

## Review Article

---

# Analysis of human mobility behavior in Smart Cities and Smart Environments: A systematic literature review and taxonomy

Carlos Eduardo Liedtke Borges<sup>\*</sup>, Jorge Arthur Schneider Aranda and Jorge Luis Victória Barbosa  
*Applied Computing Graduate Program (PPGCA), University of Vale do Rio dos Sinos (UNISINOS), São Leopoldo, Brazil*

Received 30 May 2024

Accepted 6 August 2024

**Abstract.** The analysis of human mobility behavior through computational techniques finds applications in various domains and provides valuable insights for urban planning, transportation services, and a deeper understanding of human interactions in Smart Cities and Smart Environments. In this scenario, this study presents a Systematic Literature Review (SLR) with the following main question: How are computational techniques being used to analyse human mobility behavior in Smart Cities and Smart Environments? A total of 5989 articles were initially found and filtered, resulting in 56 articles reviewed. As the main contributions, this study provides responses to 19 research questions. A list of the challenges and the computational techniques identified is provided. The algorithms, machine learning techniques and data-sources used by the reviewed studies are also presented and organized through taxonomies. A comprehensive discussion of the identified techniques is conducted, finishing with a compilation of challenges, open issues and research opportunities. To the best of our knowledge, this is the first study that reviewed human mobility behavior covering a wide range of scenarios, including urban mobility, public transport, points and regions of interest, ridesharing, bike-sharing, traffic analysis, driving behavior, electric vehicle charging stations planning, mobility on demand, crowd analysis and others.

Keywords: Smart City, Smart Environment, behavior, mobility, crowd analysis

## 1. Introduction

In the contemporary era, Information and Communication Technologies (ICT) plays a pivotal role in addressing diverse human challenges across sectors such as health, energy, transport and others. In recent decades, the advent of Smart Environments has become increasingly evident, a consequence of advancements in Ubiquitous Computing

---

<sup>\*</sup>Corresponding author: Carlos Eduardo Liedtke Borges, Applied Computing Graduate Program (PPGCA), University of Vale do Rio dos Sinos (UNISINOS), Avenida Unisinos 950, São Leopoldo, Rio Grande do Sul, Brazil. E-mail: [carlosel@edu.unisinos.br](mailto:carlosel@edu.unisinos.br).

through Internet of Things (IoT), focusing on different application areas such as healthcare [6,21,60], monitoring [12], assisting [19,49], and executing everyday tasks. In various types of Smart Environments (e.g., stadiums, convention centers, shopping malls), the analysis and monitoring of human mobility behavior offer a wide range of applications. The aforementioned applications encompass statistical monitoring and analysis of the stores that are visited most often, monitoring air quality and temperature to automatically activate corresponding devices, recommending indoor routes and visits to locations based on past user data, and optimizing energy consumption through the regulation of lighting and air conditioning in unoccupied areas [80]. In these examples, the historical movement data of individuals within the environments serves as a valuable foundation for conducting behavior analysis.

Smart Cities are also Smart Environments in a large scale, as a result of the widespread use of technology to improve the well-being of citizens [10,72]. The establishment of Smart Cities presents multiple challenges, offering extensive research opportunities with high socio-economic impact areas, such as public transport and urban mobility, security, public inspection, and the management of public and private services [3]. The utilization of smart devices and IoT for monitoring and delivering these services results in the generation of substantial datasets. Examples include the monitoring of smartphones to identify population density in certain regions [25], the use of public transport ticketing data from smart cards for mobility patterns analysis [11], use of Telecommunications BigData also for analysis of mobility behavior, and analysis of security videos to identify anomalous human behavior (e.g. violence, robberies) [10,55].

Mobility issues are part of the main challenges in big cities. According to the Brazilian Institute of Geography and Statistics (IBGE),<sup>1</sup> the vehicle fleet in Brazil increased from 45 million in 2006 to 115 million in 2022, indicating an average annual growth of 4.37 million. This is a concern not only related to Brazil, as reported in the Statistical Yearbook of China,<sup>2</sup> the average annual growth of small passenger cars between 2014 and 2019 was 19.96 million, culminating in a fleet of 220 million by the end of 2019 [45].

The growth of the active fleet, especially of small private vehicles, has a wide impact on traffic and mobility planning in cities, such as plan alternatives to traffic jams, need to widen streets to improve traffic, parking with restricted space in large business and economical centers, public transport ineffective, not covering people travel needs or economically impracticable.

Effective planning of transport services and the implementation of Mobility-on-Demand (MoD) strategies demand a thorough comprehension of travel behavior patterns and their evolution over time. This understanding should be coupled with insights into city functional regions and land-use dynamics. As an example of behavior changes, reports from the Center for Studies in Regulation and Infrastructure at FGV [20,58] highlighted the profound impact of the COVID-19 pandemic on the public transport sector, resulting in a significant decline of 70% to 85% in passenger numbers in certain capitals. The rise of remote work practices, including the adoption of hybrid work models (combining office and remote work), has further contributed to a profound shift in the usage patterns of public transport. Also related, the absence of effective transport routes and schedules in regions experiencing significant demand, coupled with a dearth of information regarding public transport options (especially buses), makes individuals to seek alternative modes of transportation, potentially leading to the acquisition of personal vehicles.

The movement of individuals and vehicles, when recorded, is usually stored through logs containing georeferenced coordinates with timestamps, generated by equipment supporting the Global Positioning System (GPS) service. Databases containing these historical records constitute a rich terrain for conducting behavior analyses, searching for mobility patterns through computational models. The application of clustering [11,36,38] and Dynamic Time Warping (DTW) analysis [36,43,68] on this historical data are some examples of context and behavior pattern exploration.

The use of Machine Learning (ML) techniques on different types of datasets (e.g. images, videos, travel history) to identify objects (vehicles, pedestrians, traffic signs etc) represents an important strategy for multiple areas, such as assisted and autonomous driving [28], plate recognition [71] and data selection and filtering [84].

---

<sup>1</sup><https://cidades.ibge.gov.br> – *Cidades IBGE* (Brazilian Institute of Geography and Statistics) website (Accessed in April/23).

<sup>2</sup><https://www.mps.gov.cn/n2254314/n6409334/c6852472/content.html> – Link of Statistical Yearbook of China (Accessed in April/23).

In this scenario, we realized a Systematic Literature Review (SLR), with the following central question: “How computational techniques are being used to analyse human mobility behavior in Smart Cities and Smart Environments?”. The review used a consolidated methodology described in [37], considering 6 databases of articles. A total of 5989 articles were initially found and filtered, resulting in 56 articles reviewed. To the best of our knowledge, this is the first study to review human mobility behavior covering a wide range of scenarios, including urban mobility, public transport, points and regions of interest, ridesharing, bike-sharing, traffic analysis, driving behavior, electric vehicle charging stations planning, mobility on demand, crowd analysis and others.

This article is organized as follows. Section 2 provides an overview of related work, specifically other literature reviews addressing similar research questions about computational solutions for identifying and predicting human mobility behavior. Section 3 outlines the applied methodology, while Section 4 presents the results of this study. Section 5 engages in a discussion of the findings, highlighting gaps and identifying research opportunities. Finally, Section 6 presents the conclusions.

## 2. Related work

This section presents reviews found during the filtering process of this study. We considered related work the studies that also review computational techniques for human mobility behavior analysis.

The state of the art of Crowd Analysis was reviewed in [10], with two subdivision axes: crowd statistics and crowd behavior analysis. The study summarizes and presents taxonomies and models of crowd behavior from other studies, focusing on: identifying the use of deep neural networks but not limited to; identifying which datasets are used (public/private) and data annotators. Pedestrian and group detection are treated as a specific session on the review, as it is considered essential tools for crowd analysis by the authors. A table is presented for each study, with date, research axis (such as: group behavior analysis, motion tracking, crowd statistics, anomaly behavior detection, crowd recognition, motion prediction, and others), dataset used, use or not of deep learning and availability of source-code. Also a table with detailed sensor and datasets types are presented. The purpose of the review is to find subareas in crowd analysis, that are still unexplored or that seem to be rarely addressed through the prism of Deep Learning. Finally, authors concludes about: the importance of data annotation tools, somehow neglected by the research community; group analysis-related tasks are not widely explored using Deep Learning methods, despite their widespread use in crowd analysis; massive crowd analysis for motion tracking and/or anomaly detection is not widely explored by the Deep Learning literature, due to the non-existence of relevant datasets.

Data generated by Telecommunication Companies (Telco) was explored in [14], presenting a small tutorial to convey a basic and advanced understanding of the unique characteristics, challenges and opportunities of Telco big data management. An overview the state-of-the-art in Telco big data analytics, on a set of basic pillars, namely: (i) background and respective architectures; (ii) real-time analytics and detection; (iii) experience, human behavior and retention analytics; (iv) privacy; and (v) storage. To quantize the data-size challenges, the example of city of Shenzend, China is presented, producing about 5TB of data per day from 10 million users. Also studies exploring customer churn predicition and groups behavior are presented. The authors conclude with open problems and future directions.

Video techniques for human activity recognition was surveyed in [55], reviewing 20 studies and classifying which algorithms/techniques recognized which activities. Going further, the authors evaluated some of the techniques (e.g. LSTF – Local Space Time Features, MIMM – Multiple Instance Markov Model, MRF – Markov Random Field, SFG – String of Features Graph) with three annotated datasets to find its precision. Finally, a table with the accuracy results is presented, with average values between 58% (SFG) to 93% (MIMM).

The use of GPS installed in taxis was reviewed in [13], to analyze mobility from various perspectives. The authors provided a formalization of the datasets and an overview of data processing methods. A timeline table was shown, categorizing the found articles into three main categories: social dynamics, traffic dynamics, and operational dynamics. Within social dynamics, articles were classified into techniques for extracting frequently visited locations and urban computing. Within operational dynamics, articles were categorized into passenger/taxi-finding strategies, route planning, route prediction and anomaly detection. Finally, discussions on the findings and directions for future research are presented.

Different from the aforementioned works, we reviewed the literature covering a wide range of scenarios related to human mobility behavior and computational techniques. This study explored six databases to contemplate more articles, following the guidelines proposed in [37], which recommend that systematic reviews encompass as much of the literature as possible. The main contributions of this study are:

- An overview of the challenges associated with analyzing behavior through computational techniques;
- A list of the techniques used by the studies;
- A compilation and taxonomy of the algorithms currently employed;
- A compilation and taxonomy of the machine learning techniques identified;
- A taxonomy of the data sources utilized, accompanied by a list of the identified datasets;
- A list of the devices and sensors employed in the reviewed studies;
- A list of the behaviors that are analysed by the techniques;
- Research opportunities and open issues related to the study theme.

### 3. Methodology

This section explains the methodology employed in this Systematic Literature Review (SLR), addressing the central question: “How computational techniques are being used to analyse human mobility behavior in Smart Cities and Smart Environments?”. The review used the methodology proposed in [37], considering 6 databases of articles to investigate 19 research questions, divided into 1 General Question, 16 Specific Questions and 2 Statistical Questions. A total of 5989 articles were initially found and filtered, seeking to identify the answers to the questions on the literature.

A systematic literature review seek to find, organize, centralize and summarize the knowledge found in existing works regarding a specific research topic, aiming to identify open gaps, trends and opportunities in the subject, through a systematic approach to provide unbiased results. Following the guidelines proposed in [37], the review was executed according to the following protocol:

1. Define the research questions;
2. Choose article bases;
3. Define the criteria for inclusion and exclusion of articles;
4. Perform selection, sorting, analysis and data extraction;
5. Organize and summarize the results found in a final report.

The research protocol was all planned and executed in the English language, due to the fact that the article databases that were searched are also in this language.

#### 3.1. Research questions

Table 1 presents the research questions defined for this study. The General Question seeks to understand how studies are using computational techniques to analyse human mobility related behaviors in Smart Cities and Smart Environments. Specific questions focus on extracting data and relevant information about the challenges, technologies, techniques, data, and other resources that were found. Finally, statistical questions are asked to evaluate publication sources and years.

#### 3.2. Performing the search process

We defined the search string with the following focus: search for “smart environments in general” (e.g. Smart Cities, smart homes) and “behavior”. To clarify, Smart City is also a Smart Environment, but on initial searches, we observed that a limited number of articles related to “smart cities” used the term “smart environment”. The same was identified with “smart homes”.

The first part of string contains terms related to Smart Cities and their synonyms. The second part and third part has terms related to Smart Environments and smart homes, respectively. The last part is related to ‘behavior’ and its

Table 1  
Research questions

Type	Question
GQ01	How computational techniques are being used to analyse human mobility behavior in Smart Cities and Smart Environments?
Specific Questions (SQ)	
SQ01	What are the challenges to analyse human mobility behavior through computational techniques?
SQ02	What techniques are currently used?
SQ03	What algorithms are being used?
SQ04	Does the study use Machine Learning (ML)? What are the ML techniques used?
SQ05	What are the data-sources used?
SQ06	What data-types are used?
SQ07	What types of devices/sensors are being used to collect the data?
SQ08	Are the datasets used by the study publicly available?
SQ09	Are the source-codes produced based on the computational models proposed in the study publicly available?
SQ10	Does the study use context histories and context prediction?
SQ11	In what kind of Smart Environment was the study carried out and/or validated?
SQ12	In which areas are the studies being applied?
SQ13	What behaviors are being observed?
SQ14	Are behaviors of a single entity (e.g. person), groups and/or crowds being observed?
SQ15	Are ontologies being used? In which scenarios are ontologies being used to identify and predict behaviors?
SQ16	What are the techniques used to validate the studies?
Statistical Questions (STQ)	
STQ01	Where the studies are being published?
STQ02	How many publications occurred per year?

Table 2  
Search terms

Domain	Synonyms	String
Smart Cities	Smart City, Smart Cities, Intelligent City, Intelligent Cities	("smart city" OR "smart cities" OR "intelligent city" OR "intelligent cities")
Smart Environments	Ambient Intelligence, Smart Environment, Smart Environments, Intelligent Environment, Intelligent Environments	("ambient intelligence" OR "smart environment" OR "smart environments" OR "intelligent environment" OR "intelligent environments")
Smart Homes	Smart Home, Smart Homes	("smart home" OR "smart homes")
Human Behavior	Behavior, Behaviors, Habit, Habits, Personality, Personalities	("behavior" OR "behaviors" OR "habit" OR "habits" OR "personality" OR "personalities")
Semantic resumed string		
("Smart Cities" OR "Smart Environments" OR "Smart Homes") AND ("Human Behavior")		
Full search string		
(("smart city" OR "smart cities" OR "intelligent city" OR "intelligent cities") OR ("ambient intelligence" OR "smart environment" OR "smart environments" OR "intelligent environment" OR "intelligent environments") OR ("smart home" OR "smart homes")) AND ("behavior" OR "behaviors" OR "habit" OR "habits" OR "personality" OR "personalities")		

synonyms. Table 2 shows the search terms organized by domains, alternative terms, the string summary semantics and the full string that was applied.

At the beginning of the protocol definition, initial searches helped to define the search terms, covering a wide range of scenarios related to behavior and computational techniques. After defining the search strings, the databases were selected, performing the search and review in April 2022. Table 3 shows the list of article databases and

Table 3  
Databases reviewed

Database	Strategy	Quantity
ACM	Search by title, abstract and keywords, considered the expansion "ACM Full-text collection". <a href="https://dl.acm.org">https://dl.acm.org</a>	174
IEEE	Search on all metadata, filtering only the type by "Magazines", "Conferences", and "Journals". <a href="https://ieeexplore.ieee.org">https://ieeexplore.ieee.org</a>	1628
Science Direct	Search filtered articles by the types "Research Articles" and "Review articles", considering title, keywords and abstract. A list of subqueries was defined, combining all of the possibilities of the search terms, to respect the database limit of 8 boolean connectors. <a href="https://www.sciencedirect.com">https://www.sciencedirect.com</a>	440
Springer	Search all metadata, removing "Preview-Only", and selecting Articles/Conference papers under the area of Computer Science, with English idiom selected. <a href="https://link.springer.com">https://link.springer.com</a>	2688
Wiley Online Library	Three searches were performed, on title, on keywords, and on abstract, with the research string, and with "Journals" selected. <a href="https://onlinelibrary.wiley.com">https://onlinelibrary.wiley.com</a>	57
Scopus	The search filtered by title, abstract, and keywords, with the document types filtered by "Article", "Conference Paper", and "Review", and domain related to Computer Science. <a href="https://www.scopus.com">https://www.scopus.com</a>	1002

Table 4  
Research criteria

Code	Description
Inclusion Criteria (IC)	
IC1	The study must be published in a conference, workshop or journal.
IC2	The study needs to contain the research string terms on title, keywords or abstract.
IC3	The study must be a full article.
IC4	The study must be written in English.
IC5	The study must be published between 2012 and April 2022 (10 years old)
Exclusion Criteria (EC)	
EC1	The study is a literature review or a systematic review.
EC2	The study does not have relation with the research theme.
EC3	The study is not available for reading or access.
EC4	The study does not focus on identifying or predicting behavior on smart homes, Smart Environments or Smart Cities.
EC5	The study is not related to human mobility behavior analysis.

summarizes the search strategy applied in each one.

### 3.3. Study filtering

After completing the search process, all data from the articles were organized and passed through a filter, considering Inclusion Criteria (IC) and Exclusion Criteria (EC). This literature considered only the articles that met all the inclusion criteria and excluded the articles that met at least one exclusion criterion. This method enables the exclusion of studies unrelated to the research topic, such as noise caused by database search algorithms, resulting in a refined selection of articles closely aligned with the review's scope. Table 4 shows all criteria.

An algorithm was developed to filter studies related to the specific criteria of IC2, which requires the article to have at least one term of each group in the title, keywords or abstract, as IEEEExplore and SpringerLink restrict the search method to fulltext. Detailing EC2, the interpretation of the main criteria is based on the selection of studies whose main objective is to analyse behaviors (e.g. identify and/or predict), based on any computational method found. Articles describing changes in user behavior only as a validation method were discarded. EC3 is only related to study access, and EC4 is mainly interpreted as the domain of studies should be related to Smart Environments, Smart Cities or smart homes.

After applying the EC1, EC2, EC3 and EC4, a total of 357 articles resulted, related to behavior analysis in Smart Environments and Smart Cities. Figure 1 shows the articles classified by study area. The majority of studies

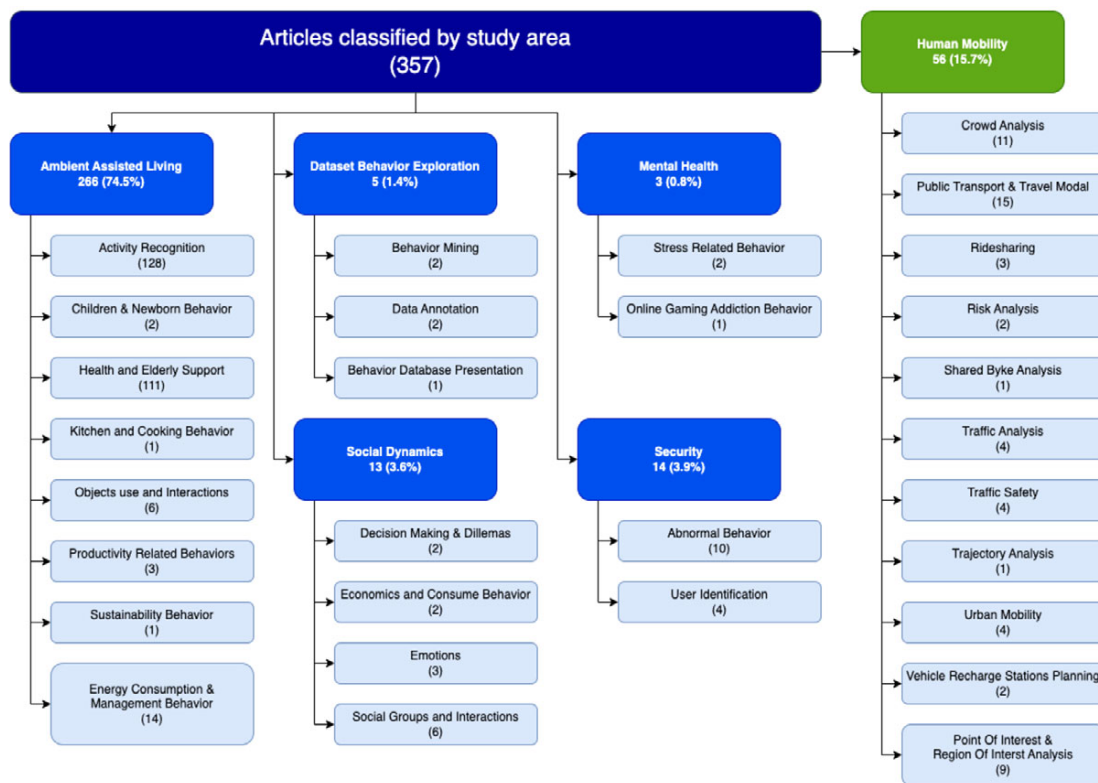


Fig. 1. Articles classified by study area. Selected articles in green.

focus on Ambient Assisted Living (AAL) with 266 articles (74.5%). Activity Recognition (AR) and ‘Health and Elderly Support’ emerge as the two primary subareas within AAL, featuring 128 and 111 articles, respectively. Human mobility related behavior represents 15.7% of the studies, with 56 articles. Other identified areas include Security (14–3.9%), Social Dynamics (13–3.6%), Database Behavior Exploration (5–1.4%), and Mental Health (3–0.8%). Some studies could be classified into two or more subareas, with the most specific association being selected.

Figure 2 presents other reviews found during this study, classified by focus area. Despite representing 15.7% of the studies identified, the analysis of human mobility behavior covering a wide range of scenarios was not yet explored by other systematic literature reviews. Therefore, the exclusion criterion EC5 was added to concentrate the research in this area. The application of this criterion removed 301 studies, resulting in 56 articles. Figure 3 shows the complete filtering process described in this section.

#### 4. Results and answers

This section aims to answer the research questions presented in Table 1. All articles selected in this review are listed in Table 14, located in the Appendix. A code was assigned to each article for better organization within the text, for example, ‘A01’.

##### 4.1. SQ01 – What are the challenges to analyse human mobility behavior through computational techniques?

The following items present the main challenges found in the articles:

- **Behavior identification and modeling:** Part of the main focus of all studies, explored with different techniques, as detailed in SQ02.

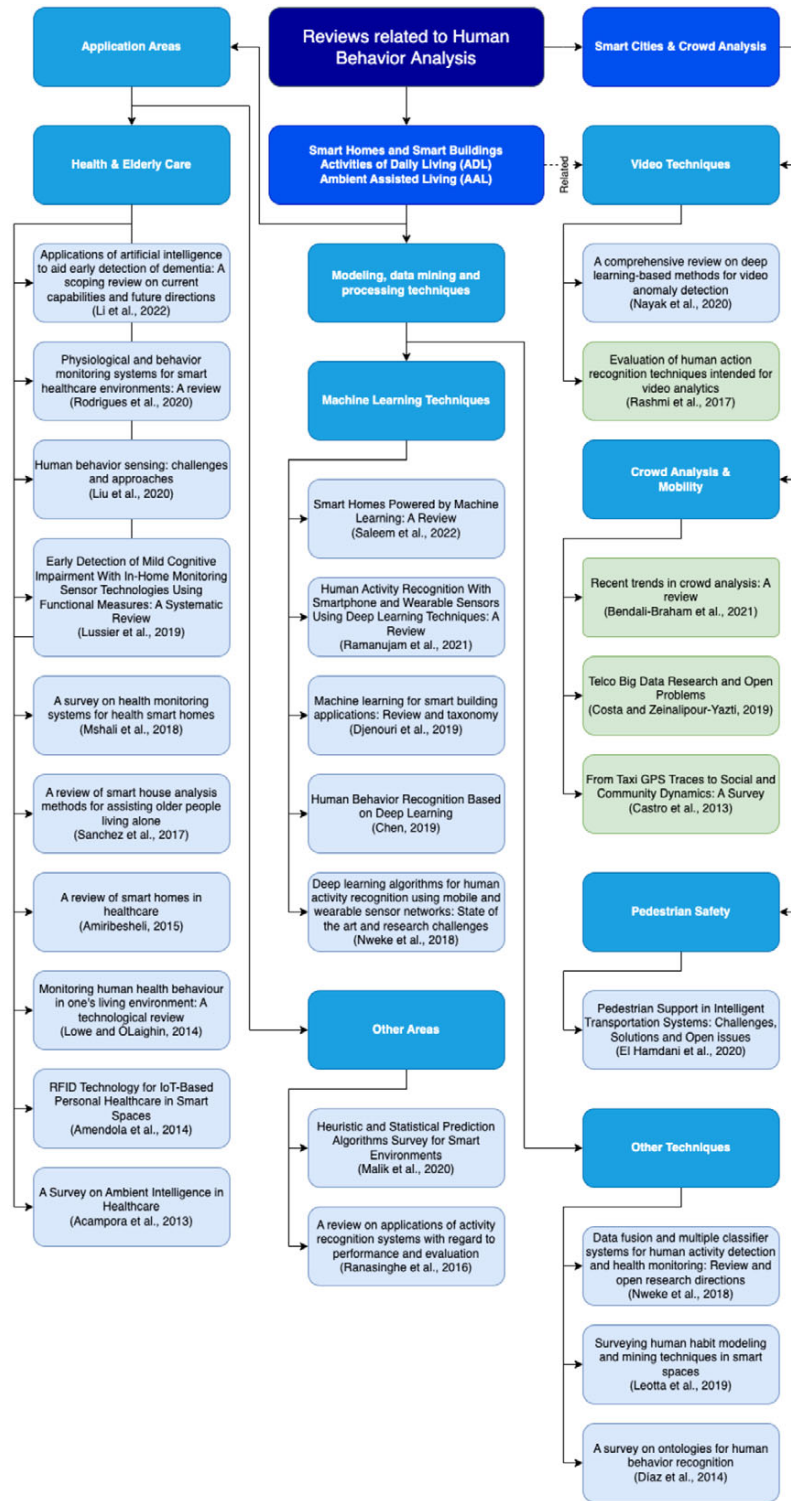


Fig. 2. Reviews classified by focus areas. Related works in green.



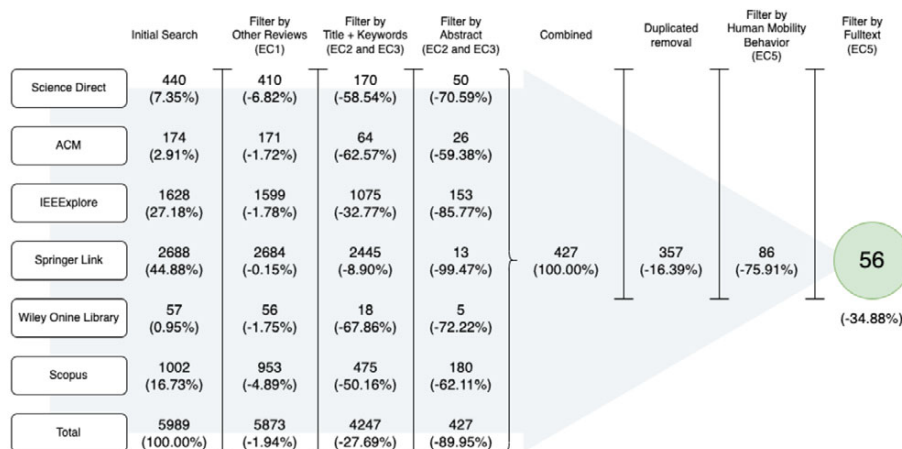


Fig. 3. Review study filtering overview.

- **Issues related to a large volume of data – Bigdata:** With sensors producing data in real time or GPS sensors generating data every 30 seconds from thousands of people or vehicles, the volume of data quickly generated scale to large proportions. Trajectories of private vehicles was collected by [45] and the generated dataset was made publicly available, arguing that the data quickly arrived in the order of terabytes. With regard to data from Telecommunications companies (Telco or Telecom), georeferencing data from users by radio triangulation can reach Terabytes per day [32]. These large amounts of data lead to other directly related problems, such as storage capacity, processing, and transfer speed of this data to data centers.
- **Real-time processing demand:** Applications that use video cameras in high resolution formats (e.g. FullHD) also quickly reach the order of gigabytes per day. In this specific case, some applications such as monitoring and security demand real-time processing, not always available on edge computing devices. The ability to transfer this data to servers that can perform the processing becomes an important issue to be explored, as well as processing latency, which can directly impact the final result of the application [75].
- **Inaccurate or incomplete data:** Data generated by Telecom has significant geospatial inaccuracy. According to some studies, the accuracy varies between 25 m, 50 m, 100 m [22], and in some cases may exceed 200 m, in addition to eventually having sparse temporal fragmentation (e.g. hourly) [24,59]. Other data sources, such as those generated by public Wifi Routers also have georeferencing inaccuracy, directly related to the range of these equipment, around 25 m [9]. This inaccuracy may make it unfeasible to use the data for some applications. Data generated from experiments with volunteers or crowdsourcing can also generate inaccurate results, since the forms may not always be filled out correctly, or not all data may have been collected by user forgetting [63]. On platforms that are gamified, or that have some aspect of competition, users may also tend to fill only advantageous data about their behavior [35].
- **Cost of acquisition, installation and maintenance of devices and sensors:** The problem of costs for acquisition and maintenance of resources is very common, and generally affects almost any solution. Studies that use devices and sensors that cover certain regions inevitably need to acquire and install new sensors to expand their scope, such as: public Wifi Routers [9,25]; use of cameras and optical sensors for crowd monitoring [7,75,80,84]; use of RFID sensors or presence identifiers [15], among others. Studies that use sensing on the observed entity itself (e.g. person or vehicle) have the same cost problem, for example: cameras in vehicles [56]; Use of GPS in vehicles and accelerometer sensors [67,68], among others.
- **Privacy, legal or commercial restrictions on data availability:** Several of the articles found used mobility datasets based on Telecom datasets [17,18,22,32,59,73,74,86], Public Transport systems [11,38,40,41,46,50] or private companies trajectory databases [24,33,39,78,79], but they did not make the data available due to legal, privacy or commercial restrictions, despite anonymizing the data when using them in the research, which leads directly to the next problem.

- **Few mobility datasets available or datasets not annotated:** Many of the studies used restricted access datasets, which could not be made available. Probably due to these restrictions, a shortage of public mobility datasets to be explored by other researchers was observed. This was quite evident when observing the study by [47], published in 2017, using a dataset from 2007 (10 years at the time), and even more evident in the study by [34], published also in 2017, using a dataset generated in 2004 (13 years at the time), before the popularization of smartphones with GPS. Another problem is the lack of annotated datasets. In the study of [77] researchers manually annotated 1000 vehicles to generate the training database. This problem is also discussed by [10], a review in which the authors express concern about the non-existence of these annotated datasets. DiDi,<sup>3</sup> one of the largest ridesharing company in the world created the Gaia Initiative,<sup>4</sup> addressing to provide anonymized mobility datasets to support research in transport area.
- **Lack of tools for database annotation:** As also identified by [10], there is a lack of studies related to building tools and computational models to support database annotation. Only two studies, out of the initially identified 357 in this review, specifically addressed data annotation-related issues.

#### 4.2. SQ02 – What techniques are currently used?

Table 5 presents the techniques applied on each article. The study found several combinations of techniques, types of sensors and data crossings, ranging from the use of smartphone sensors to identify aggressive driving, use

Table 5  
Techniques used to identify and predict behaviors

Article	Technique
A01	Use of Wifi Access Points to collect data from smartphones that travel through the region to identify anomalies.
A02	Use of algorithms in Telecom data for analysis of urban activity by region.
A03	Use of Bluetooth (BT) readers to identify the presence of users through data emitted by their smartphone.
A04	Identification and prediction of presence clusters in Regions of Interest (ROI) and/or urban functional regions through the use of algorithms in databases of private vehicle trips.
A05	Identification and prediction of travel intentions through algorithms applied in Telecom datasets.
A06	Crossing data from the California Household Travel Survey (CHTS), Twitter and Google Places, and applying algorithms to identify and predict users' travel intentions.
A07	Identification of travel patterns by applying algorithms in databases of private vehicle trajectories.
A08	Analysis of pedestrian emergency evacuation behavior through simulation.
A09	Use of ERI (Electronic Registration Identification) in vehicles to record the history of trajectories, through RFID devices.
A10	Use of smartphone sensors (accelerometer, gyroscope, magnetometer and GPS) to identify aggressive driving behaviors.
A11	Use of historical GPS trajectories provided by users, collected through an application for smartphones, also containing questionnaires about user behavior.
A12	Use of thermal cameras in public spaces to identify pedestrian movement patterns. Use of T-Analyst data annotation software.
A13	Use of a smartphone application to track users' transport mode and trajectory, and use forms to collect information about the transport mode.
A14	Use of smartphone application to collect user coordinates every 30 minutes, and apply algorithms to evaluate the influence of walkability on traffic behavior.
A15	Use of vehicle cameras to identify traffic violations.
A16	Use of Wifi Access Points in the city to monitor people traffic, and points of origin and destination.
A17	Use of public transport transaction data with smart cards to identify travel patterns, and passenger's boarding and disembarkation.
A18	Use of weighted methods for better precision in the DTW algorithm, aiming to identify similarity patterns in traffic data.

<sup>3</sup><https://web.didiglobal.com> – Access in June 2023.

<sup>4</sup><https://www.businesswire.com/news/home/20180111005837/en/DiDi-Expands-GAIA-Initiative-Worldwide-to-Facilitate-Data-driven-Research-in-Transportation> – Access in June 2023.

Table 5  
(Continued)

Article	Technique
A19	Use of Multi-object tracking algorithms to identify pedestrian trajectories in camera videos.
A20	Use of deep-learning algorithms to identify abnormal pedestrian trajectories and behaviors in public camera videos.
A21	Use of transaction data in public transport through payments via smart card and information filled in by the user, to identify travel, boarding and disembarking patterns.
A22	Use of LIDAR (Light Detection and Ranging) for crowd movement tracking, combined with applications collecting accelerometer and microphone data for indoor location identification.
A23	Use of Heavy-Goods Vehicles traffic datasets to identify mobility patterns and regions of interest.
A24	Use of Telecom data for urban mobility analysis.
A25	Application of algorithms in videos from security cameras to identify user behavior.
A26	Use of game theory and deep-learning to predict pedestrian movement based on photographs.
A27	Use of the smartphone accelerometer to identify sudden behavior and aggressive driving by the driver.
A28	Use of Telecom data for urban mobility analysis, in conjunction with population sense information and household surveys.
A29	Use of a gamified platform to analyze and influence the means of transport chosen by users, crossing public transport data collected through the GTFS (General Transit Feed Specification) standard.
A30	Crossing of usability data from an electric vehicle charging application and readings at charging stations, aiming at analysis and planning for the installation of new stations.
A31	Use of user trajectories data on a Bigdata scale, for analysis of origin-destination travel patterns.
A32	Article performs a descriptive analysis of data from the 2017 National Household Travel Survey, applying the Multiple Discrete-Continuous Extreme Value (MDCEV) method to analyze the dynamics of people's "transport modal style", also analyzing Mobility on Demand.
A33	This study proposes a trajectory summarization and noise reduction technique for the extraction of the locations of interest from symbolic trajectories based on Telecom dataset.
A34	Use of security cameras connected to a server to process images using Social Force Model (SFM) algorithms to identify anomalous behavior. When behaviors are identified, or a large volume of people in a location is identified, a notification is triggered according to the user's category and interest in the monitored region.
A35	Simulation of electric vehicle charging behavior changes in Smart Cities when introducing variable energy price.
A36	An application was created, integrated with the Sii-Mobility API and the KM4City data structure, to identify behaviors and send notifications with information and survey feedback request from users.
A37	Uses vehicular monitoring within Smart Cities through Automatic License Plate Recognition (ALPR) techniques and vehicle discrimination object detection, in order to obtain timely statistical data.
A38	Use of the smartphone's accelerometer to identify changes in user behavior, extrapolating this collectively to identify emergency situations and notify users nearby.
A39	Use of data from Telecom 'Big Data Challenge' to process origin-destination matrices in the city of Milano – Italy, in conjunction with other data on traffic, crime, pollution, events, among others, proposing an adaptive routing strategy algorithm to reduce the overall traffic in a Smart City.
A40	Use of Telco big data to compute mobility patterns for Smart Cities. Several big data mining algorithms have been discussed to support the prediction and estimation of vehicular traffic conditions, speed profiles for roads, flows of people moving among subway stations and around POIs of the city, and also O/D matrices, providing technical aspects to support real-time big data analysis.
A41	Use of engine speed, accelerometer, GPS, speed, gear, pedal (brake, clutch or accelerator) sensors to collect driving behavior data to achieve a comprehensive perspective of driving performance. Data were collected from 4 volunteer drivers, on the same stretch, during 4 trips.
A42	Use of a smartphone application to monitor and suggest sustainable transport routes, aiming to change users behavior.
A43	Processing smartcard data, separated into clusters of weekly and temporal usage patterns, for statistical evaluation of these user patterns and year-to-year changes.
A44	This work proposed a framework for Recognizing the Crowd Mobility Patterns in Cities using Location-Based Social Networks (LBSNs) data. The framework comprises four main components: data gathering, recurrent crowd-mobility patterns extraction, temporal functional regions detection, and visualization component.
A45	The article uses stereoscopic cameras to count people and Wifi routers, counting signals and MAC ADDRESS emitted by cell phones, and cross-references the final counts and anonymizes necessary data. Three elements are analyzed: Crowd estimation, flow of people and length of stay. The data is anonymized and sent to the cloud through the framework's data exchange standards.

Table 5  
(Continued)

Article	Technique
A46	Article proposes the Adaptive Input Infinite Hidden Markov Model (AI-iHMM), aiming to consider different time variables to identify state patterns in the analyzed behavioral data. It is an adaptation to HMM.
A47	This work proposed a framework named TBI2Flow, based on the Travel Behavioral Inertia (TBI) from taxi GPS records, which embodies Driver Inertia (DI) and Passenger Inertia (PI). TBI was integrated with other features to construct multi-dimensional features and predict taxi passenger flow based on a deep learning algorithm.
A48	The study analyzed the density of users in Shanghai city from geolocation data of Weibo to compare their density through univariate and bivariate density estimation techniques: Point Density and Kernel Density Estimation (KDE), with the following findings: (i) characteristics of users' spatial behavior, the center of activity based on their check-ins, (ii) the feasibility of check-in data to explain the relationship between users and social media, among others.
A49	Using Call Detail Records (CDR) from users to determine the classification of specific regions of the city, example: Commercial, Industrial, Residential, Offices, etc.
A50	Analysis of transactions on users smart cards to statistically verify the passenger behavior change in bus lines during peak traffic and non-peak traffic situations.
A51	Article uses "Network Science" and graph science to analyze travel patterns and travel reasons extracted from the two datasets used. Several histograms are presented performing analyzes in relation to the data and raising behavioral hypotheses.
A52	Use of a shared bicycle trajectories dataset (dockless) from Mobike to perform analysis of users' mobility patterns. A data-driven framework was established to integrate multiple data sources, including transportation network data, road characteristics, and urban land use, to achieve a detailed, accurate analysis of cycling patterns at both the individual and group levels.
A53	A descriptive analysis of data from two bike-sharing and ride-hailing datasets is presented. Through this analysis, the authors report having found a double logarithmic power law distribution in certain time intervals of the analyzed data, among other findings.
A54	This article proposed a multi-pattern passenger flow prediction framework, MPGCN, based on Graph Convolutional Network (GCN). Firstly, a sharing-stop network to model relationships between passengers based on bus record data was constructed, then, employed GCN to extract features from the graph by learning useful topology information and introduce a deep clustering method to recognize mobility patterns hidden in bus passengers. To fully utilize spatio-temporal information, proposed the GCN2Flow to predict passenger flow based on various mobility patterns.
A55	The study proposes an entropy estimator named ContextTransition entropy that can capture both the sequential orders of human mobility and these contextual information, to derive the limits of predictability.
A56	Analysed the transfer rate of empty cars and taxis using a dataset provided by DiDi Chuxing Gaia Initiative, to classify urban functional regions through taxi behavior. The authors proposed an attentional spatio-temporal model (Attentional Gated Recurrent Unit, AGRU), based on three modules, which are the spatial feature extraction module, the temporal feature extraction module, and the attentional pooling mechanism.

of public transport ticketing databases for analysis and prediction of travel patterns, to the use of Telco bases for urban mobility analysis.

#### 4.3. SQ03 – What algorithms are being used?

Table 6 lists the algorithms used in the articles. Not all studies identify the names of the used algorithms. The aim of this review is not to extensively and comprehensively explore the algorithms used, but rather to identify the methods employed. Figure 4 presents a taxonomy of the algorithms grouped by type. The articles IDs associated with each algorithm are enclosed in brackets.

Multi-Object Tracking algorithms are represented in Table 6 only as MOT, because that is a vast area of research, which explores video processing techniques for tracking objects, involving a wide range of algorithms (e.g. FairMOT, TransCenter, CenterTrack, CTracker). This link<sup>5</sup> contains several articles, open-source repositories of algorithms, datasets, and other information about it.

<sup>5</sup><https://github.com/luanshiyinyang/awesome-multiple-object-tracking> – Multi-Object Tracking (MOT) information repository.

Table 6  
Algorithms found

Short name	Name	Articles
AA	Average Accuracy	A51
AI-HMM	Adaptive Input Hidden Markov Model	A46
AMI	Adjusted Mutual Information	A18
ANCOVA	Analysis of Covariance	A50
ANOVA	Analysis of Variance	A48
ARIMA	Autoregressive Integrate Moving Average	A04, A16, A47, A54
BFS	Breadth First Search	A20
BP	Bayes Predictor	A55
BPF	Band Pass Filter	A10
BWT	Burrows-Wheeler Transform	A51
CC	Correlation Coefficient	A47, A54
CEM	Classification Expectation Maximization	A43
CTE	ContextTransition Entropy	A55
DA	Dijkstra Algorithm	A52
DBSCAN	Density-Based Spatial Clustering of Applications with Noise	A02, A17, A20
DT	Decision Tree	A51
DTW	Dynamic Time Warping	A10, A18
EM	Expectation Maximization	A43
FMC	Factorized Markov Chains	A03
FPMC	Factorizing Personalized Markov Chains	A03
GK	Gaussian Kernel	A44
GWDTW	Gaussian weighted DTW	A18
HA	Historical Average	A04
HAC	Hierarchical Agglomerative Clustering	A44
HMM	Hidden Markov Model	A20, A31
ICL	Integrated Completed Likelihood	A43
IERP	Improved Edit distance with Real Penalty	A07
IPP	Inhomogeneous Poisson Process	A02
JA	Jenks Algorithm	A33
K-fold	K-fold	A40
K-means	K-means	A16, A43, A46, A54
K-means++	K-means++	A04
K-medoids	K-medoids	A16, A18
KDE	Kernel Density Estimation	A44, A48
KNN	K-Nearest Neighbors	A05, A20
KPCA	Kernel Principal Component Analysis	A07
LDA	Latent Dirichlet Allocation	A44
LinR	Linear Regression	A40
LogiR	Logistic Regression	A40, A51, A54
MAE	Mean Absolute Error	A47, A54
MC	Markov Chains	A3
MDCEV	Multiple Discrete Continuous Extreme Value	A32
MDP	Markov Decision Process	A26
MF	Matrix Factorization	A03, A55
MLE	Maximum Likelihood Estimator	A16
MOT	Multi-Object Tracking	A19
NMF	Non-negative Matrix Factorization	A44

Table 6  
(Continued)

Short name	Name	Articles
OD-Matrix	Origin Destination Matrix	A16, A28, A40, A54
PCA	Principal Component Analysis	A46
PCC	Pearson Correlation Coefficient	A52
PD	Point Density	A48
PL	Power Law	A53
PoissonR	Poisson Regression	A52
PolyR	Polynomial Regression	A40
RDI	Richness Diversity Index	A33
RE	Real Entropy	A55
RLE	Run-Length Encoding	A33
RMSE	Root Mean Square Error	A47, A54
ROCAUC	Receiver Operating Characteristic/Area Under the Curve	A20
S-H-ESD	Seasonal Hybrid Extreme Studentized Deviate	A01
SC	Spectral Clustering	A44
SDI	Simpson Diversity Index	A33
SeqScan-D	SeqScan-D	A33
SFM	Social Force Model	A34
SLIC	Simple Linear Iterative Clustering	A25
SMA	Simple Moving Average	A10
SWEI	Shannon-Weiner Entropy Index	A33, A55
TF-IDF	Term Frequency-Inverse Document Frequency	A44
VA	Viterbi Algorithm	A52
VW	Visvalingam-Whyatt	A28
WCL	Weighted Centroid Localization	A45
WMF	Weighted Matrix Factorization	A55

#### 4.4. SQ04 – Does the study use Machine Learning (ML)? What are the ML techniques used?

This question delves deeper into understanding the Machine Learning (ML) techniques identified in this review. Despite this, the objective of this research is not to comprehensively explore the entire field of ML, which would require dedicated research on the subject. Table 7 lists the ML techniques identified.

During the review, we observed that 20 articles used ML, with a total of 35 techniques. The articles that most used ML are: i) [45] brought 10 references; ii) [38] with 8; iii) [78] with 7; iv) [39] with 4. Most of these algorithms was used as comparison baselines to validate the proposed models. Articles in ii) and iv) has authors in common, suggesting a potential affiliation within the same research group. Figure 5 provides an additional taxonomy of Machine Learning techniques organized by type.

Figure 6 presents a bar chart of techniques ordered according to their frequency of usage. The Support Vector Machine (SVM) emerges as the most frequently employed method in the reviewed articles, with a total of 4 occurrences. This is followed by Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), each appearing 3 times. Support Vector Regression (SVR), Diffusion Convolutional Recurrent Neural Network (DCRNN), and Convolutional Neural Network (CNN) are each utilized twice. All other techniques are represented by a single instance of usage.

Figure 7 illustrates the temporal evolution of ML usage. The data show an initial increase in use beginning in 2015 and continuing until 2017, followed by a subsequent increase from 2019 to 2022, indicating a trend of growing interest.

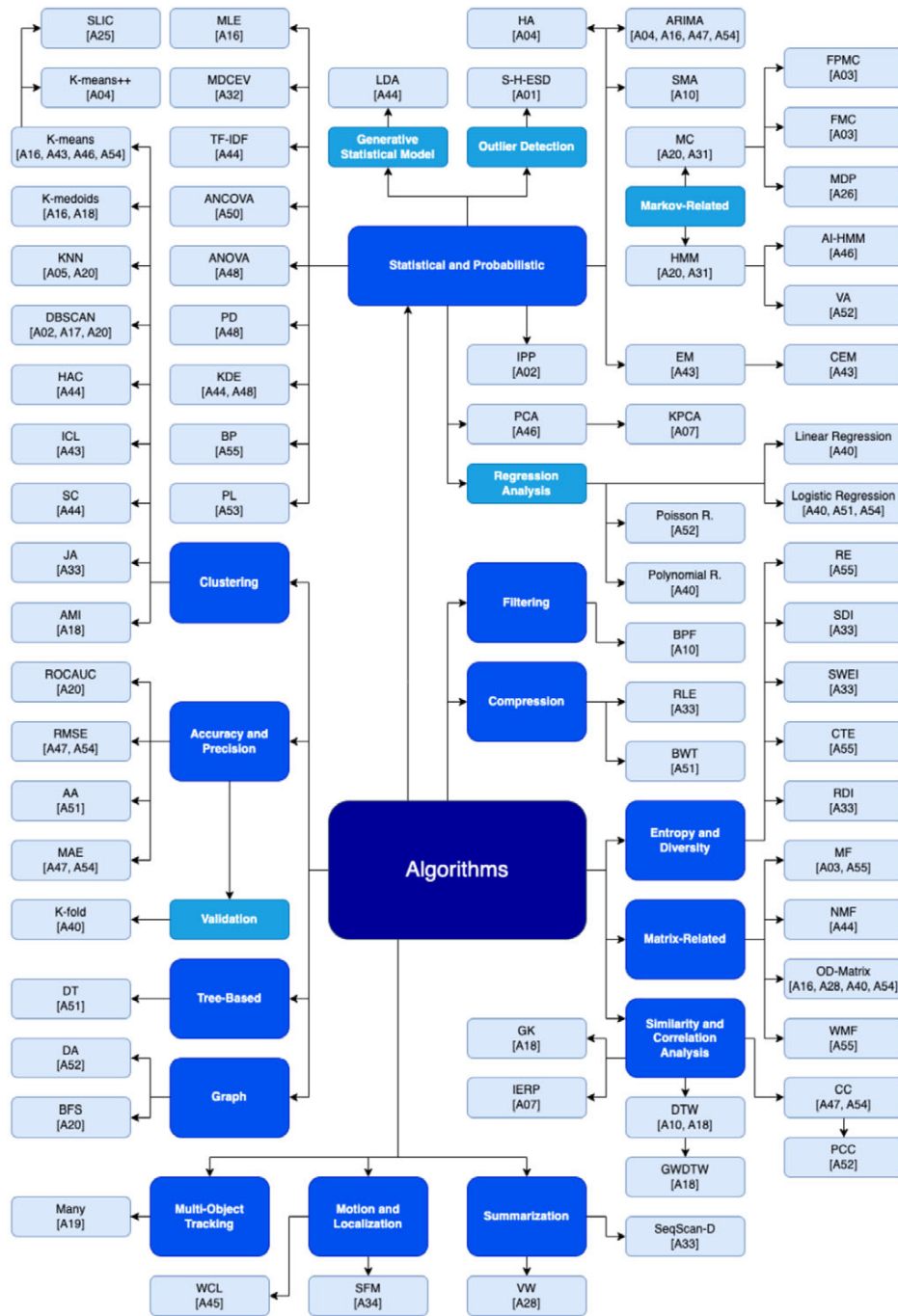


Fig. 4. Algorithms being used on studies.

4.5. SQ05 – What are the data-sources used?

Figure 8 presents a taxonomy of the data sources. The dataset usage description are listed below:

- Telecom databases was explored by 8 articles.
- Big data containing GPS trajectories collected through industry applications was used in 8 studies.

Table 7  
Machine Learning algorithms

Short name	Name	Articles
AARNN	Attentional pooling in Attentional Recurrent Neural Network	A56
AGRU	Attentional Gated Recurrent Unit	A56
ALSTM	Attentional Long-Short Term Memory	A56
ARNN	Attentional Recurrent Neural Network	A56
BMA	Bayesian Model Averaging	A14
BO-SVR	Support Vector Regression approach based on Bayesian Optimization	A04
BPNN	Back Propagation Neural Network	A47
CNN	Convolutional Neural Network	A20, A37
ConvLSTM	Convolutional LSTM	A25
DBN	Dynamic Bayesian Network	A06
DBN	Deep Belief Network	A47
DCRNN	Diffusion Convolutional Recurrent Neural Network	A04, A54
DenseGCN	Dense Graph Convolutional Networks	A04
GBRT	Gradient Boost Regression Tree	A04
GBDT	Gradient Boost Decision Tree	A54
GCN	Graph Convolutional Network	A54
GMAN	Graph Multiattention Network	A04
GNB	Gaussian Naive Bayes	A51
GRU	Gated Recurrent Unit	A04, A46, A56
GWN	Graph WaveNet	A04
LDA	Latent Dirichlet Allocation	A44
LSTM	Long Short Term Memory	A54, A55, A56
MGDCN	Multigraph Dense Convolutional Network	A04
MLNN	Multi-Layered Neural Network	A47
MP-LSTM	Multi-Pattern LSTM	A54
MP-GCN	Multi-Pattern Graph Convolutional Network	A54
Multi-LSTM	Multidimensional LSTM	A04
RF	Random Forest	A49
S-BPR	Sequential Bayesian Personalized Ranking	A03
SSD	Single Shot MultiBox Detector	A15
ST-GCN	Spatio-Temporal Graph Convolutional Networks	A54
SVM	Support Vector Machines	A11, A13, A23, A51
SVR	Support Vector Machine for Regression	A47, A54
T-GCN	Temporal GCN	A04
TrAdaBoost	Boosting for Transfer Learning Algorithm	A07

- 6 articles used ticketing data through Bus/Public transport.
- One article used a shared-bike pick/drop dataset.
- One article used an electric vehicle recharge dataset.
- Local Position Registering was explored by 8 articles, with 5 using WIFI technology, 3 using Bluetooth (two also used WIFI), and 2 using RFID readers.
- User-filled data was used by two articles, with one utilizing Location Based Social Networks (LBSN) and one using form/survey data.
- Smartphone sensors were explored by 9 articles, with 4 using accelerometers, 6 using GPS, and 1 using a microphone.
- Image-based sources were used by 11 articles, with 8 using public/security cameras, 1 using thermal cameras, 3 using traffic cameras, 1 using vehicle internal camera, 2 using LIDAR, 1 using a vehicle counter, and 1 using a vehicle plate recognizer.



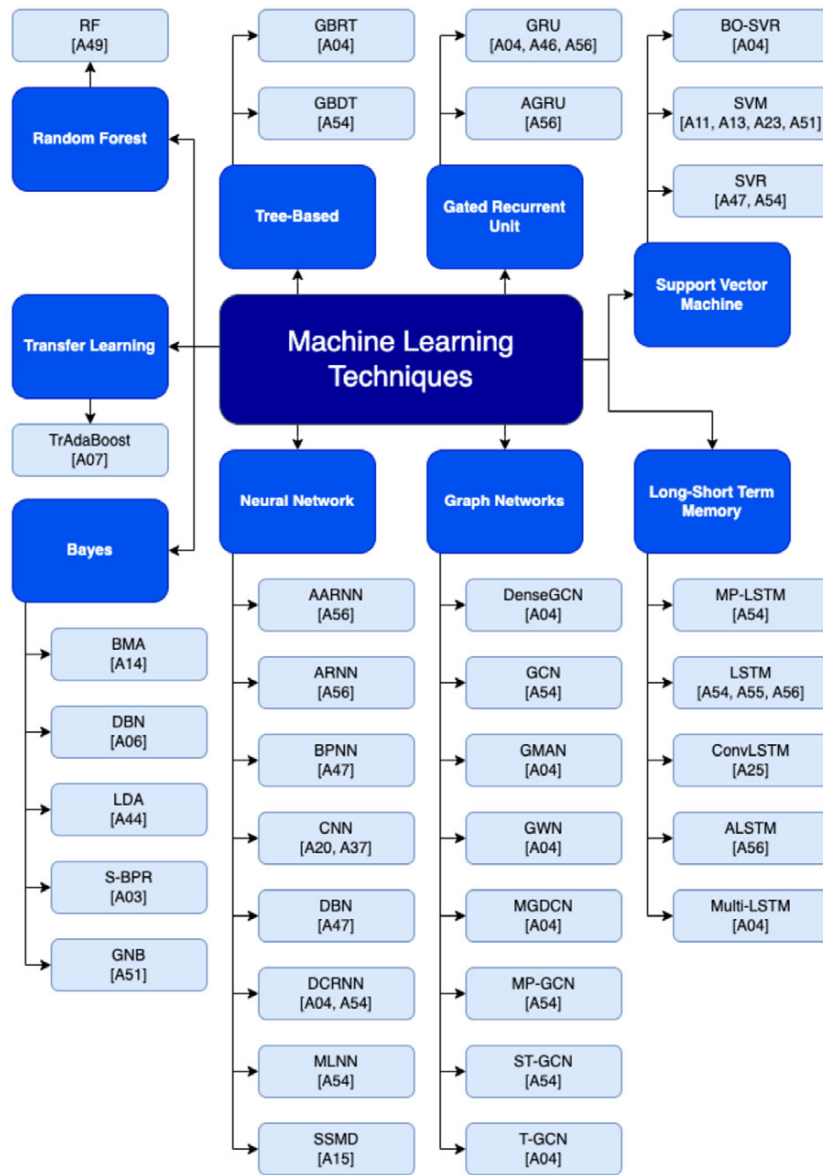


Fig. 5. Machine Learning techniques being used on studies.

- Government or Public Data was explored by 13 articles, including one using GTFS, one using land-use datasets, 3 using public surveys, and 7 using Points of Interest (POIs) datasets.
- GPS Travel History datasets were explored by the majority of the articles, with 16 articles. 8 used users’ smartphones to generate the datasets, 1 used shared-bike history, 1 used Goods transport history, 1 used public transport, 4 used Ridesharing datasets, and 3 used private vehicles datasets.

In the studies of [62] and [63], data was generated through volunteering by crowdsourcing with smartphone app, with 1,000 people generating data on 125,000 trips [62], and 8,000 generating 150,000 trips [63]. Both articles have authors in common, probably being the evolution of the research over time. Data collection through volunteering also poses major challenges, from coordination, motivation, and others already reported in SQ01.

Figure 9 presents a bar chart of data-source type usage by articles, demonstrating the predominant interest in utilizing GPS Travel History datasets from various segments in research, with 16 articles. Nonetheless, the importance

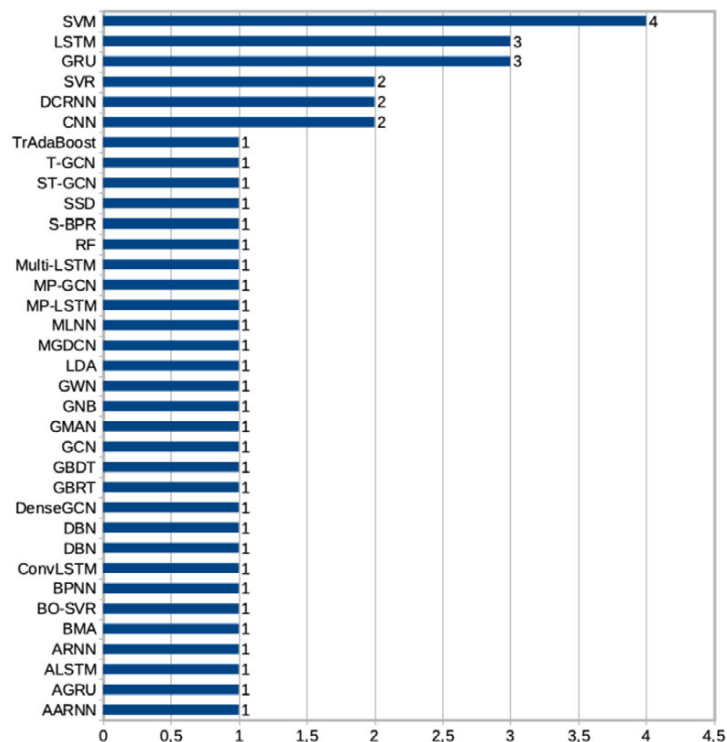


Fig. 6. Machine Learning techniques: usage count.

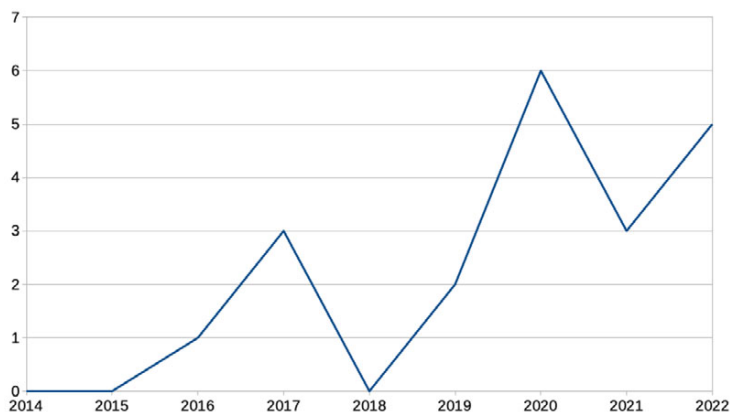


Fig. 7. Yearly count of articles using machine learning.

of public datasets and open data policies is also evident, as identified by the use of Government or Public Data in 12 articles. The same rationale can be applied to Usage/Ticketing databases, used by 11 articles, and mobility databases based on Telecom data, utilized in 8 studies. Together, these four groups represent 35 articles (some articles use more than one type of data-source), highlighting the importance of strengthening partnerships with industry and public open data policies to support scientific research.

#### 4.6. SQ06 – What data-types are used?

In both industry and academia, trajectory storage is typically accomplished by collecting multiple GPS coordinates along with the record date-time (Timestamp), forming the contextual history of the observed moving entity.

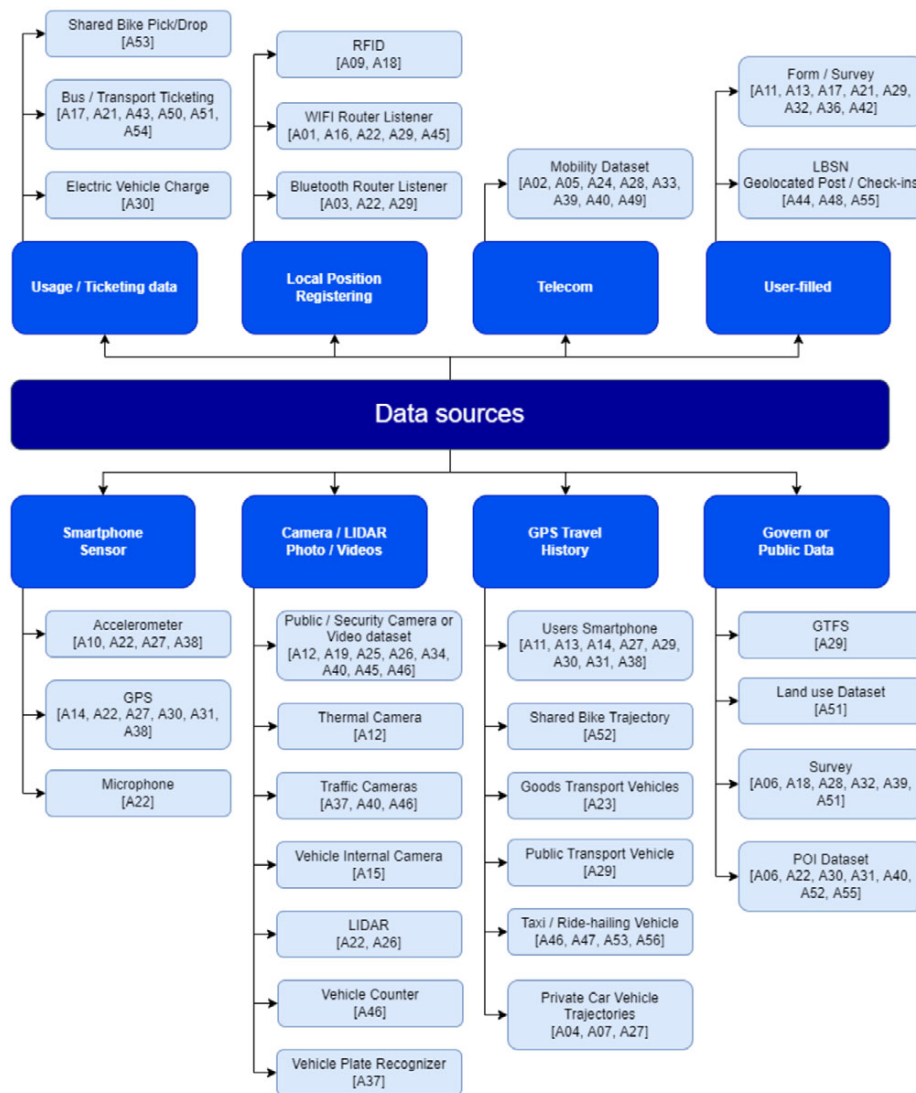


Fig. 8. Data sources used on the studies.

This question is intended to explore other types of data that are utilized and analyzed in conjunction to identify behaviors. Table 8 lists data-types identified.

4.7. SQ07 – What types of devices/sensors are being used to collect the data?

Installing sensors to cover entire cities and geographically wide environments is a challenge that demands many resources, already explored in SQ01. Table 9 shows devices and sensors identified on the review. Common types of equipment used for data collection were identified, including Telecom using Radio Network Controllers (RNC); Bigdata using location data generated by users’ smartphones; videos generated by cameras; smart card transactions being processed by Bus Ticketing equipment, among others.

It was also observed the use of less explored sensors, such as accelerometers of smartphones to identify driving behaviors [5,67,68]. The use of the accelerometer is also explored by [80] to monitor displacement behavior between indoor points, in addition to the use of audio collected by the microphone to identify crowding levels around a pedestrian.

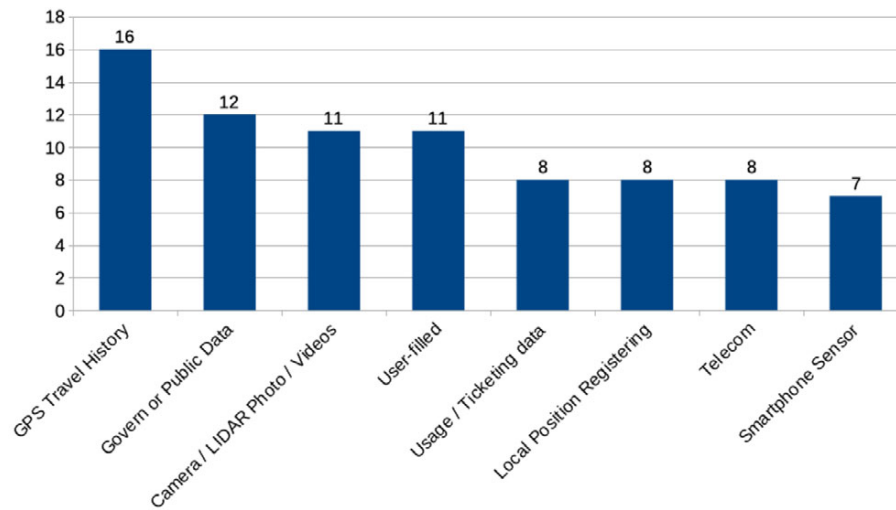


Fig. 9. Total or articles using each data source group.

Table 8  
Data-types identified in review

Data-type	Articles
Accelerometer data	A10, A22, A27, A38, A41
Bike-sharing Stations Location	A51, A53
Building Shape/Blueprint	A08
Bus Stops Location	A02, A06, A13, A17, A21, A29, A36, A42, A43, A50, A52, A54
Electric Vehicle Recharge Station Location	A30, A35
Electric Vehicle Recharge Event	A30, A35
GPS Points + Timestamps	A02-07, A09, A11, A13, A14, A16, A17, A21, A23, A27-31, A33, A35, A36, A38-44, A46, A47, A49, A50, A52, A53, A54, A56
Human Body Images	A20, A22, A26, A34
Human Skeleton Poses Points	A26
Images (Videos/Photos)	A12, A15, A19, A25, A26, A34, A37, A40, A45
Indoor Location + Timestamps	A08, A22, A45
LIDAR sensor data	A22, A26
LBSN Check-in or Geolocated Post	A06, A22, A44, A48, A55
Magnetometer data	A10
Parking Lots Location	A02, A13, A30, A36
Points of Interest (POIs) Location	A02, A04, A06, A13, A16, A22, A31, A35, A36, A40, A44, A48, A52, A55
Public Geolocated Statistics (ex. Land-use datasets or surveys)	A06, A18, A28, A32, A39, A51
RFID Readings	A09
Smartphone Bluetooth MAC Address	A03, A22, A29
Smartphone WIFI Mac Address	A01, A16, A22, A29, A45
Gyroscope data	A10, A22
Streets GPS Lines	A02, A04, A13, A14, A16, A21, A23, A28, A29, A35, A40, A41, A42, A47, A48, A50, A52, A53
Ticketing/Payment Register	A17, A21, A43, A50, A51, A53, A54
Train Stations Location	A13, A17, A21, A28, A29, A40, A42, A51, A52
Vehicle Traffic Count	A18, A40, A46
Weather data	A47, A52

Table 9  
Sensors/devices used to collect data by articles

Sensor	Articles
Accelerometer	A10, A22, A27, A41
Access Point (WIFI & Bluetooth)	A01, A03, A16, A29, A45
Camera	A15, A19, A20, A25, A26, A34, A37, A45, A46
Electric Vehicle Charging Station	A30
Gyroscope	A10, A22
GPS	A04, A07, A10, A11, A13, A14, A17, A21, A23, A27, A29, A30, A31, A41, A43, A47, A52, A53
LBSN Check-ins or Geolocated Post	A06, A44, A48
LIDAR camera	A22, A26
Magnetometer	A10
Microphone	A22
RFID Reader	A09
Radio Network Controller (RNC)	A02, A05, A24, A28, A33, A40, A49
Smartphone	A05, A06, A10, A11, A13, A14, A24, A27, A28, A29, A30, A31, A33, A36, A38, A39, A40, A42, A44, A48, A49, A55, A56
Surveillance camera	A19, A20, A25, A26
On-board Devices (OBD)	A04, A07, A23, A41, A47, A52
Thermal camera	A12
Ticketing equipment	A17, A21, A43, A50, A51, A53, A54
Vehicle camera	A15
Vehicle counter	A46

Figure 10 presents a bar chart of sensor type usage, demonstrating a strong tendency towards the usage of smartphones and GPS data, with 23 and 18 studies, respectively. Camera were used by 9 studies, followed by Radio Network Controller (RNC) and Ticketing equipment, each used in 7 studies.

#### 4.8. SQ08 – Are the datasets used by the study publicly available?

As already explained in SQ01, one of the great challenges is the scarcity of databases to be explored by researchers, annotated or not. During the execution of the research the following databases were found.

- **PrivateCarTrajectoryData**:<sup>6</sup> Dataset containing trajectories of private vehicles [45,77].
- **Traffic Count Vehicle Classification 2014-2017**:<sup>7</sup> Dataset containing data from the survey carried out in the city of Melbourne – Australia, measuring vehicle traffic from 2014 to 2017 on the city’s roads [43].
- **GTI Data**:<sup>8</sup> Databases containing images to support research by the Image Treatment Group of the Polytechnic University of Madrid – Spain [56].
- **Humbi Behavioral Dataset**:<sup>9</sup> Open databases of behaviors and body expressions [7].
- **MOTChallenge**:<sup>10</sup> Database for creating and testing Multi-Object Tracking (MOT) algorithms [84].
- **MOT repository**:<sup>11</sup> Github repository containing datasets, algorithms, and other MOT-related content [84].
- **Crowd 11**:<sup>12</sup> Video dataset containing crowd movements for creating and testing Crowd Analysis algorithms [75].

<sup>6</sup><https://github.com/HunanUniversityZhuXiao/PrivateCarTrajectoryData> (access in June 2023).

<sup>7</sup><https://data.melbourne.vic.gov.au/explore/dataset/traffic-count-vehicle-classification-2014-2017/information/> (access in June 2023).

<sup>8</sup><http://www.gti.ssr.upm.es/data> (access in June 2023).

<sup>9</sup><https://humbi-data.net/> (access in June 2023).

<sup>10</sup><https://motchallenge.net/> (access in June 2023).

<sup>11</sup><https://github.com/luanshiyinyang/awesome-multiple-object-tracking> (access in June 2023).

<sup>12</sup><https://paperswithcode.com/dataset/crowd11> (access in June 2023).

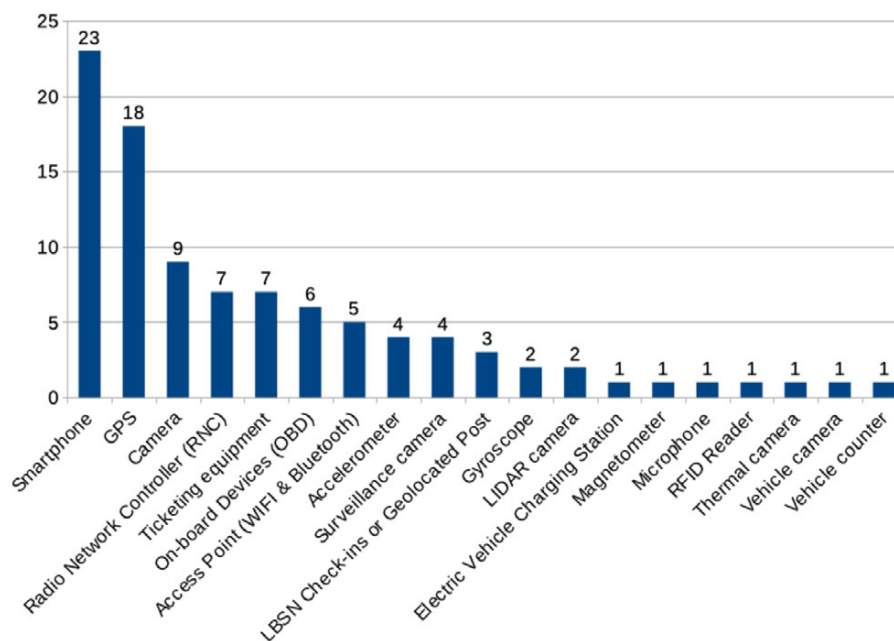


Fig. 10. Sensor type use count by article.

- **TownCentre dataset:**<sup>13</sup> Dataset containing video footage from security cameras in downtown Oxford – England [47].
- **Singapore Statistics:**<sup>14</sup> Singapore Census Data and General Household Survey [73].
- **National Household Travel Survey:**<sup>15</sup> Conducted by the Federal Highway Administration, the NHTS is the authoritative source on the travel behavior of the American public [65].
- **Monitoring Human Activity:**<sup>16</sup> A project of the Artificial Intelligence, Robotics and Vision Laboratory from University of Minnesota, Department of Computer Science and Engineering [64].
- **Mobike users Dataset:**<sup>17</sup> This dataset contains 102,361 trips made by 16,887 Mobike users on 79,062 bicycles in August 2016 [44].
- **Users’ check-ins data from New York City and Tokyo:** Users’ check-in data from Foursquare, collected from New York City and Tokyo from April 2012 to September 2013 [81,82] – used by [83].

#### 4.9. SQ09 – Are the source-codes produced based on the computational models proposed in the study publicly available?

The source-codes produced by the articles that are publicly available are listed below. Source-codes related to Multi-Object Tracking (MOT) was already mentioned in SQ08.

- **SeqScan-D**<sup>18</sup> [17]
- **Context Transition Predictability**<sup>19</sup> [83]

<sup>13</sup>[https://exposing.ai/oxford\\_town\\_centre/](https://exposing.ai/oxford_town_centre/) (access in June 2023).

<sup>14</sup><https://www.singstat.gov.sg/find-data/search-by-theme/population/mode-of-transport/publications-and-methodology> (access in June 2023).

<sup>15</sup><https://nhts.orl.gov/> – (access in June 2023).

<sup>16</sup><http://mha.cs.umn.edu/> – (access in June 2023).

<sup>17</sup>[https://figshare.com/articles/dataset/data\\_and\\_code\\_for\\_Mobike\\_v4\\_zip/11493420/1](https://figshare.com/articles/dataset/data_and_code_for_Mobike_v4_zip/11493420/1) – (access in June 2023).

<sup>18</sup><https://github.com/SeqScan/SeqScan-D> – (access in June 2023).

<sup>19</sup><https://github.com/zcfinal/ContextTransitionPredictability> – (access in June 2023).

Table 10  
Articles that uses context history and context prediction

Type	Articles
History	A01–A07, A09–A14, A16, A17, A18, A21–24, A27–31, A33, A35, A36, A41–47, A50–56
Prediction	A03, A04, A05, A07, A08, A16, A17, A21, A26, A30, A36, A46, A51, A54, A55

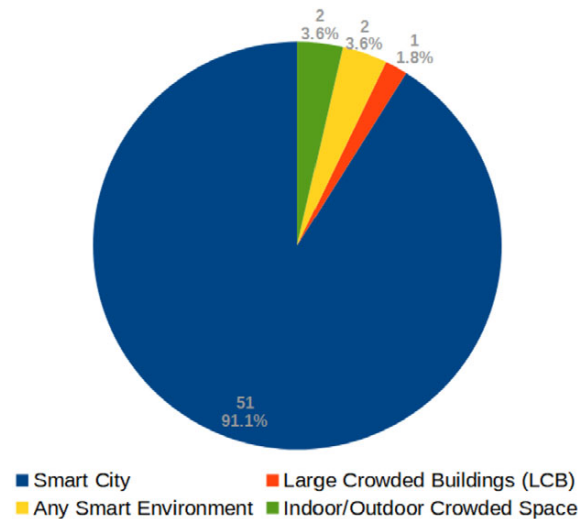


Fig. 11. Studies classified by Smart Environment.

#### 4.10. SQ10 – Does the study use context histories and context prediction?

Context can be understood as any specific information of some entity stored in a certain period of time, this entity can be anything, like a vehicle, a person, an object, a place etc [66]. This information can be its geolocation, its size, speed, temperature, color, weight, or other information about its state at the given moment [27]. Storing a collection of these contexts over time allows exploration of the past behavior of the entity being analyzed. In the literature, the name “Contexts History” is given to this collection [16].

Context histories allow the exploration of patterns in the past [23], similarity [26] and make predictions of future contexts [16]. Predictive models use past contexts combined with present context to make predictions of future contexts [16,48].

None of the articles found literally express in the text the use of Context History and Context Prediction, in the terms presented above. Despite this, as a ‘History’ criterion, articles using historical data, such as the geolocated trajectories of users/vehicles, were considered, and as a ‘Prediction’ criterion, articles that predicted future contexts. The result of this classification is shown in Table 10.

#### 4.11. SQ11 – In what kind of Smart Environment was the study carried out and/or validated?

Of all the studies reviewed, 51 (91.1%) were conducted and validated in the context of Smart Cities. Study A08 (1.8%) was conducted in Large Crowded Buildings (LCB), while A19 and A25 are related to any type of Smart Environment (3.6%). Studies A22 and A34 make reference to any Indoor/Outdoor Crowded Space (3.6%).

Figure 11 presents a chart categorizing the studies by their environment type. As the figure shows, mobility-related behavior is extensively studied within the context of Smart Cities.

Table 11  
Areas that the studies are applied

Area	Articles
<b>Crowd Analysis:</b> It seeks to identify behavior patterns of displacement of people/vehicles, groups or crowds, in the perimeter to which the study is being applied, carried out through cameras, LIDAR, Wifi Access Point, among others.	A01, A12, A16, A19, A20, A22, A24, A25, A26, A29, A31, A34, A44, A45, A55
<b>Bike-sharing:</b> Related to Bike-sharing services, with or without fixed stations.	A51, A52, A53
<b>Goods Mobility:</b> Seeks to explore specific solutions and problems in the mobility of cargo and consumer goods.	A23
<b>Mobility on Demand:</b> Seeks to promote means of transport according to demand patterns in certain regions at certain times, normally by crossing public or private, collective or individual transport data (e.g. taxis, app drivers, autonomous vehicles, etc.).	A21, A32
<b>POI Popularity:</b> ( <i>Point of Interest</i> ): Aims to identify patterns of presence and frequency of people and/or vehicles in certain places (e.g. restaurants, gas stations, parking lots, among others).	A06, A22, A36, A40, A44, A55
<b>Electric Vehicle Charging Stations Planning:</b> This area is self-explanatory on its behalf.	A30, A35
<b>Public Transport:</b> Articles that seek to carry out analyzes directly related to public and collective transport modes, and/or to promote insights for transport planning.	A04, A05, A11, A13, A17, A21, A28, A29, A32, A36, A40, A42, A43, A50, A51, A54
<b>Regions of Interest (ROI):</b> When the study seeks to identify, through its methods, the popularity and population density patterns in certain regions.	A16, A17, A21, A28, A33, A40, A44, A48, A49, A51, A55, A56
<b>Ridesharing:</b> Related to carpooling or ridesharing services.	A09, A47, A53, A56
<b>Risk and Emergency Management:</b> This item is self-explanatory on its behalf.	A08, A38
<b>Traffic Analysis:</b> It seeks to monitor and measure the number of vehicles and people moving in the regions to which the research perimeter is delimited, performing analyzes on the numbers found.	A18, A27, A28, A29, A32, A36, A37, A39, A40, A42, A43, A46, A47, A50, A52, A53, A54, A56
<b>Traffic Safety:</b> Researches that seek to identify anomalous behavior in driving vehicles, and/or seek to promote mechanisms to avoid collisions in traffic.	A10, A15, A27, A41
<b>Urban Mobility:</b> This is a broad area of research that encompasses several of the subareas already listed, such as analysis of modes of transport used by people, analysis of pedestrian and vehicle traffic, demand x supply of means of transport, among others.	A01, A03-07, A09, A11-18, A21, A23, A24, A27, A28, A29, A31, A32, A33, A39, A40, A42, A43, A44, A46-56

#### 4.12. SQ12 – In which areas are the studies being applied?

Table 11 lists the identified areas in which the studies have been applied. Many articles are associated with more than one area, demonstrating the interdisciplinary nature of the researches. For example, articles related to Crowd Analysis and Regions of Interest often overlap with Urban Mobility and Traffic Analysis.

Figure 12 presents the number of articles by area. The broad field of Urban Mobility has a significant number of articles related to mobility behavior, distributed across various subareas, such as Traffic Analysis, Public Transportation, Crowd Analysis, Regions of Interest, and others.

Electric Vehicle Charging Stations Planning is under-explored within the mobility behavior perspective. There is a need for more targeted research to address the unique challenges in this field, such sustainable infrastructure development, given the growing interest in the topic.

The use of technologies such as cameras, LIDAR, and GPS in areas like Crowd Analysis and Traffic Analysis indicates the integration of advanced technologies in research. Future discussions could explore how emerging technologies like AI (Artificial Intelligence) and IoT can further enhance these studies.

Research in areas like Public Transport, Risk and Emergency Management, and Traffic Safety has direct implications for public policy and urban planning. Modern mobility solutions like Bike-sharing, Ridesharing and Mobility-on-Demand are also being studied within the mobility behavior perspective, with 3, 4 and 2 articles respectively.



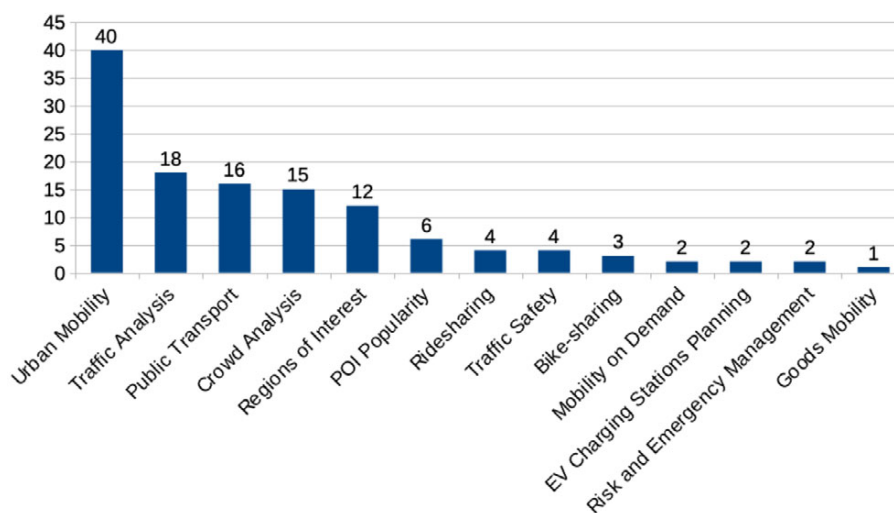


Fig. 12. Area of application by studies.

#### 4.13. SQ13 – What behaviors are being observed?

Table 12 lists the behaviors that are being observed by the studies. The meaning of each behavior is also described.

The substantial number of studies related to Movement Patterns (29), as shown in Fig. 13, indicates a significant interest in understanding the movement of people and vehicles across different environments. Additionally, the studies related to Transport Modal (14), Semantic Travel Intention (2), Presence, Stay-time or Arrive-Stay-Leave (15), and Boarding and Disembarking Patterns (7) demonstrate a growing interest in comprehending these behaviors, which assists in planning and guiding public policies related to urban mobility.

Analyzing the mobility behaviors of electric vehicles and measuring the demand for recharging represents an important research area, given its potential to contribute to the development of sustainable urban mobility solutions.

Many behaviors overlap with multiple articles, demonstrating the interdisciplinary nature of research in this domain. For example, Transport Modal and Movement Patterns often intersect with other behaviors like Presence, Stay-time or Arrive-Stay-Leave.

#### 4.14. SQ14 – Are behaviors of a single entity (e.g. person), groups and/or crowds being observed?

Studies that allow the identification of a specific entity, and then extrapolate this analysis to the collective (micro to macro) were identified in Table 13 as observing single and multiple entities. Studies that conducted analyzes of specific groups of entities within the observed population were categorized accordingly. Studies that exclusively analyzed the collective, without the ability to identify individual entities, were designated as only observing crowd behavior.

#### 4.15. SQ15 – Are ontologies being used? In which scenarios are ontologies being used to identify and predict behaviors?

Studies of [9] and [4] used data collected through an API called Sii-Mobility [52], built based on the KM4City [8] ontology. These three articles share common authors, suggesting that they were likely conducted by related research group. Additional information about this ontology can also be found in a specific disclosure link.<sup>20</sup>

<sup>20</sup><https://www.km4city.org/> – Link with KM4City ontology information (Access in June 2023).

Table 12  
Behaviors observed by the studies

Behavior	Articles
<b>Abnormal Crowd Behavior:</b> Unusual, emergency or security related crowd behaviors.	A34, A38
<b>Driving Behavior:</b> Act of driving, including aggressive driving or driving violations.	A10, A15, A27, A41
<b>Boarding and disembarking patterns:</b> Identification of locations where passengers normally board and disembark from transport vehicles.	A17, A21, A43, A50, A51, A53, A54
<b>Building evacuation:</b> The act of an entity fleeing or moving out of a building under safety risk (e.g. football stadium, theater)	A08
<b>Indoor travel between points:</b> Refers to displacement within closed environments, not easily monitored by GPS technologies existing in smartphones nowadays.	A22, A34, A45
<b>Presence, Stay-time or Arrive-Stay-Leave:</b> The act of being at a certain point or region at a certain time, staying and/or leaving.	A01, A02, A03, A04, A05, A07, A16, A23, A24, A28, A36, A45, A46, A48, A49
<b>Semantic travel intention:</b> It is related to the purpose of the travel behavior, for example: The person left home to go to the market, probably with the purpose of shopping, among others.	A05, A06
<b>Transport Modal:</b> Defined by the modals of transport that the person used to travel (e.g. by foot, car, bicycle, bus, etc.).	A05, A11, A13, A14, A28, A29, A32, A36, A37, A40, A42, A50, A51, A53
<b>Movement Patterns:</b> Determined by the two or more spatio-temporal location of an entity (e.g. person, vehicle), not necessarily being this point the origin or destination of this entity.	A01, A03, A05, A06, A07, A09, A12, A13, A14, A16, A17, A19, A20, A21, A23, A28, A29, A31, A33, A36, A43, A44, A45, A46, A47, A52, A54, A55
<b>Vehicle Recharging:</b> Behavior of charging an electric vehicle.	A30, A35

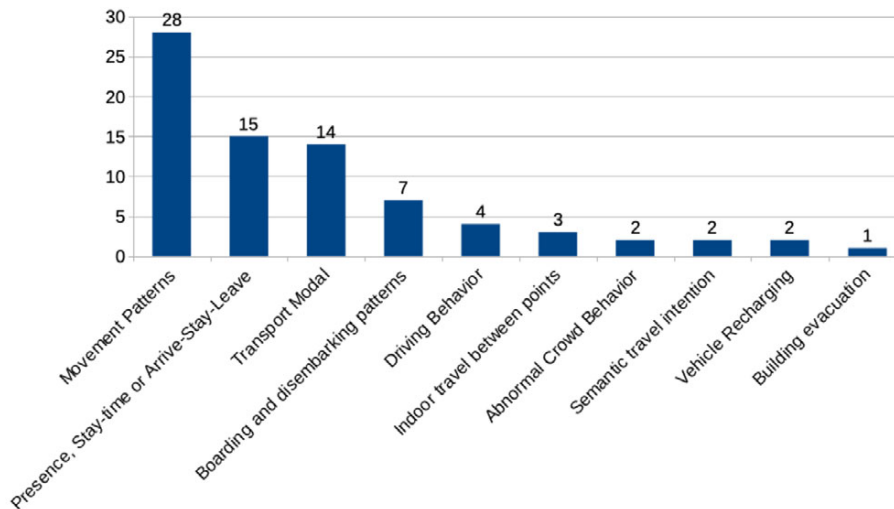


Fig. 13. Number of studies by observed behavior.

Table 13  
Articles that uses context history and context prediction

Type	Articles
Single	A07, A08, A09, A10, A12, A16, A17, A21, A22, A24, A29, A34, A35, A36, A41, A42, A43
Groups	A19, A20, A37, A43
Multiple	A01, A02, A03, A04, A05, A06, A08, A09, A11, A12, A13, A14, A15, A16, A17, A18, A19, A20, A21, A22, A23, A24, A25, A26, A27, A28, A29, A30, A31, A32, A33, A34, A35, A38, A39, A40, A42, A43, A44, A45, A46, A47, A48, A49, A50, A51, A52, A53, A54, A55, A56

The study from [70] utilized the Ontology FIESTA-IOT with the specific M3-lite taxonomy [1]. The objective was to provide seamless interoperability and information transparency from IoT systems to crowd management applications. These two articles also share common authors, suggesting that they were conducted by related research group.

#### 4.16. SQ16 – What are the techniques used to validate the studies?

Various methodologies were identified and applied in the validation of the studies. The following list provides an overview of these techniques:

- Studies that conducted their research through the application of computational models in datasets that the researchers had access, in a general view, validated their research in two ways:
  1. The results of the computational models proposed by the research were compared with results of related computational models, as explored by [45]. The study proposal was compared with 9 other computational models.
  2. Some of the studies based on Machine Learning models used part of the dataset for training, leaving another part for validating the results:
    - \* [34] used 90% of the dataset for training, and 10% for validation.
    - \* [45] used 70% for training, 10% for validation, and 20% for testing.
    - \* [51] used 80% for training and 20% for validation.
    - \* [63] used 75% of the dataset for training and 25% for validation.
- The studies that used transaction data from smart cards in public transport [50,65] validated their boarding and disembarking prediction algorithms based on data from subsequent days generated by users.
- Computational models that observe the behavior of aggressive driving [67,68], using the accelerometer of the smartphone, validated the research in real time with tests in vehicles monitored by the researchers.
- The research that used computer simulation [85], analyzed the simulation data itself to validate the models.
- One study had access to multiple datasets [56], and used a dataset for training and another for validation.

#### 4.17. STQ01 – Where the studies are being published?

Figure 14 illustrates the number of publications by database and year. In ScienceDirect, 12 publications related to mobility and behavior were selected. In the ACM, 3 publications on the subject were found. IEEE Xplore represented the majority of the publications, with 23 articles selected. In Springer Link, 3 articles were selected, and one from Wiley Online Library.

The Scopus database indexes articles from other databases, potentially leading to duplicate articles. In Scopus, 14 articles were found (out of a total of 28 articles), and 14 articles were identified as duplicates. However, these duplicates were attributed to their original ScienceDirect, ACM, IEEE Xplore, and Springer Link databases.

#### 4.18. STQ02 – How many publications occurred per year?

Figure 14 also illustrates the number of publications per year. The initial publications related to mobility and behavior emerged in 2014, showing fluctuations and a general upward trend in subsequent years. In 2017, a total of 9 articles were identified. The peak of articles occurred in 2020 with 13 publications. In 2021, a total of 6 articles were found. This review collected articles until April/2022, with 7 additional articles identified during that period.

## 5. Discussion

This section explores and discusses the results presented in the previous section. Firstly, the central question of the review is addressed. Subsequently, a discussion on challenges and open research questions is provided.

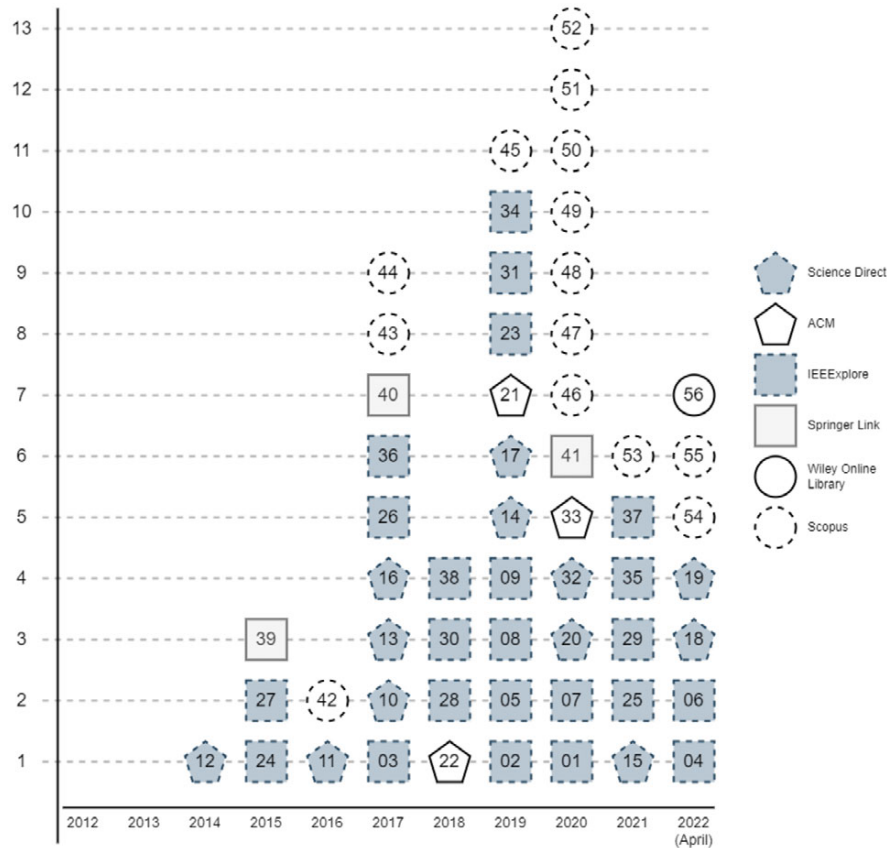


Fig. 14. Publications per year and databases.

### 5.1. How computational techniques are being used to analyse human mobility behaviors in Smart Cities and Smart Environments?

The analysis of mobility behaviors through computational techniques finds applications in various domains and provides valuable insights for urban planning, transportation services, and a deeper understanding of human interactions in Smart Environments. A significant number of studies have focused on examining how different transportation modes are being utilized, 14 in total. Research efforts include the analysis of incentives to use environmentally sustainable modes of transportation, such as bike-sharing and others, resulting in benefits for individuals (health) as well for the population, with the reduction of CO<sub>2</sub> emissions and improvements in traffic flow [62,63].

The identification of origin-destination patterns, through passenger boarding and disembarking data, provides valuable information for public transportation demand planning, achieving accuracy rates of up to 85% in predicting this behaviors on specific regions [46,50]. With known origin-destination patterns, the implementation of mobility-on-demand services are facilitated, enabling the deployment of other last-mile transport modals near mass transit boarding and disembarking stations (such as trains and buses) [65]. Taxis and ridesharing services are crucial modes for providing this last-mile transportation. Analyzing the mobility patterns of these vehicles offers even more granular and specific insights into the reality of urban mobility [33,39,78].

The analysis of trajectories from small private vehicles provides crucial information about traffic patterns and congestion in large economic districts [77]. Specifically, a significant subset of articles, totaling 18, focuses on traffic behavior analysis. The techniques employed to achieve these objectives are diverse. A private vehicle datasets was utilized in [45,77] to cluster urban regions, evaluating Arrive-Stay-Leave patterns in these areas. Wifi routers was used in [25] and [9] to collect people's crossing information by capturing the MAC addresses of smartphones passing

within the routers' range. With a similar purpose, [71] employed vehicle counters and vehicle type identifiers for traffic analysis.

Eight studies explored algorithms applied into Telecom databases, with various focuses such as the analysis of Points and Regions of Interest (POIs and ROIs), and the mobility of users between these regions. Accessing this type of dataset involves challenges and resources, as some Telecom companies commercialize this access.<sup>21</sup> Another strategy explored is the cross-referencing of datasets from surveys and public research. Telecom data was used by [73], cross-referenced with public traffic and pollution survey data to identify origin-destination matrices. NHTS survey data was utilized in [33] for dynamic identification of transport modes used, with a focus on mobility-on-demand.

Individual behavior analysis, specifically driving behavior, was explored by 4 articles. Data collected from the smartphone's accelerometer and GPS was used by [68] and [67] to identify driving behaviors and traffic situations. Videos from vehicular cameras and image processing was used in [56] to identify traffic violations and illicit maneuvers performed by other drivers. A test vehicle with various sensors installed was used by [5] to collect comprehensive information about driving behavior, including pedals pressure, gear changes, brakes, and accelerations.

A total of 15 articles explored crowd mobility behaviors (crowd analysis). The use of cameras and object tracking processing algorithms is one of the most commonly employed techniques [84]. A thermal camera was used in [53] for this monitoring, thus preserving users' privacy. The use of LIDAR in indoor movement tracking was explored by [80]. In addition to LIDAR, the authors used the accelerometer and gyroscope of the mobile phone to monitor this movement, combined with the use of the microphone to identify the volume of conversations around the person, thus estimating occupancy and visitation of indoor locations.

Related to Machine Learning techniques (ML), Fig. 7 illustrates the temporal evolution of ML usage. The data show an initial increase in use beginning in 2015 and continuing until 2017, followed by a subsequent increase from 2019 to 2022, indicating a trend of growing interest.

About the data-sources used, Fig. 9 demonstrates the predominant interest in utilizing datasets of GPS Travel History from various segments in research, with 16 articles. Nonetheless, the importance of public datasets and open data policies is also evident, as identified by the use of Government or Public Data in 12 articles. The same rationale can be applied to "Usage / Ticketing data" databases, used by 11 articles, and mobility databases based on Telecom data, utilized in 8 studies. Together, these four groups represent 35 articles (some articles use more than one type of data-source), highlighting the importance of strengthening partnerships with industry and public open data policies to support scientific research.

The majority of studies related to mobility behavior are extensively realized within the context of Smart Cities, as displayed on Fig. 11. Another observation is the low number of studies that shared the source code produced in their research. Only [17] and [83] made their model source code available in online repositories, a practice that seems to be more common in the field of Multi-Object Tracking [84].

About what behavior was observed, the substantial number of studies related to Movement Patterns (29), as shown in Fig. 13, indicates a significant interest in understanding the movement of people and vehicles across different environments. Additionally, the studies related to Transport Modal (14), Semantic Travel Intention (2), Presence, Stay-time or Arrive-Stay-Leave (15), and Boarding and Disembarking Patterns (7) demonstrate a growing interest in comprehending these behaviors, which assist in planning and guiding public policies related to urban mobility.

The strategies for analyzing behavior, in general, are directly related to the datasets used in the research. Figure 8 provides an overview of the types of datasets explored, and Fig. 9 presents usage count of each type. The complete list of techniques employed are listed in Section 4 (Results and answers – SQ02, in Table 5).

During the review, we noticed the limited use of ontologies. The studies [9] and [4] utilized the KM4City and Sii-Mobility API ontologies, while [70] employed the FIESTA-IOT ontology. In both cases, the ontologies shared common authors.

## 5.2. Challenges and open issues

This section aims to discuss the open issues and research opportunities, especially related to SQ01 (Section 4.1).

---

<sup>21</sup><https://www.vivo.com.br/para-empresas/produtos-e-servicos/digitais/big-data/smartsteps> – Access in November 2023.

- **Behavior identification and modeling:** The analysis of human mobility behaviors is a central focus of this systematic review. No universally applicable or generic method has been identified as a reference for all other studies. The methods and strategies are diverse, as detailed in SQ02 and throughout the Results (Section 4). The range of applicable strategies for behavior analysis is directly related to the datasets available.
- **Issues related to large volumes of data – Big Data:** Regarding data volume, the application of algorithms and Machine Learning techniques for selecting image frames or specific intervals through edge computing can reduce the use of computational resources. Studies focused on filtering and sending reduced data segments for cloud processing were little explored. This issue was explored only by [75]. The goal of this approach is to reduce the consumption of these computational resources, such as data traffic, processing and storage.
- **Real-time image processing demand:** Algorithms and ML for object identification in images executed directly on capture devices can also reduce the consumption of computational resources on cloud data traffic. An example of a study in this area is YOLO (You Only Look Once) [57,76], used by [71] and [84].
- **Imprecise or incomplete data:** Challenges related to the imprecision and correction of spatio-temporal data, such as those generated by Telecom or imprecise GPS sensors, are issues to be explored in future research. The study conducted by [44] addresses this problem, presenting results with and without correction algorithms, achieved through the intersection of street datasets. Sparse or incomplete data also represent significant challenges to be explored. GPS devices may experience moments of loss of geolocation or connection signal, creating intervals of missing data. Users travel history was explored by [46] to compose trajectory patterns, given the absence of this information in the research dataset.
- **Few mobility datasets available or datasets not annotated:** The absence of anonymized, large-scale datasets, as previously mentioned in SQ01, poses a significant research challenge. Future work could focus on creating and providing annotated mobility datasets to support further research. This issue becomes evident when observing studies that used datasets generated with more than 10 years of difference between the study and the dataset creation [34,47]. As previously mentioned, DiDi, one of the largest ridesharing companies, founded the Gaia Initiative to provide anonymized mobility datasets for research purposes.
- **Lack of tools for dataset annotation:** The lack of dedicated research in creating data annotation tools was related by the review conducted by [10] and is also identified in this review. The research field has a high demand for annotated datasets, but the development of specific tools for this purpose has been overlooked by researchers. The use of crowdsourcing for generating these datasets is a promising solution, as explored by [62] and [63]. However, this approach may lead to biased results depending on how the data is annotated and collected by users. As reported by the researchers and also by [35], solutions that encourage the use of sustainable transportation methods may result in inaccurate or missing data, as users may hesitate to report the use of polluting modes of transportation.

### 5.3. Limitations

This review has limitations that may affect the scope of the results. We reviewed six databases to reduce bias, and defined the search string considering major terms and the broader spectrum of computational techniques to analyse behavior. The search process requires a combination of string parts that form the search string, which may limit the studies selection, since related studies can address behavior, Smart Cities and Smart Environments, without specific terms, such as “smart spaces” or “behavioral”. This review considers studies published only in the English language and thus may exclude relevant articles published in other languages. Finally, the search engine algorithms from each database may influence the results, as some databases could implement synonyms automatically or other search strategies.

## 6. Conclusion

This study presented a Systematic Literature Review (SLR) with the following central question: How computational techniques are being used to analyse human mobility behaviors in Smart Cities and Smart Environments? A total of 5989 articles were initially found, filtered, resulting in 56 articles reviewed. As the main contribution,

this study provides responses to 19 research questions, starting from an exploration of challenges, to listing computational techniques utilized. The algorithms, machine learning techniques and data-sources used by the reviewed studies are also listed and presented through taxonomies. A list of sensors/devices employed on the reviewed studies, and the data-types utilized are also presented. Finally, a comprehensive discussion of the identified techniques is conducted, finishing with a compilation of challenges, open issues and research opportunities.

The analysis of mobility behaviors through computational techniques finds applications in various domains and provides valuable insights for urban planning, transportation services, and a deeper understanding of human interactions in Smart Environments. A significant number of studies have focused on examining how different transportation modes are being utilized. Research efforts include the analysis of incentives to use environmentally sustainable modes of transportation, resulting in benefits for individuals as well for the population. The identification of origin-destination patterns through passenger boarding and disembarking data provides valuable information for public transportation demand planning, achieving accuracy rates of up to 85% in predicting this behaviors on specific regions [46,50].

The analysis of taxis and ridesharing mobility patterns offers granular and specific insights into the reality of urban mobility [33,39,78]. Trajectories from small private vehicles also provides crucial information about traffic patterns and congestion in large economic districts [77], with a significant number of studies in this area.

The exploration of Telecom databases affords important mobility insights, allowing the identification of Points and Regions of Interest (POIs and ROIs), and the mobility of users between these regions, specially when cross-referenced with additional geospatial information [73]. Smartphone's GPS and accelerometer also allows the a analysis of individual's driving behavior [67,68]. The use of cameras and object tracking processing algorithms is one of the most commonly employed techniques for crowd mobility behavior analysis [84], with some studies using thermal cameras [53] and LIDAR sensors [80] to indoor and outdoor applications. These large datasets offer a rich foundation for the exploration of patterns in the data through Machine Learning techniques, as identified in this study.

### 6.1. Future work

Future reviews may extend this study by considering deeper behavior-specific research by application area. New studies can explore the challenges and open issues listed in Section 5.2. Strategies to analyse human mobility behavior are directly related to the datasets available. Research efforts can be focused on generating annotated mobility datasets, development of annotation tools, behavior analysis, Machine Learning techniques and big-data computational challenges.

## Acknowledgements

The authors would like to thank CNPq (National Council for Scientific and Technological Development) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. We would also like to thank the University of Vale do Rio dos Sinos (Unisinos) for supporting the development of the present study.

## Funding

This work was supported by CAPES [001] and CNPQ [307137/2022-8].

## Conflict of interest

The authors have no competing interests to declare.

## Appendix. Selected articles

Table 14  
Review selected articles

Code	Title	Citation
A01	Detection and Analysis of Anomalies in People Density and Mobility Through Wireless Smartphone Tracking	[25]
A02	RevUrb: Understanding Urban Activity and Human Dynamics through Point Process Modelling of Telecoms Data	[22]
A03	Matrix factorization on semantic trajectories for predicting future semantic locations	[34]
A04	Exploiting Spatiotemporal Correlations of Arrive-Stay-Leave Behaviors for Private Car Flow Prediction	[45]
A05	Research on User Behavioral Intention Based on Telecommunication Data	[32]
A06	Towards the Inference of Travel Purpose with Heterogeneous Urban Data	[51]
A07	On Extracting Regular Travel Behavior of Private Cars Based on Trajectory Data Analysis	[77]
A08	An Analysis of Pedestrians' Behavior in Emergency Evacuation Using Cellular Automata Simulation	[85]
A09	A Carpooling Service for Private Vehicles Using Electronic Registration Identification Data	[15]
A10	A smartphone based technique to monitor driving behavior using DTW and crowdsensing	[68]
A11	Crowdsourcing mobility insights – Reflection of attitude based segments on high resolution mobility behaviour data	[62]
A12	Taking the Temperature of Pedestrian Movement in Public Spaces	[53]
A13	Spatial context mining approach for transport mode recognition from mobile sensed big data	[63]
A14	Big-data analysis for carbon emission reduction from cars: Towards walkable green smart community	[79]
A15	Smart traffic control: Identifying driving-violations using fog devices with vehicular cameras in smart cities	[56]
A16	Wi-Fi based city users' behaviour analysis for smart city	[9]
A17	Domain-specific data mining for residents' transit pattern retrieval from incomplete information	[46]
A18	Weighted dynamic time warping for traffic flow clustering	[43]
A19	Multi-object tracking based on attention networks for Smart City system	[84]
A20	Deep learning for pedestrian collective behavior analysis in smart cities: A model of group trajectory outlier detection	[7]
A21	BuScope: Fusing Individual and Aggregated Mobility Behavior for "Live" Smart City Services	[50]
A22	A Human Tracking and Sensing Platform for Enabling Smart City Applications	[80]
A23	Monitoring Urban-Freight Transport Based on GPS Trajectories of Heavy-Goods Vehicles	[30]
A24	Estimation of mobility direction of a people flux by using a live 3G radio access network and smartphones in non-connected mode	[59]
A25	Visual Attention based Cognitive Informative Frame Extraction Method for Smart Crowd Surveillance	[75]
A26	Forecasting Interactive Dynamics of Pedestrians with Fictitious Play	[47]
A27	Traffic expression through ubiquitous and pervasive sensorization: Smart cities and assessment of driving behaviour	[67]
A28	A Comparative Study of Urban Mobility Patterns Using Large-Scale Spatio-Temporal Data	[73]
A29	A Gamification Platform to Analyze and Influence Citizens' Daily Transportation Choices	[35]
A30	Planning Electric Vehicle Charging Stations Based on User Charging Behavior	[42]
A31	CityTracker: Citywide Individual and Crowd Trajectory Analysis Using Hidden Markov Model	[24]
A32	Dynamics of travelers' modality style in the presence of mobility-on-demand services	[65]
A33	On Location Relevance and Diversity in Human Mobility Data	[17]
A34	Crowd Detection Management System	[64]
A35	Electric Vehicles Trips and Charging Simulator Considering the User Behaviour in a Smart City	[69]
A36	User Engagement Engine for Smart City Strategies	[4]
A37	Smart Cities Mobility Monitoring through Automatic License Plate Recognition and Vehicle Discrimination	[71]
A38	System for Detection of Emergency Situations in Smart City Environments Employing Smartphones	[29]
A39	Personalized routing for multitudes in smart cities	[18]
A40	Cell phone big data to compute mobility scenarios for future smart cities	[74]
A41	Extraction of Naturalistic Driving Patterns with Geographic Information Systems	[5]
A42	Policy 2.0 platform for mobile sensing and incentivized targeted shifts in mobility behavior	[61]
A43	Analyzing year-to-year changes in public transport passenger behaviour using smart card data	[11]



Table 14  
(Continued)

Code	Title	Citation
A44	RCMC: Recognizing crowd-mobility patterns in cities based on location based social networks data	[2]
A45	Toward understanding crowd mobility in smart cities through the internet of things	[70]
A46	Probabilistic modelling for unsupervised analysis of human behaviour in smart cities	[54]
A47	TBI2Flow: Travel behavioral inertia based long-term taxi passenger flow prediction	[39]
A48	Role of big data in the development of smart city by analyzing the density of residents in shanghai	[31]
A49	Utilizing digital traces of mobile phones for understanding social dynamics in urban areas	[86]
A50	Optimal planning of real-time bus information system for user-switching behavior	[40]
A51	Inferring temporal motifs for travel pattern analysis using large scale smart card data	[41]
A52	Understanding intra-urban human mobility through an exploratory spatiotemporal analysis of bike-sharing trajectories	[44]
A53	Dynamic Study of Intelligent Traffic Behaviour Based on Multiple Traffic Modes	[33]
A54	Exploring Human Mobility for Multi-Pattern Passenger Prediction: A Graph Learning Framework	[38]
A55	Beyond The Limits of Predictability in Human Mobility Prediction: Context-Transition Predictability	[83]
A56	A classification method for urban functional regions based on the transfer rate of empty cars	[78]

## References

- [1] R. Agarwal, D.G. Fernandez, T. Elsaiah, A. Gyrard, J. Lanza, L. Sanchez, N. Georgantas and V. Issarny, Unified iot ontology to enable interoperability and federation of testbeds, in: *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*, 2016, pp. 70–75. doi:[10.1109/WF-IoT.2016.7845470](https://doi.org/10.1109/WF-IoT.2016.7845470).
- [2] H. Assem, T.S. Buda and D. O’Sullivan, Rcmc: Recognizing crowd-mobility patterns in cities based on location based social networks data, *ACM Transactions on Intelligent Systems and Technology* (2017).
- [3] J.C. Augusto, *Smart Cities: State of the Art and Future Challenges*, Springer International Publishing, Cham, 2020, pp. 1–12.
- [4] C. Badii, P. Bellini, D. Cenni, A. Difino, M. Paolucci and P. Nesi, User engagement engine for smart city strategies, in: *2017 IEEE International Conference on Smart Computing (SMARTCOMP)*, 2017, pp. 1–7.
- [5] J. Balsa-Barreiro, P.M. Valero-Mora, M. Menéndez and R. Mehmood, *Extraction of Naturalistic Driving Patterns with Geographic Information Systems. Mobile Networks and Applications*, 2020.
- [6] R. Bavaresco, Y. Ren, J. Barbosa and G.P. Li, An ontology-based framework for workers health reasoning enabled by machine learning, *Computers & Industrial Engineering* **193** (2024), 110310. doi:[10.1016/j.cie.2024.110310](https://doi.org/10.1016/j.cie.2024.110310).
- [7] A. Belhadi, Y. Djenouri, G. Srivastava, D. Djenouri, J.C.-W. Lin and G. Fortino, Deep learning for pedestrian collective behavior analysis in smart cities: A model of group trajectory outlier detection, *Information Fusion* **65** (2021), 13–20. doi:[10.1016/j.inffus.2020.08.003](https://doi.org/10.1016/j.inffus.2020.08.003).
- [8] P. Bellini, M. Benigni, R. Billero, P. Nesi and N. Rauch, Km4city ontology building vs data harvesting and cleaning for smart-city services, *Journal of Visual Languages & Computing* **25**(6) (2014), 827–839.
- [9] P. Bellini, D. Cenni, P. Nesi and I. Paoli, Wi-fi based city users’ behaviour analysis for smart city, *Journal of Visual Languages & Computing* **42** (2017), 31–45. doi:[10.1016/j.jvlc.2017.08.005](https://doi.org/10.1016/j.jvlc.2017.08.005).
- [10] M. Bendali-Braham, J. Weber, G. Forestier, L. Idoumghar and P.-A. Muller, Recent trends in crowd analysis: A review, *Machine Learning with Applications* **4** (2021), 100023. doi:[10.1016/j.mlwa.2021.100023](https://doi.org/10.1016/j.mlwa.2021.100023).
- [11] A.-S. Briand, E. Côme, M. Trépanier and L. Oukhellou, Analyzing year-to-year changes in public transport passenger behaviour using smart card data. *Transportation Research Part C: Emerging Technologies*, 2017.
- [12] P.C. Büttgenbender, E.G. de Azevedo Neto, W.F. Heckler and J.L.V. Barbosa, A computational model for identifying behavioral patterns in people with neuropsychiatric disorders, *IEEE Latin America Transactions* **20**(4) (2022), 582–589. doi:[10.1109/TLA.2022.9675463](https://doi.org/10.1109/TLA.2022.9675463).
- [13] P.S. Castro, D. Zhang, C. Chen, S. Li and G. Pan, From taxi gps traces to social and community dynamics: A survey, *ACM Comput. Surv.* **46**(2) (2013). doi:[10.1145/2543581.2543584](https://doi.org/10.1145/2543581.2543584).
- [14] C. Costa and D. Zeinalipour-Yazti, Telco big data research and open problems, in: *2019 IEEE 35th International Conference on Data Engineering (ICDE)*, 2019, pp. 2056–2059. doi:[10.1109/ICDE.2019.00238](https://doi.org/10.1109/ICDE.2019.00238).
- [15] C. Cui, L. Zheng, D. Xia, L. Chen and D. Sun, A carpooling service for private vehicles using electronic registration identification data, in: *2019 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)*, 2019, pp. 1046–1054.
- [16] J.H. Da Rosa, J.L.V. Barbosa and G.D. Ribeiro, Oracon: An adaptive model for context prediction, *Expert Systems with Applications* **45** (2016), 56–70. doi:[10.1016/j.eswa.2015.09.016](https://doi.org/10.1016/j.eswa.2015.09.016).
- [17] M.L. Damiani, F. Hachem, C. Quadri, M. Rossini and S. Gaito, On location relevance and diversity in human mobility data, *ACM Trans. Spatial Algorithms Syst.* **7**(2) (2020).
- [18] M. De Domenico, A. Lima, M.C. González and A. Arenas, Personalized routing for multitudes in smart cities, *EPJ Data Science* **4**(1) (2015), 1. doi:[10.1140/epjds/s13688-015-0038-0](https://doi.org/10.1140/epjds/s13688-015-0038-0).

- [19] L.R. de Souza, R. Francisco, J.E. da Rosa Tavares and J.L.V. Barbosa, Intelligent environments and assistive technologies for assisting visually impaired people: a systematic literature review, *Universal Access in the Information Society* (2024).
- [20] G.C.L. de Souza Lima, R. Schechtman, L.C. Brizon and M.Z. Figueiredo, Transporte público e covid-19 – o que pode ser feito? in: *FGV CERI*, Rio de Janeiro, RJ, Brasil, 2020, Centro de Estudos em Regulação e Infraestrutura da Fundação, Getúlio Vargas (FGV CERI).
- [21] L.P.S. Dias, H.D. Vianna and J.L.V. Barbosa, Human behaviour data analysis and noncommunicable diseases: A systematic mapping study, *Behaviour & Information Technology* **42**(14) (2023), 2485–2503. doi:[10.1080/0144929X.2022.2128422](https://doi.org/10.1080/0144929X.2022.2128422).
- [22] R. Doorley, A. Noyman, Z. Xiong, L. Alonso, A. Grignard and K.L. Revurb, Understanding urban activity and human dynamics through point process modelling of telecoms data, in: *2019 Smart City Symposium Prague (SCSP)*, 2019, pp. 1–6.
- [23] D. Dupont, J.L.V. Barbosa and B.M. Alves, Chspam: A multi-domain model for sequential pattern discovery and monitoring in contexts histories, *Pattern Analysis and Applications* **23**(2) (2020), 725–734. doi:[10.1007/s10044-019-00829-9](https://doi.org/10.1007/s10044-019-00829-9).
- [24] S.-H. Fang, L. Lin, Y.-T. Yang, X. Yu and Z. Xu, Citytracker: Citywide individual and crowd trajectory analysis using hidden Markov model, *IEEE Sensors Journal* **19**(17) (2019), 7693–7701. doi:[10.1109/JSEN.2019.2916693](https://doi.org/10.1109/JSEN.2019.2916693).
- [25] A. Fernández-Ares, P. García-Sánchez, M.G. Arenas, A.M. Mora and P.A. Castillo-Valdivieso, Detection and analysis of anomalies in people density and mobility through wireless smartphone tracking, *IEEE Access* **8** (2020), 54237–54253. doi:[10.1109/ACCESS.2020.2979367](https://doi.org/10.1109/ACCESS.2020.2979367).
- [26] A.S. Filippetto, R. Lima and J.L.V. Barbosa, A risk prediction model for software project management based on similarity analysis of context histories, *Information and Software Technology* **131** (2021), 106497. doi:[10.1016/j.infsof.2020.106497](https://doi.org/10.1016/j.infsof.2020.106497).
- [27] P.S. Gandodhar and S.M. Chaware, Context aware computing systems: A survey, in: *2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics, 2018, and Cloud (I-SMAC)I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*.
- [28] J. Gao, D. Wang, C.-P. Lin, C. Luo, Y. Ruan and M. Yuan, Detecting and learning city intersection traffic contexts for autonomous vehicles, *Journal of Smart Cities and Society* **1** (2022). doi:[10.3233/SCS-220010](https://doi.org/10.3233/SCS-220010).
- [29] L. García, L. Parra, M. Taha and J. Lloret, System for detection of emergency situations in smart city environments employing smartphones, in: *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2018, pp. 266–272. doi:[10.1109/ICACCI.2018.8554654](https://doi.org/10.1109/ICACCI.2018.8554654).
- [30] S. Hadavi, S. Verlinde, W. Verbeke, C. Macharis and T. Guns, Monitoring urban-freight transport based on gps trajectories of heavy-goods vehicles, *IEEE Transactions on Intelligent Transportation Systems* **20**(10) (2019), 3747–3758. doi:[10.1109/TITS.2018.2880949](https://doi.org/10.1109/TITS.2018.2880949).
- [31] S.A. Haidery, H. Ullah, N. Ullah Khan, K. Fatima, S.S. Rizvi and S.J. Kwon, Role of big data in the development of smart city by analyzing the density of residents in Shanghai, *Electronics (Switzerland)* (2020).
- [32] Y. Han, D. Fu, L. Zhao, C. Cheng, X. Cheng and L. Xu, Research on user behavioral intention based on telecommunication data, in: *2019 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)*, 2019, pp. 1499–1504.
- [33] H. Jing, X. Li, G. Xu, M. Zhu, L. Shen, F. Liu and H. Peng, Dynamic study of intelligent traffic behaviour based on multiple traffic modes, *Scientific Programming* (2021).
- [34] A. Karatzoglou, S.C. Lamp and M. Beigl, Matrix factorization on semantic trajectories for predicting future semantic locations, in: *2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2017, pp. 1–7.
- [35] R. Kazhiamiakin, E. Loria, A. Marconi and M. Scanagatta, A gamification platform to analyze and influence citizens' daily transportation choices, *IEEE Transactions on Intelligent Transportation Systems* **22**(4) (2021), 2153–2167. doi:[10.1109/TITS.2021.3049792](https://doi.org/10.1109/TITS.2021.3049792).
- [36] P. Kisters, H. Schreiber and J. Edinger, Categorization of crowd-sensing streaming data for contextual characteristic detection, *Journal of Smart Cities and Society* **2** (2023).
- [37] B. Kitchenham and S. Charters, Guidelines for performing systematic literature reviews in software engineering, in Technical Report EBSE 2007-001, Keele University and Durham University Joint Report, 2007.
- [38] X. Kong, K. Wang, M. Hou, F. Xia, G. Karmakar and J. Li, Exploring human mobility for multi-pattern passenger prediction: A graph learning framework, *IEEE Transactions on Intelligent Transportation Systems* (2022).
- [39] X. Kong, F. Xia, Z. Fu, X. Yan, A. Tolba and Z. Almahadmeh, Tbi2flow: Travel behavioral inertia based long-term taxi passenger flow prediction, *World Wide Web* **23**(2) (2020), 1381–1405. doi:[10.1007/s11280-019-00700-1](https://doi.org/10.1007/s11280-019-00700-1).
- [40] Z. Lee, S. Hwang and J. Kim, Optimal planning of real-time bus information system for user-switching behavior, *Electronics (Switzerland)* (2020).
- [41] D. Lei, X. Chen, L. Cheng, L. Zhang, S.V. Ukkusuri and F. Witlox, Inferring temporal motifs for travel pattern analysis using large scale smart card data. Transportation Research Part C: Emerging Technologies, 2020.
- [42] J. Li, X. Sun, Q. Liu, W. Zheng, H. Liu and J.A. Stankovic, Planning electric vehicle charging stations based on user charging behavior, in: *2018 IEEE/ACM Third International Conference on Internet-of-Things Design and Implementation (IoTDI)*, 2018, pp. 225–236.
- [43] M. Li, Y. Zhu, T. Zhao and M. Angelova, Weighted dynamic time warping for traffic flow clustering, *Neurocomputing* **472** (2022), 266–279. doi:[10.1016/j.neucom.2020.12.138](https://doi.org/10.1016/j.neucom.2020.12.138).
- [44] W. Li, S. Wang, X. Zhang, Q. Jia and Y. Tian, Understanding intra-urban human mobility through an exploratory spatiotemporal analysis of bike-sharing trajectories, *International Journal of Geographical Information Science* (2020).
- [45] C. Liu, Z. Xiao, D. Wang, L. Wang, H. Jiang, H. Chen and J. Yu, Exploiting spatiotemporal correlations of arrive-stay-leave behaviors for private car flow prediction, *IEEE Transactions on Network Science and Engineering* **9**(2) (2022), 834–847. doi:[10.1109/TNSE.2021.3137381](https://doi.org/10.1109/TNSE.2021.3137381).
- [46] Y. Liu, J. Li, Z. Ming, H. Song, X. Weng and J. Wang, Domain-specific data mining for residents' transit pattern retrieval from incomplete information, *Journal of Network and Computer Applications* **134** (2019), 62–71. doi:[10.1016/j.jnca.2019.02.016](https://doi.org/10.1016/j.jnca.2019.02.016).

- [47] W.-C. Ma, D.-A. Huang, N. Lee and K.M. Kitani, Forecasting interactive dynamics of pedestrians with fictitious play, in: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 4636–4644. doi:[10.1109/CVPR.2017.493](https://doi.org/10.1109/CVPR.2017.493).
- [48] M.G. Martins, L. Nesi, P.R. da Silva Pereira and J.L.V. Barbosa, Delfos: A model for multitemporal analysis based on contexts history, *IEEE Latin America Transactions* **19** (2021).
- [49] C.M. Matos, V.K. Matter, M.G. Martins, J.E. da Rosa Tavares, A.S. Wolf, P.C. Bittenbender and J.L.V. Barbosa, Towards a collaborative model to assist people with disabilities and the elderly people in smart assistive cities, *JUCS – Journal of Universal Computer Science* **27**(1) (2021), 65–86. doi:[10.3897/jucs.64591](https://doi.org/10.3897/jucs.64591).
- [50] L. Meegahapola, T. Kandappu, K. Jayarajah, L. Akoglu, S. Xiang and A. Misra, Buscope: Fusing individual & aggregated mobility behavior for “live” smart city services, in: *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys’19*, Association for Computing Machinery, New York, NY, USA, 2019, pp. 41–53.
- [51] C. Meng, Y. Cui, Q. He, L. Su and J. Gao, Towards the inference of travel purpose with heterogeneous urban data, *IEEE Transactions on Big Data* **8**(1) (2022), 166–177. doi:[10.1109/TBDDATA.2019.2921823](https://doi.org/10.1109/TBDDATA.2019.2921823).
- [52] P. Nesi, C. Badii, P. Bellini, D. Cenni, G. Martelli and M. Paolucci, Km4city smart city api: An integrated support for mobility services, in: *2016 IEEE International Conference on Smart Computing (SMARTCOMP)*, 2016, pp. 1–8.
- [53] S.Z. Nielsen, R. Gade, T.B. Moeslund and H. Skov-Petersen, Taking the temperature of pedestrian movement in public spaces, *Transportation Research Procedia* **2** (2014), 660–668. The Conference on Pedestrian and Evacuation Dynamics 2014 (PED 2014), 22–24 October 2014, Delft, The Netherlands. doi:[10.1016/j.trpro.2014.09.071](https://doi.org/10.1016/j.trpro.2014.09.071).
- [54] Y. Qarout, Y.P. Raykov and M.A. Little, Probabilistic modelling for unsupervised analysis of human behaviour in smart cities, *Sensors (Switzerland)* (2020).
- [55] S.R. Rashmi, S. Bhat and V.C. Sushmitha, Evaluation of human action recognition techniques intended for video analytics, in: *2017 International Conference on Smart Technologies for Smart Nation (SmartTechCon)*, 2017, pp. 357–362. doi:[10.1109/SmartTechCon.2017.8358396](https://doi.org/10.1109/SmartTechCon.2017.8358396).
- [56] M.M. Rathore, A. Paul, S. Rho, M. Khan, S. Vimal and S.A. Shah, Smart traffic control: Identifying driving-violations using fog devices with vehicular cameras in smart cities, *Sustainable Cities and Society* **71** (2021), 102986. doi:[10.1016/j.scs.2021.102986](https://doi.org/10.1016/j.scs.2021.102986).
- [57] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, 2016, You only look once: Unified, real-time object detection.
- [58] D.L. Romeiro, F.L. Cardoso, R. Schechtman, L.C. Brizon and M.Z. Figueiredo, Transporte público e covid-19: o abandono do setor durante a pandemia, in: *FGV CERI*, Rio de Janeiro, RJ, Brasil, 2021, Centro de Estudos em Regulação e Infraestrutura da Fundação, Getúlio Vargas (FGV CERI).
- [59] A. Scaloni and D. Micheli, Estimation of mobility direction of a people flux by using a live 3g radio access network and smartphones in non-connected mode, in: *2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC)*, 2015, pp. 1869–1873. doi:[10.1109/EEEIC.2015.7165457](https://doi.org/10.1109/EEEIC.2015.7165457).
- [60] G.L. Schroeder, W. Heckler, R. Francisco and J.L.V. Barbosa, Problematic smartphone use on mental health: A systematic mapping study and taxonomy, *Behaviour & Information Technology* **42**(16) (2023), 2808–2831. doi:[10.1080/0144929X.2022.2149422](https://doi.org/10.1080/0144929X.2022.2149422).
- [61] I. Semanjski, A.J.L. Aguirre, J. De Mol and S. Gautama, Policy 2.0 platform for mobile sensing and incentivized targeted shifts in mobility behavior, *Sensors (Switzerland)* (2016).
- [62] I. Semanjski and S. Gautama, Crowdsourcing mobility insights – reflection of attitude based segments on high resolution mobility behaviour data, *Transportation Research Part C: Emerging Technologies* **71** (2016), 434–446. doi:[10.1016/j.trc.2016.08.016](https://doi.org/10.1016/j.trc.2016.08.016).
- [63] I. Semanjski, S. Gautama, R. Ahas and F. Witlox, Spatial context mining approach for transport mode recognition from mobile sensed big data, *Computers, Environment and Urban Systems* **66** (2017), 38–52. doi:[10.1016/j.compenvurbsys.2017.07.004](https://doi.org/10.1016/j.compenvurbsys.2017.07.004).
- [64] W.M. Shalash, A.A. AlZahrani and S.H. Al-Nufaii, Crowd detection management system, in: *2019 2nd International Conference on Computer Applications Information Security (ICCAIS)*, 2019, pp. 1–8.
- [65] A. Shamshiripour, E. Rahimi, R. Shabanpour and A.K. Mohammadian, Dynamics of travelers’ modality style in the presence of mobility-on-demand services, *Transportation Research Part C: Emerging Technologies* **117** (2020), 102668. doi:[10.1016/j.trc.2020.102668](https://doi.org/10.1016/j.trc.2020.102668).
- [66] L. Shuai, Z. Xueyan, S. Xiaodong, Y. Xiaohan, T. Ruichun and J. Qingyun, Survey on context-aware systems and their applications, in: *IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, 2019.
- [67] F. Silva, C. Analide and P. Novais, Traffic expression through ubiquitous and pervasive sensorization: Smart cities and assessment of driving behaviour, in: *2015 International Conference on Pervasive and Embedded Computing and Communication Systems (PECCS)*, 2015, pp. 33–42.
- [68] G. Singh, D. Bansal and S. Sofat, A smartphone based technique to monitor driving behavior using dtw and crowdsensing, *Pervasive and Mobile Computing* **40** (2017), 56–70. doi:[10.1016/j.pmcj.2017.06.003](https://doi.org/10.1016/j.pmcj.2017.06.003).
- [69] J. Soares, F. Lezama, A. Trindade, S. Ramos, B. Canizes and Z. Vale, Electric vehicles trips and charging simulator considering the user behaviour in a smart city, in: *2021 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, 2021, pp. 1–6.
- [70] G. Solmaz, F.-J. Wu, F. Cirillo, E. Kovacs, J.R. Santana, L. Sanchez, P. Sotres and L. Munoz, Toward understanding crowd mobility in smart cities through the Internet of things, *IEEE Communications Magazine* (2019).
- [71] M. Spanu, M. Bertolusso, G. Bingöl, L. Serreli, C.G. Castangia, M. Anedda, M. Fadda, M. Farina and D.D. Giusto, Smart cities mobility monitoring through automatic license plate recognition and vehicle discrimination, in: *2021 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2021, pp. 1–6.
- [72] M.J. Telles, R. Santos, J.M. da Silva, R. da Rosa Righi and J.L.V. Barbosa, An intelligent model to assist people with disabilities in smart cities, *Journal of Ambient Intelligence and Smart Environments* **13**(4) (2021), 301–324. doi:[10.3233/AIS-210606](https://doi.org/10.3233/AIS-210606).
- [73] The Anh Dang, J. Chiam and Y. Li, A comparative study of urban mobility patterns using large-scale spatio-temporal data, in: *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, 2018, pp. 572–579.

- [74] D. Tosi, Cell phone big data to compute mobility scenarios for future smart cities, *International Journal of Data Science and Analytics* **4**(4) (2017), 265–284. doi:[10.1007/s41060-017-0061-2](https://doi.org/10.1007/s41060-017-0061-2).
- [75] E.B. Varghese and S.M. Thampi, Visual attention based cognitive informative frame extraction method for smart crowd surveillance, in: *2021 IEEE Conference on Norbert Wiener in the 21st Century (21CW)*, 2021, pp. 1–9.
- [76] C.-Y. Wang, A. Bochkovskiy and H.-Y.M. Mark Liao, Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, 2022.
- [77] Z. Xiao, S. Xu, T. Li, H. Jiang, R. Zhang, A.C. Regan and H. Chen, On extracting regular travel behavior of private cars based on trajectory data analysis, *IEEE Transactions on Vehicular Technology* **69**(12) (2020), 14537–14549. doi:[10.1109/TVT.2020.3043434](https://doi.org/10.1109/TVT.2020.3043434).
- [78] Z. Xu, J. Li, Z. Lv, C. Dong and L. Fu, A classification method for urban functional regions based on the transfer rate of empty cars, *IET Intelligent Transport Systems* **16**(2) (2022), 133–147. doi:[10.1049/itr2.12134](https://doi.org/10.1049/itr2.12134).
- [79] Y. Yamagata, D. Murakami, Y. Wu, P.P.-J. Yang, T. Yoshida and R. Binder, Big-data analysis for carbon emission reduction from cars: Towards walkable green smart community, *Energy Procedia* **158** (2019), 4292–4297, Innovative Solutions for Energy Transitions.
- [80] H. Yamaguchi, A. Hiromori and T. Higashino, A human tracking and sensing platform for enabling smart city applications, in: *Proceedings of the Workshop Program of the 19th International Conference on Distributed Computing and Networking, Workshops ICDCN'18*, Association for Computing Machinery, New York, NY, USA, 2018.
- [81] D. Yang, D. Zhang, L. Chen and B. Qu, Nantotelescope, Monitoring and visualizing large-scale collective behavior in lbsns, *Journal of Network and Computer Applications* **55** (2015), 170–180. doi:[10.1016/j.jnca.2015.05.010](https://doi.org/10.1016/j.jnca.2015.05.010).
- [82] D. Yang, D. Zhang and B. Qu, Participatory cultural mapping based on collective behavior data in location-based social networks, *ACM Trans. Intell. Syst. Technol.* **7**(3) (2016).
- [83] C. Zhang, K. Zhao and M. Chen, Beyond the limits of predictability in human mobility prediction: Context-transition predictability, *IEEE Transactions on Knowledge and Data Engineering* (2022).
- [84] X. Zhou, Y. Jia, C. Bai, H. Zhu and S. Chan, Multi-object tracking based on attention networks for smart city system, *Sustainable Energy Technologies and Assessments* **52** (2022), 102216. doi:[10.1016/j.seta.2022.102216](https://doi.org/10.1016/j.seta.2022.102216).
- [85] Z. Zhou, H. Cai and Z. Zhou, An analysis of pedestrians' behavior in emergency evacuation using cellular automata simulation, in: *2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, 2019, pp. 2644–2650.
- [86] O. Zinman and B. Lerner, Utilizing digital traces of mobile phones for understanding social dynamics in urban areas, *Personal and Ubiquitous Computing* **24**(4) (2020), 535–549. doi:[10.1007/s00779-019-01318-w](https://doi.org/10.1007/s00779-019-01318-w).