach combines state-of-the-art principles with frequent pattern mining and density-based clustering techniques for a superior inference of alighting stops for non-commuting passengers.

Procedure: The proposed alighting estimation model operates in three key stages, each with specific principles and parameters to enhance accuracy.

Stage 1—Trip-Chaining Principles: The model begins by applying foundational trip-chaining principles. This stage includes baseline principles gathered from the state-of-the-art review that are able to generalize commuting patterns. Among the subsets of candidate alighting stops, the optimal alighting stops is selected based on proximity to reference boarding. The reference boarding stop is dependent on the following principles:

- *[E1] After alighting, the passenger will walk to the next boarding, whose transfer distance must be less than a given threshold (except for the last trip of the day).*
- *[E2] For the last trip of the day, the passenger will alight close to the first boarding of the day, whose distance must be less than a given threshold*
- *[A1] Two consecutive boarding stops will not occur in the same location.*

Table 6 Estimation rate for each phase of the estimation process

Stage		ER (%)	CER [HQ	Transfer distance (meters)				X Confidence (%)	
				x	Q1	Q2	Q3	Distance based	Frequency based
1. Trip-chaining Algorithm	El	5–4.44	–	148	26.2	80.5	171.3	95.1	–
	E2	19.62	74.06	198	23.4	74.8	240.1	92.1	–
	Bl	2.16	76.22	197	20.5	74.2	212.5	91.4	–
	B2	3.22	79.44	145	22.2	76.9	175.4	94.4	–
2. Clustering	Single trips	5.13	84.57	86	0	21.8	102.2	97.1	–
	Others	5.80	90.37	105	0	31.9	120.1	96.2	–
3.Frequent Pattern		0,17	90.54	–	–	–	–	–	59.6
Not Estimated		9.46	100	–	–	–	–	–	–

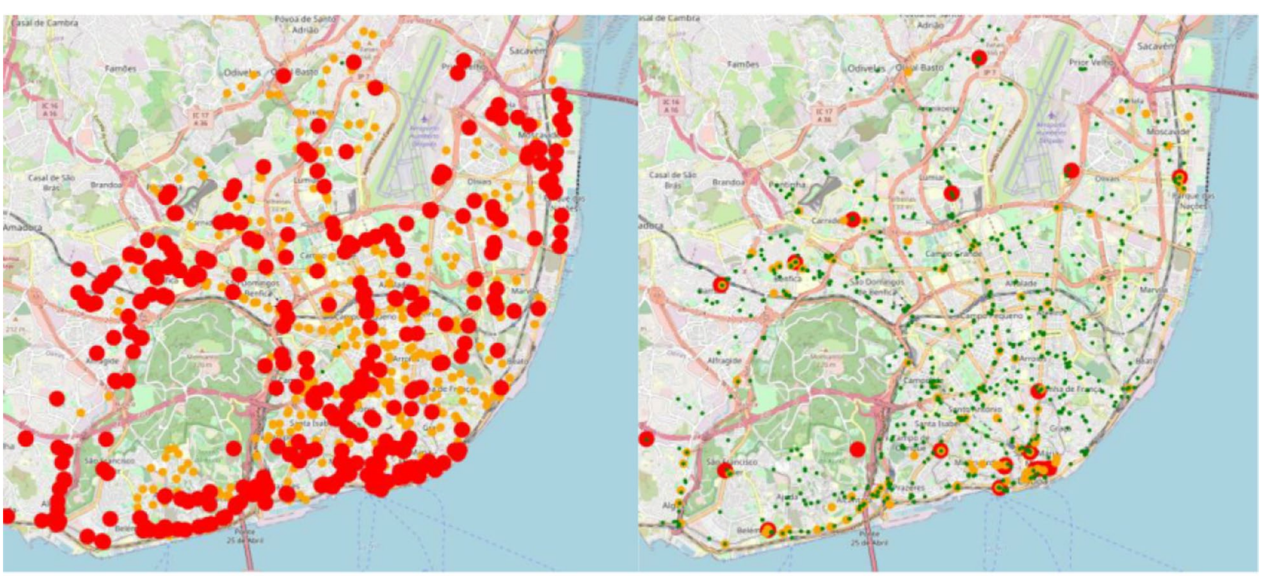
* ER(%) = Estimation Rate; CER(%) = Cumulative Estimation Rate; *x* = Average; Ql to Q3 = Quartiles

- *[A2] For all passengers, a day starts when the network has the lowest activity, for instance, from 4 AM to 3:59 AM of the next day.*

By observing the unique patterns in the mobility data, the work proposes enhanced principles (B1 and B2) to improve estimations for cases where commuting patterns are less clear:

- *[B1] Principle states that an alighting stop for the last trip of the day (target trip) can be estimated by assuming that it is close to the first boarding that occurs in the next days.*
- *[B2] Principle states that a candidate alighting stop can be among the stops upstream from the stop across the street, in the opposite route (applicable only to last stops of the route service)*

Stage 2—Density-Based Clustering: Previous research suggests tracking the sequence of primary clusters for each passenger to help estimating alighting points. By applying density-based clustering, GPS data on passenger locations—along with boarding and alighting coordinates—are grouped to identify movement patterns. This approach assumes that passengers often follow routines, using similar boarding stops, times, and routes. Parameters such as a 600 m clustering distance and a 1000 m maximum transfer distance ensure cluster cohesion and accuracy. The identified cluster serve as potential alighting location.



(a) Results after Stage 1 (b) Results after stages 2 and 3

Fig. 10 Spatial distribution of trips without alighting stops estimated, considering the boarding stop location

Stage 3—Frequent Pattern Mining: In this final stage, frequent patterns in boarding times and locations are analysed for trips with missing alighting data. An alighting stop is inferred if it appears with sufficient frequency and within a 2-h boarding time range from other similar trips. This threshold limits overestimation and improves reliability.

Sensitive analysis: From the results presented in Table 6, the alighting estimation model reaches a total estimation rate of 92.84% through its three stages. In the first stage, which applies principles E1, E2, B1, and B2, the model achieves a 79.44% estimation rate, with E1 and E2 providing the largest contributions. The second stage, using density-based clustering, further boosts the estimation rate by 10.93 percentage points (pp) and is especially effective for single-trip estimations, adding 5.13 pp. In the third stage, which uses patterns from frequent trips with similar boarding times, there's a smaller increase of 0.17 pp.

An analysis of transfer distances in the first two stages reveals values reaching up to 240 m, with confidence levels between 91.4% and 97.1%. The third stage (frequent patterns), however, shows a lower confidence values (59.6%).

Unimodal model validation: To model was further tested in a bus-only (unimodal) context, where it achieved an 83% estimation rate, slightly lower than the 92.84% rate in the multimodal setup. Both models aligned on 73.32 pp of estimated stops at an 83% confidence level. Additionally, 9.5 pp of these matches were estimated in stage 1 of the multimodal model but were achieved through stages 2 and 3 in the unimodal model, indicating that stages 2 and 3 can partially fill gaps in multimodal data.

Spatial Exploratory Analysis: Fig. 10 presents the spatial distribution of trips without estimated alighting stops after stage 1 (left figure) and stages 2–3 (right figure). Red circles represent boarding stops where alighting could not be estimated for over 20% of trips. In the distribution after stage 1, red circles are scattered widely across the map without a clear pattern. However, following stages 2 and 3, the red circles become concentrated mainly within Lisbon's city boundary and historic centre. These high-failure zones correspond to boarding stops where passengers transfer to other public transport operators. Despite the overall positive results, this concentration highlights limitations in alighting estimation when trips involve operators that are not captured in the available data.

Acknowledgements

The authors thank CARRIS, METRO and Câmara Municipal de Lisboa for the provision of data and valuable support. This research was funded by Fundação para a Ciência e Tecnologia (FCT) under PhD Grant number 2022.13483.BD to SC, ADMIRAL (HORIZON-CL5-2022-D6-02-01 under agreement 101104163), project ILU (DSAIPA/DS/0111/2018), INESC-ID plurianual (UIDB/50021/2020) and ATE (02/C05-i01.02/2022 under agreement PC644914747-00000023).

Author contributions

. All authors contributed to the conceptual development of the solution. SC implemented the software and writting (original draft). EA, JB and RH ensured the supervision of the implemented works and analysis. SC, EA, JB and RH contributed to the writing - review and editing of the manuscript. EA - funding. All authors read and approved the final manuscript.

Funding

This article has been published open access with support of the TRA2024 project funded by the European Union.

Data availability

The smart card data that support the findings of this study cannot be shared at this time due to privacy aspects and legal constraints. NDA-mediated protocols with CARRIS and METRO required for data access.

Declarations

Competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests.

Received: 22 May 2024 Accepted: 26 March 2025
Published online: 22 May 2025

References

1. Alsger, A. A., Mesbah, M., Ferreira, L., & Safi, H. (2015). Use of smart card fare data to estimate public transport origin–destination matrix. *Transportation Research Record, 2535*(1), 88–96.
2. Hamedmoghadam, H., Vu, H. L., Jalili, M., Saberi, M., Stone, L., & Hoogendoorn, S. (2021). Automated extraction of origin-destination demand for public transportation from smartcard data with pattern recognition. *Transportation Research Part C: Emerging Technologies, 129*, 103210.
3. Seaborn, C., Attanucci, J., & Wilson, N. H. (2009). Analyzing multimodal public transport journeys in London with smart card fare payment data. *Transportation Research Record, 2121*(1), 55–62.
4. Trépanier, M., Tranchant, N., & Chapleau, R. (2007). Individual trip destination estimation in a transit smart card automated fare collection system. *Journal of Intelligent Transportation Systems, 11*(1), 1–14.
5. Hussain, E., Bhaskar, A., & Chung, E. (2021). Transit OD matrix estimation using smartcard data: Recent developments and future research challenges. *Transportation Research Part C: Emerging Technologies, 125*, 103044.
6. Mohammed, M., & Oke, J. (2023). Origin-destination inference in public transportation systems: A comprehensive review. *International Journal of Transportation Science and Technology, 12*(1), 315–328.
7. Anda, C., Medina, S. A. O., & Fourie, P. (2018). Multi-agent urban transport simulations using OD matrices from mobile phone data. *Procedia Computer Science, 130*, 803–809.
8. Zannat, K. E., & Choudhury, C. F. (2019). Emerging big data sources for public transport planning: A systematic review on current state of art and future research directions. *Journal of the Indian Institute of Science, 99*(4), 601–619.
9. Barry, J. J., Freimer, R., & Slavin, H. (2009). Use of entry-only automatic fare collection data to estimate linked transit trips in New York City. *Transportation Research Record, 2112*(1), 53–61.
10. He, L., & Trépanier, M. (2015). Estimating the destination of unlinked trips in transit smart card fare data. *Transportation Research Record, 2535*(1), 97–104.
11. Nassir, N., Khani, A., Lee, S. G., Noh, H., & Hickman, M. (2011). Transit stop-level origin–destination estimation through use of transit schedule

and automated data collection system. *Transportation Research Record, 2263*(1), 140–150.
12. Nunes, A. A., Dias, T. G., & e Cunha, J. F. (2015). Passenger journey destination estimation from automated fare collection system data using spatial validation. *IEEE Transactions on Intelligent Transportation Systems, 17*(1), 133–142.
13. Assemi, B., Alsger, A., Moghaddam, M., Hickman, M., & Mesbah, M. (2020). Improving alighting stop inference accuracy in the trip chaining method using neural networks. *Public Transport, 12*(1), 89–121.
14. Cheng, Z., Trépanier, M., & Sun, L. (2021). Probabilistic model for destination inference and travel pattern mining from smart card data. *Transportation, 48*(4), 2035–2053.
15. Jung, J., & Sohn, K. (2017). Deep-learning architecture to forecast destinations of bus passengers from entry-only smart-card data. *IET Intelligent Transport Systems, 11*(6), 334–339.
16. Lee, S., Lee, J., Bae, B., Nam, D., & Cheon, S. (2021). Estimating destination of bus trips considering trip type characteristics. *Applied Sciences, 11*(21), 10415.
17. Lei, D., Chen, X., Cheng, L., Zhang, L., Wang, P., & Wang, K. (2021). Minimum entropy rate-improved trip-chain method for origin–destination estimation using smart card data. *Transportation Research Part C: Emerging Technologies, 130*, 103307.
18. Yan, F., Yang, C., & Ukkusuri, S. V. (2019). Alighting stop determination using two-step algorithms in bus transit systems. *Transportmetrica A: Transport Science, 15*(2), 1522–1542.
19. Cerqueira, S., Arsenio, E., Barateiro, J., & Henriques, R. (2024). Moving from classical towards machine learning stances for bus passengers' alighting estimation: A comparison of state-of-the-art approaches in the city of Lisbon. *Transportation Engineering, 16*, 100239.
20. Munizaga, M., Devillaine, F., Navarrete, C., & Silva, D. (2014). Validating travel behavior estimated from smartcard data. *Transportation Research Part C: Emerging Technologies, 44*, 70-79.
21. Alsger, A., Assemi, B., Mesbah, M., & Ferreira, L. (2016) Validating and improving public transport origin–destination estimation algorithm using smart card fare data. *Transportation Research Part C: Emerging Technologies*, 68490–68506. https://doi.org/10.1016/j.trc.2016.05.004.
22. Nunes, A. A., Dias, T. G., & e Cunha, J. F. (2016). Passenger journey destination estimation from automated fare collection system data using spatial validation. *IEEE Transactions on Intelligent Transportation Systems, 17*(1), 133–142. https://doi.org/10.1109/TITS.2015.2464335.
23. Faroqi, H., & Mesbah, M. (2021). Inferring trip purpose by clustering sequences of smart card records. *Transportation Research Part C: Emerging Technologies, 127*, 103131.
24. Gordon, J. B., Koutsopoulos, H. N., Wilson, N. H., & Attanucci, J. P. (2013). Automated inference of linked transit journeys in London using fare-transaction and vehicle location data. *Transportation Research Record, 2343*(1), 17–24.
25. Jafari Kang, M., Ataeian, S., & Amiripour, S. M. (2021). A procedure for public transit OD matrix generation using smart card transaction data. *Public Transport, 13*, 81–100.
26. Kumar, P., Khani, A., & He, Q. (2018). A robust method for estimating transit passenger trajectories using automated data. *Transportation Research Part C: Emerging Technologies, 95*, 731–747.
27. Munizaga, M. A., & Palma, C. (2012). Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile. *Transportation Research Part C: Emerging Technologies, 24*, 9–18.
28. Yap, M. D., Cats, O., van Oort, N., & Hoogendoorn, S. P. (2017). A robust transfer inference algorithm for public transport journeys during disruptions. *Transportation Research Procedia, 27*, 1042–1049.
29. Zhao, D., Wang, W., Li, C., Ji, Y., Hu, X., & Wang, W. (2019). Recognizing metro-bus transfers from smart card data. *Transportation Planning and Technology, 42*(1), 70–83.
30. Nassir, N., Hickman, M., & Ma, Z. L. (2015). Activity detection and transfer identification for public transit fare card data. *Transportation, 42*, 683–705.
31. Bokolo, A. J. (2023). Inclusive and safe mobility needs of senior citizens: Implications for age-friendly cities and communities. *Urban Science, 7*(4), 103.
32. Fatima, K., Moridpour, S., De Gruyter, C., & Saghapour, T. (2020). Elderly sustainable mobility: Scientific paper review. *Sustainability, 12*(18), 7319.
33. World Health Organization. (2015). Measuring the age-friendliness of cities: A guide to using core indicators.
34. Aparicio, J. T., Arsenio, E., Santos, F. C., & Henriques, R. (2024). Walkability defined neighbourhoods for sustainable cities. *Cities, 149*, 104944.
35. Arsenio, E., Martens, K., & Di Ciommo, F. (2016). Sustainable urban mobility plans: Bridging climate change and equity targets? *Research in Transportation Economics, 55*, 30–39.

Publisher's Note