Identification and Characterisation of Passenger Archetypes based on annual Long-distance Travel Patterns

J. Bueno-González, M. de Boissieu, O.G. Cantú-Ros, R. Herranz
Nommon Solutions and Technologies
Madrid, Spain

***Abstract*— The European transport policy envisions a multimodal transport system where different networks and services are planned and managed in a coordinated manner to maximize the efficiency, predictability, environmental sustainability, and resilience of the door-to-door passenger journey. To achieve this goal, transport planners need an in-depth understanding of the behaviors, preferences and needs of the different types of travelers within Europe. This paper presents a methodology for the identification and characterization of long-distance passenger archetypes based on the application of unsupervised learning algorithms to a set of travel behavior indicators extracted from anonymized mobile network data. The proposed methodology is demonstrated and evaluated through its application to long-distance travel in Spain.**

Keywords- mobile network data; passenger archetypes; clustering; data fusion, transport planning.

I. Introduction

A. Background and motivation

The European Commission's Sustainable and Smart Mobility Strategy [1] has defined a multitude of goals and respective flagships that pave the way towards zero-emission, resilient and inclusive mobility, creating seamless and efficient connectivity and establishing the European Union as a connectivity hub. In line with this objective, the long-term vision for the European aviation sector outlined in the report 'Flightpath 2050 - Europe's Vision for Aviation' [2] envisages a passenger-centric air transport system thoroughly integrated with other modes to ensure a seamless passenger experience. The need for better integration between air transport and other modes is also acknowledged by SESAR Strategic Research and Innovation Agenda [3], one of whose key research areas, 'Multimodality and passenger experience', focuses on coordinated planning and collaborative decision-making solutions that improve the integration of air transport in the intermodal transport system.

To design an efficient multimodal transport system and improve passenger experience, it is crucial to understand the underlying demand and the passengers' expectations. Long-distance travelers encompass a wide variety of profiles, from business professionals prioritizing efficiency and reliability to families seeking comfort and affordability. Each of these groups has distinct expectations and varying sensitivities to different trip characteristics, such as the price, the CO_2 emissions, the number of transfers, and the total travelling time. These different sensitivities will determine their mode and path choice when booking a trip. Moreover, beyond sociodemographic characteristics, mode choice and route election are also influenced by the traveler's long-term travel patterns [4]. This means that groups of travelers sharing long-term travel behaviors tend to show similar attitudes towards modal choice. By identifying and characterizing these groups, transport planners can create more targeted and effective services that better meet the needs of each group, ultimately enhancing the overall passenger experience. The concept of 'passenger archetype' — or 'passenger persona' — provides a way to model, summarize and communicate the profile and behavior of distinct passenger groups [5].

The analysis of passenger behavior has traditionally relied on surveys. Surveys provide a detailed characterization of the respondent, but are expensive and time-consuming, which limits the sample size and the frequency of update. In recent years, different studies have explored how the digital traces generated by personal mobile devices can be exploited to study passenger behavior. Mobile network data (MND) are particularly suitable for this purpose, thanks to the possibility of working with large, well-distributed population samples with high temporal and spatial resolution ([6],[7]). The longitudinal nature of MND enables the continuous monitoring of long-distance travel demand: passenger travel diaries can be reconstructed and the location of passengers' overnights throughout a long period of time can be identified. These diaries can then be extrapolated to the whole population using census data and other sociodemographic statistics, in a similar process to the sample expansion of a traditional travel survey ([6],[7]). In addition, MND provide sociodemographic characteristics such as age and gender. On the other hand, MND do not directly provide certain key features, such as the trip purpose (e.g., business vs leisure) or the passenger income level, which can have a strong influence on passenger behavior; to overcome this limitation, different approaches based on a combination of data fusion and machine learning techniques have recently been developed ([8],[9]).

In this paper, we investigate how to use unsupervised machine learning methods, in particular clustering techniques, on different mobility indicators extracted from MND to identify and characterize passenger archetypes based on their annual long-distance travel patterns.

SESAR Innovation Days 2024
12 - 15 November 2024. Rome
EUROPEAN PARTNERSHIP

B. Previous work

While models have been developed to analyze and forecast interurban travel demand ([9],[11]), most of them do not consider the specificities of different groups of passengers. Regarding the analysis of passenger personas, most existing studies are based on the passengers' demographics rather than on their travel patterns. The work developed in [12] proposes a typology of travelers based on people's annual intercity travel patterns, using data from the 2013 Longitudinal Survey of Overnight Travel; this survey provides a year of data, addressing the nonroutine nature of long-distance travel, but is limited in sample size (1,220 respondents). Similarly, the classification of leisure traveler types conducted in [13] is based on applying latent class analysis to two surveys. The use of MND instead of survey data, as proposed in the present study, ensures a wider and better-distributed sample and eliminates the potential bias of survey respondents when describing their own behavior.

C. Objectives of the study

This paper presents a methodology to extract passenger archetypes based on the long-distance travel mobility patterns observed through MND. This methodology comprehends: (i) the application of unsupervised machine learning to identify clusters of passengers according to their travel behavior, (ii) the characterization of these archetypes by analyzing the common sociodemographic factors within each cluster, and (iii) the combination of these two factors to formulate passenger archetypes. The proposed approach is applied to the case study of long-distance travel of the residents in Spain using a whole year of MND.

The rest of the paper is structured as follows. Section II describes the proposed methodology. Section III shows the results of the application of the methodology to the Spain case study. Section IV presents some additional analysis to validate the results obtained. Section V discusses the main conclusions of the study and suggests directions for future research.

II. DATA AND METHODOLOGY

A. Data

The methodology presented relies on anonymized MND as the main source for identifying long-distance trips. MND consists of records of interactions between mobile devices and the network of antennas managed by a mobile network operator (MNO), capturing data from both local users (the customers of the MNO) and foreign visitors (roaming-in users). Each record includes a device ID, a timestamp, and the identifier of the antenna in communication with the device. In addition to these records, the dataset contains information about the location of the antennas and basic sociodemographic information for each anonymous user, such as age and gender.

The data is geolocated at the cell level, i.e., the position is known at the level of the coverage area of each antenna. Therefore, the geolocation accuracy depends on the density of antennas, ranging from dozens or hundreds of meters in urban areas to a few kilometers in rural regions. The temporal resolution depends on the type of interaction: active events (such as calls or data sessions) generally occur at intervals of 20-30 minutes, while passive events (network scans) produce more frequent records, typically every 5-10 minutes.

The sample size is usually substantial, with access to MND from a single MNO often providing data from 15-30% of the mobile devices in the target country. Since mobile phones are widely used across all population segments, except for children, the sample provides good representativeness of the adult population.

B. Methodology

The proposed methodology consists of four main steps, which are described in the following subsections.

1) Generation of long-distance travel diary

Based on the methodology developed in previous studies ([6],[7]) for the identification of long-distance trips from MND, a longitudinal analysis has been conducted to extract the annual travel plan of each individual in the sample. To do this, a nighttime interval is defined, during which the anonymized mobile phone records are analyzed, along with the minimum stay time within the nighttime interval to determine that the user has spent the night within the coverage area of the antenna the user is connected to. The result of this analysis is the overnight diary, which consists of the sequence of locations where the individual spends the night throughout the year, as shown in Figure 1. This data provides insights into the user's annual travel behavior (i.e., where they travel, how often they travel and how long they travel for), along with sociodemographic information such as age, gender, income level, and home location.

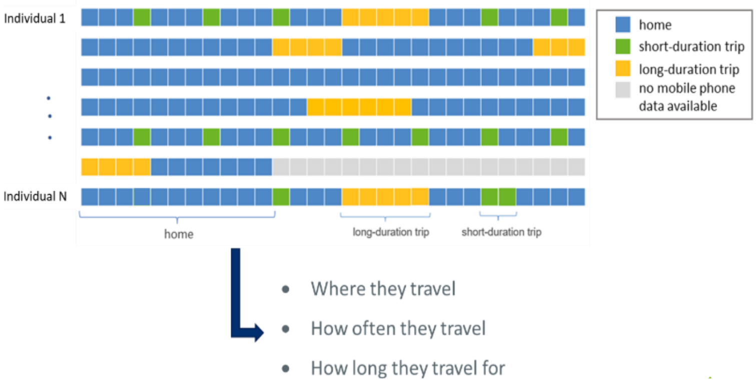

Figure 1- Overnights diary for each individual

2) Feature extraction

This step involves identifying features that can be derived from the data and are representative of specific long-distance travel behaviors. A crucial preliminary task is to clearly define what constitutes long-distance travel. Following the usual convention [4], a one-way distance threshold of 100 km is set. All the trips falling below this threshold are filtered out from the overnight diary.

The selected features aim to capture different travel patterns in the population along three dimensions: (i) volume of long-distance trips; (ii) average behavior; and (iii) longitudinal variation. The following list of features is considered:

- Number of trips performed,
- Number of overnight stays outside home,
- Average one-way distance travelled,
- Average duration of the trips,
- Ratio of most common destination trips to total number of trips,
- Ratio of weekday trips to total trips, where a weekday trip excludes any overnight stay on Friday, Saturday, or Sunday,
- Ratio of overnight stays abroad to total overnight stays, which capture the frequency of international destinations,
- Coefficient of variation of distance,
- Coefficient of variation of duration,
- Destination entropy, which assesses the degree of dispersion or concentration of trips across different destinations,
- Ratio of overnight stays during holidays to total overnights stays, where holidays are defined as the periods covering July, August, Christmas, and Easter,
- Coefficient of variation of overnight stays per month,
- Overnight stays per month, resulting in twelve variables, one for each month.

Once the relevant features are calculated, a correlation analysis is conducted to identify any highly correlated features. Features with a Pearson correlation coefficient greater than 0.9 are considered for removal.

After feature selection, the next step is to standardize the data using min-max scaling. This technique rescales each feature to a range between 0 and 1, ensuring that all features contribute equally to the clustering, preventing variables with larger ranges (e.g., average one-way distance travelled) from disproportionately influencing the clustering outcome.

To further reduce the dimensionality of the dataset, Principal Component Analysis (PCA) is employed. PCA identifies the most informative components that capture the maximum variance in the data. In our case, we retained components that explained 98% of the total variance. The choice to retain 98% of the variance in the PCA was driven by a balance between maximizing data representativeness and maintaining computational efficiency, i.e., achieving sufficient reduction in dimensionality while preserving the most meaningful information in the dataset.

3) K-means clustering

The algorithm selected for clustering is k-means, an unsupervised machine learning algorithm used to partition a dataset into K distinct clusters. K-means works by iteratively assigning each individual to one of the K clusters based on the similarity of their travel patterns, measured using Euclidean distance. The algorithm begins by allocating K cluster centroids randomly. Each individual is then assigned to the nearest centroid, determined by the lowest distance between its feature values and the centroids' values, forming clusters. The centroids are then recalculated as the mean of all the individuals assigned to each cluster. This process is repeated until the centroids stabilize, and no further significant changes in cluster assignments occur.

The k-means algorithm was chosen primarily due to its computational efficiency when dealing with large datasets. Since the data was obtained from MND, we had access to a vast sample of the population, which made scalability a crucial factor. Other algorithms, such as hierarchical clustering, encountered memory issues when processing such large datasets, whereas k-means is well-suited for handling large-scale data with lower memory requirements and faster convergence.

A sequential k-means clustering was employed to handle the large dataset effectively. This approach involves two distinct stages. First, we separate individuals with zero trips during the studied period from the rest of the dataset. This step is necessary because the non-travelers do not provide meaningful information to the clustering of travel patterns and could skew the results if included. Then, the travelers are portioned into K clusters based on their travel patterns.

A key hyperparameter of the k-means algorithm is the number of clusters, K. Choosing the right K is crucial for obtaining meaningful and actionable clusters. The within-cluster sum of squares (WCSS) and the silhouette score are calculated to determine the optimal number of clusters. WCSS measures the variance within each cluster. As K increases, WCSS generally decreases because the clusters become smaller and more specific. The goal is to find the "elbow" point in the plot, where the rate of decrease sharply slows down. The silhouette score assesses the quality of clustering by evaluating how well each point is clustered within its own cluster compared to other clusters. The score ranges from -1 to 1, where a value close to 1 indicates that the point is well-clustered and distinctly separate from other clusters.

4) Characterization of clusters and prototyping of passenger archetypes

Once the travel patterns are identified through clustering, we characterize these patterns by analyzing various sociodemographic attributes within each cluster. The sociodemographic characteristics examined include age, gender, home municipality, and income level. Age and gender are obtained from the customer profiling data provided by the MNO; home municipality is inferred from the MND; income level is obtained by combining the place of residence inferred from mobile phone data with official statistics on average income level per postal code published by the Spanish National Statistics Office. This analysis helps understand the demographic profiles and geographic distribution of each cluster.

Finally, by examining both long-distance travel mobility patterns and sociodemographic characteristics within each cluster, we extract and define prototype passenger archetypes. This comprehensive examination allows us to identify distinct passenger profiles that represent the key travel behaviors and demographic traits of different segments within the dataset.

III. RESULTS

The data used for this analysis comes from MND provided by a major mobile network operator in Spain, which holds a market share of around 25%. The study covers the period from March 2019 to February 2020, inclusive. A random sample of 259,000 individuals was selected from the data, representing more than 0.5% of the total Spanish population.

We apply the methodology outlined in Section II. Initially, we finalize the set of features used for clustering. From the list presented in Section II, we exclude two features: the ratio of the most common destination trips to the total number of trips, and the number of overnight stays per month. The first feature aimed to identify individuals with a preferred destination (e.g., those frequently traveling to a second residence), but it proved to be insignificant in the clustering process. The second feature, which included 12 variables to capture longitudinal variation throughout the year, introduced unnecessary complexity and did not contribute effectively to the clustering. Therefore, the final set of features used for clustering is: number of trips performed, number of overnight stays outside home, average one-way distance travelled, average duration of the trips, ratio of weekday trips to total trips, ratio of overnight stays abroad to total overnight stays, coefficient of variation of distance, coefficient of variation of duration, destination entropy and ratio of overnight stays during holidays to total overnight stays.

To determine the optimal number K of clusters, the k-means algorithm was executed with various values of K, and both the sum of squared distances and the silhouette scores were calculated. These metrics are illustrated in Figure 2 and Figure 3, respectively. While Figure 2 does not clearly indicate an elbow point, Figure 3 suggests that six clusters (K=6) result in the best clustering quality. Consequently, we have chosen 6 as the optimal number of clusters for this study.

The travel patterns of the six clusters identified through the clustering analysis, along with the non-traveler group, are detailed in Table 1. For each feature used in the clustering process, the average value within each cluster is provided.

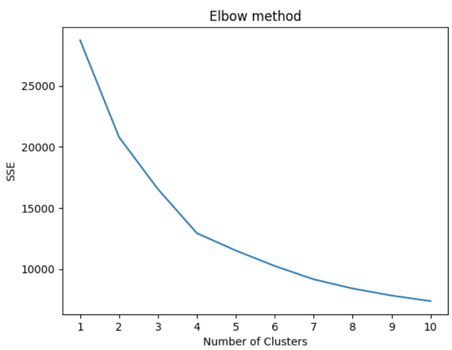

Figure 2: Sum of Squared Errors in function of the number of clusters

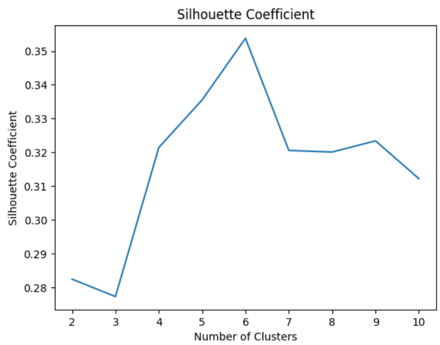

Figure 3: Silhouette coefficient in function of the number of clusters

To further characterize the clusters, we analyzed the sociodemographic attributes of each group. This analysis is illustrated in Figures 4 to 8. Figure 4 shows the age distribution across four categories: 0-24, 25-44, 45-64, and 65+. Gender distribution is presented in Figure 5. Income and home municipality size are displayed in Figures 6 and 7, respectively. Both income and municipality size are divided into five categories, based on quintiles observed across the entire dataset. A horizontal line in each figure represents the overall average for the total dataset, providing a reference point for comparison. Additionally, Figure 8 illustrates the geographical distribution of each cluster, where red regions indicate a higher presence of a given cluster compared to the average, and blue regions indicate a lower presence. Finally, Figure 9 provides a boxplot for each cluster, showing the distribution of overnight stays per month. This visualization offers a more detailed, longitudinal analysis of travel patterns over time.

With this information, seven distinct passenger archetypes have been identified:

- non-traveler: this archetype, representing 30% of the population, does not engage in any long-distance travel throughout the year. It is characterized by an older age, with a slight predominance of females and individuals with limited income. They tend to live in smaller cities;
- long-distance actives: the second-largest archetype, representing 25% of the sample, is characterized by a higher frequency of travel activity compared to all other clusters. The trips undertaken by these individuals cover various destinations and distances, and are distributed throughout the year, with a slightly higher concentration during the summer. They are young, high-income individuals residing in large cities;
- sporadic international traveler: this group occasionally embarks on medium-duration trips, very often preferring international destinations. Their travel pattern does not follow any specific seasonality. They are middle-aged and do not align with any particular category of gender, income, or home municipality size;
- international urbanite: they explore international destinations during holiday periods (summer, Christmas, and Easter). Their trips are long, averaging more than 7 days in duration. This group includes middle-aged individuals with a very high-income, who live in big cities, especially in large metropolitan areas such as Barcelona and Madrid;

- sporadic long-haul traveler: this group rarely engages in long-distance travel, typically taking fewer than two trips per year. However, when they do travel, they cover long distances, often choosing off-peak periods such as September and June, which may be related to their low income. They are not defined by any gender or age group, but they are more prevalent in rural and coastal regions;
- domestic summer traveler: this passenger profile does not usually travel, only twice a year, but the duration of their trips is the longest among all the archetypes. They prefer domestic travel and almost exclusively travel in July or August. Demographically, they correspond to individuals with average income level and home municipality size, but they are concentrated in the range of 45 to 64 years old. Representing 18% of the population, this cluster closely mirrors the typical Spanish family living in the central, often hot, regions of the country and vacationing in coastal areas during the summer for one or two weeks;
- occasional weekday traveler: this group tends to travel during the workweek, with medium frequency, and prefers shorter-duration trips. Their travel is spread throughout the year, with higher activity in autumn and spring. They are low-income individuals who live in medium or small cities. This archetype is most similar to business travelers.

IV. VALIDATION

To validate the results, two additional analyses have been conducted. The first one evaluates the impact of a different number of clusters, while the second one assesses the impact of the sample size.

The choice of K=6 clusters, despite the absence of a clear elbow point, was based on achieving a balance between interpretability and practical application. We considered alternative cluster evaluation metrics, such as silhouette scores, and found that while these scores did not indicate a definitive optimal point, K=6 consistently offered an effective balance between within-cluster cohesion and separation from other clusters. This configuration allowed us to avoid excessive overlap or fragmentation, preserving the practical value of the clusters for actionable insights. To assess the robustness of the results and the impact of this hyperparameter, we repeated the analysis with K=5. The resulting clusters were almost identical, with the exception of the sporadic international traveler and international urbanite archetypes, which merged into a single cluster. This suggests that the remaining archetypes are well-separated, but the two international travel archetypes share more similarities. We chose to retain both clusters because, although they both favor international destinations and are key targets for air travel, there are significant differences between them. International urbanites tend to live in large metropolitan areas, have high incomes, and primarily travel during the summer, whereas sporadic international travelers are more geographically dispersed, often near the coast, with medium incomes and no specific seasonal travel pattern. These distinctions are crucial for designing policies and developing transport networks.

To assess the representativeness of the results and the impact of sample size, we repeated the analysis with smaller samples of 100,000, 50,000, and 10,000 individuals. The results showed that with a sample size of 10,000, the clusters differed significantly from those presented here. However, with the 100,000 and 50,000 samples, the clusters were very similar to the original results, confirming that the chosen sample size of 259,000 was appropriate, balancing representativeness and computational efficiency.

TABLE 1: CLUSTER CHARACTERISTICS

	Non-traveler	Long-distance actives	Sporadic international traveler	International urbanite	Sporadic long-haul traveler	Domestic summer traveler	Occasional weekday traveler
Number of trips	0	9.54	3.35	3.54	1.85	2.28	4.25
Number of overnights	0	38.65	13.80	21.09	7.25	16.23	10.15
Average one-way distance [km]	0	328.66	199.33	218.99	400.01	362.04	332.31
Average duration	0	4.37	4.72	7.14	3.94	7.40	2.40
Ratio weekday trips to total trips	0	0.08	0.13	0.04	0.01	0.02	0.67
Ratio overnight stays abroad to total overnight stays	0	0.11	0.83	0.80	0.01	0.02	0.05
Coefficient of variation of distance	0	0.40	0.09	0.11	0.13	0.16	0.21
Coefficient of variation of duration	0	0.69	0.38	0.52	0.20	0.40	0.31
Destination entropy	0	1.30	0.41	0.49	0.32	0.39	0.67
Ratio overnight stays during holidays to total overnight stays	0	0.50	0.09	0.82	0.03	0.91	0.09
Coefficient of variation of overnights stays per month	0	1.46	2.45	2.42	2.84	2.78	2.33
Size [%]	30	25	6	6	11	18	4

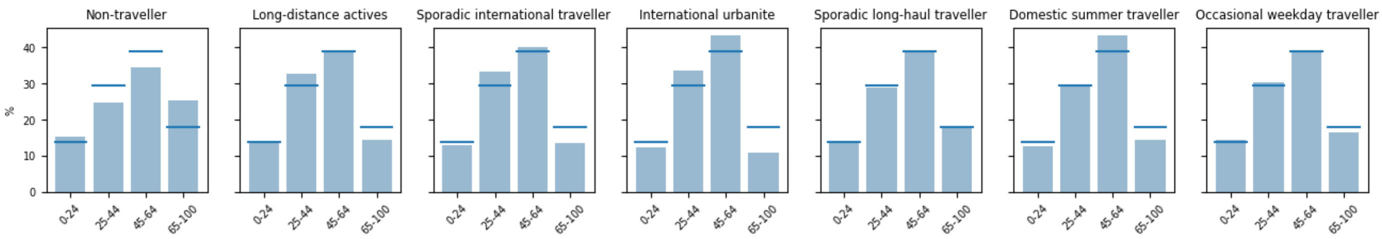

Figure 4: Age distribution within the clusters

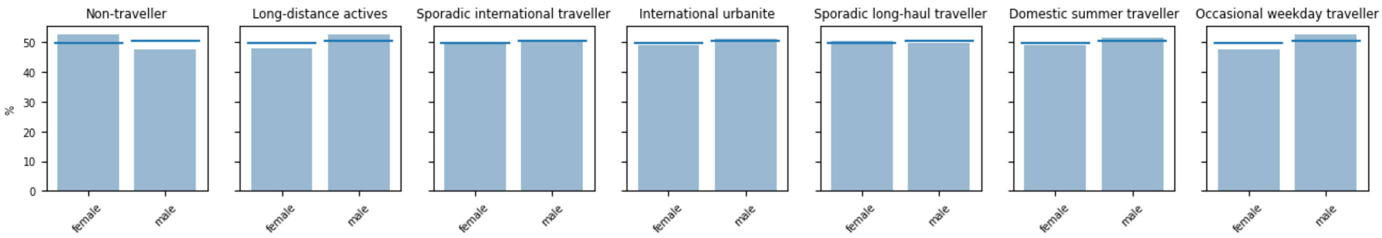

Figure 5: Gender distribution within the clusters

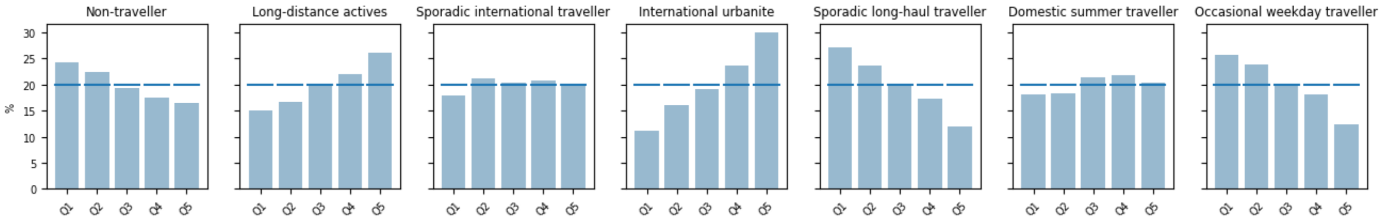

Figure 6: Income distribution within the clusters

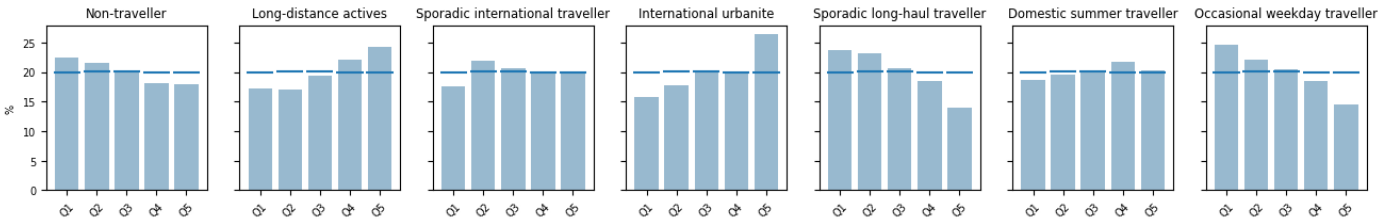

Figure 7: Home municipality size distribution within the clusters

Figure 8: Geographical distribution within the clusters

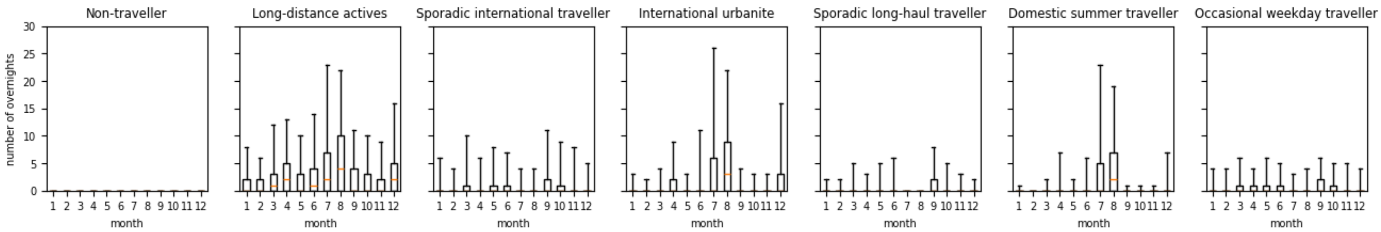

Figure 9: Overnights per month distribution within the cluste

V. Conclusions and Future Research

This paper shows how the longitudinal nature of new big data sources, such as MND, open a valuable opportunity for the continuous and detailed monitoring of door-to-door journeys, enabling the early detection of emerging patterns and trends and providing airlines, airports, and ground transport operators with a deeper understanding of traveler behavior. In particular, the paper contributes to the advancement in the study of long-distance travel by providing a methodology for deriving long-distance passenger archetypes from the analysis of travel patterns extracted from MND. The study offers insights into the preferences and needs of different types of passengers – including when and where they travel – that can help optimize multimodal transport systems in order to better align transportation supply with travel demand and enhance passenger experience.

Future work will encompass several research lines:

- Future work could explore more advanced data preprocessing, dimensionality reduction, and clustering techniques to enhance the robustness of passenger archetype identification. Specifically, applying transformations like power or Box-Cox could improve data distribution for clustering, while comparing dimensionality reduction methods such as t-SNE or UMAP with PCA might improve cluster separability. Additionally, experimenting with alternative clustering algorithms, like DBSCAN, could address the limitations of K-means, especially regarding cluster shape assumptions and sensitivity to K-selection.
- Extending the analysis to incorporate additional features, such as the mode(s) of transport used for long-distance journeys, will enrich our understanding of how different factors, such as travel time, cost, CO_2 emissions and the number of transfers, influence passenger choices. Furthermore, the analysis of how and from where these passengers access airports may reveal opportunities for enhanced coordination of multimodal networks, transforming airports into effective intermodal hubs.
- The proposed methodology will serve as a basis for developing passenger archetypes at worldwide level. Expanding the analysis to other countries will enable a broader examination of long-distance travel patterns, offering a more comprehensive picture of passenger behavior across various regions. Additionally, this will allow us to compare the outcomes of our methodology with the passenger personas defined by other studies that use more conventional approaches based on demographics and/or questionnaires, such as [12] and [13].
- The defined passenger archetypes will be used to assess traveler sensitivities to different multimodal travel options. This will facilitate the development of a multimodal modelling and evaluation framework that supports the design, development, and assessment of different multimodal solutions, such as scheduling and disruption management solutions.

In conclusion, by providing large-scale insights into passenger behavior, the proposed methodology opens up new opportunities for the development of a more efficient and integrated multimodal transport system.

Acknowledgment

The project leading to these results has received funding from the SESAR Joint Undertaking under Grant Agreement No 101114815 under European Union's Horizon Europe research and innovation program. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the SESAR Joint Undertaking.

References

[1] European Commission (2020). Sustainable and Smart Mobility Strategy. Putting European transport on track for the future.

[2] European Commission (2011). Flightpath 2050. Europe's Vision for Aviation. Report of the High-Level Group on Aviation Research.

[3] SESAR JU (2020). Strategic Research and Innovation Agenda – Digital European Sky.

[4] Axhausen, K., Madre, J.-L., Polak, J., & Toint, P. (Eds.) (2002). Capturing Long-Distance Travel. Research Science Press.

[5] Airport Council International (2017). ACI Passenger Personas: A new approach to passenger profiling

[6] Garcıa, P., Ramasco, J. J., Andrienko, G., Adler, N., Ciruelos, C., & Herranz, R. (2016). Big data analytics for a passenger-centric ATM system: A case study of door-to-door intermodal passenger journey inferred from mobile phone data. Proceedings of the 6th SESAR Innovation Days.

[7] García Albertos, P., García-Cantú Ros, O., Ciruelos, C., & Herranz, R. (2017). Understanding Door-to-Door Travel Times from Opportunistically Collected Mobile Phone Records: A Case Study of Spanish Airports. Proceedings of the 7th SESAR Innovation Days.

[8] Blasco-Puyuelo, J., Blasco, C., Jordá-Muñoz, R., Burrieza-Galán, J., Cantú-Ros, O. G., & Mocholí, D. (2022). Data fusion for the analysis of air travel behavior: application to Palma de Mallorca Airport. In Proceedings of the 12th SESAR Innovation Days.

[9] Burrieza-Galán, J., Martín, I., Nicolai, F., Monroig, J., Torres, J., & Picornell, M. (2023). Identification of long-distance tour purpose through ML-based data fusion techniques for combining mobile network and survey data. Transportation Research Procedia, 71, 93

[10] Lu, M., Zhu, H., Luo, X., & Lei, L. (2014). Intercity travel demand analysis model. Advances in Mechanical Engineering, 6, 108180.

[11] Yao, E., & Morikawa, T. (2005). A study of an integrated intercity travel demand model. Transportation Research Part A: Policy and Practice, 39(4), 367-381.

[12] Bacon, B., & LaMondia, J. J. (2016). Typology of travelers based on their annual intercity travel patterns developed from 2013 longitudinal survey of overnight travel. Transportation Research Record, 2600(1), 12-19.

[13] Miriam Magdolen, Sascha von Behren, Bastian Chlond, Peter Vortisch, Long-distance travel in tension with everyday mobility of urbanites – A classification of leisure travellers, Travel Behaviour and Society, Volume 26, 2022, Pages 290-300, ISSN 2214-367X, https://doi.org/10.1016/j.tbs.2021.10.010.