Urban data analysis for the public transportation system of Montevideo, Uruguay

Renzo Massobrio and Sergio Nesmachnow

Universidad de la República, Uruguay
{renzom,sergion}@fing.edu.uy

Abstract. This article presents a study of the public transportation system in Montevideo, Uruguay, following a data science approach. More than 20 million records from the Intelligent Transportation System (ITS) are analyzed in order to characterize mobility in the city. Several useful pieces of information are obtained through data analysis, related to tickets sold, patterns of smart card utilization, most used bus lines and stops, and socioeconomic insights about passengers behavior. Practical case studies are also presented: anomaly detection in space and time, and a study of potential safety hazards due to reckless driving. The work reported in this article constitutes one of the first steps towards using data from the ITS in Montevideo to understand mobility in the city.

Keywords: data science · public transportation · smart cities

1 Introduction

Public transportation plays a major role in urban mobility, as they are the most efficient, sustainable, and socially fair mode of transportation [9]. Understanding the interaction between citizens and public transportation systems is paramount in order to design and implement policies aimed at improving mobility.

Modern *smart cities* take advantage of technology to improve urban services [6]. Urban traffic and transportation systems are addressed under the paradigm of smart cities, in what is referred to as *smart mobility* [3]. Related to this concept are Intelligent Transportation Systems (ITS), which use technology to develop and enhance transportation. In addition to improving mobility, ITS allow collecting large volumes of urban data [7]. Large repositories of data offer a unique opportunity to gain valuable insights into the mobility of citizens [12]. In this context, urban data analysis arises as a tool to extract meaningful information from raw urban data to help decision-making processes in cities.

In 2010 an urban mobility plan was implemented in Montevideo, Uruguay, with the goal of restructuring and modernizing public transportation [1]. Under this plan, public transportation in the city was integrated into a unified system, which incorporates many of the characteristics common to ITS. Buses were equipped with on-board GPS units and ticket selling machines operated with smart cards. These devices represent new sources of urban data, which have a huge potential to help authorities understand mobility in Montevideo.

This article presents a characterization of the use of the transportation system through urban data analysis, along with a series of case studies of potential use for these rich data sources. The article is organized as follows. Next section describes the data analysis process in smart cities. Section 3 describes the public transportation system in Montevideo, Uruguay. Specific case studies are described and analyzed in Section 4. Finally, Section 5 presents the conclusions and the main lines for future work.

2 Data analysis in smart cities

This section introduces ITS in the context of smart cities, and urban data analysis as an efficient tool to extract meaningful information from urban data. Then, a brief review of related works in the literature is presented.

2.1 Smart cities and ITS

The paradigm of smart cities proposes taking advantage of information and communication technologies to improve the quality and efficiency of urban services [6]. Smart devices embedded into traditional physical systems deployed on cities, generate vast volumes of data for the analysis. Extracting insights from the gathered data is crucial to improve decision-making in cities and to achieve quality improvements and increase the efficiency of public services.

Related to smart mobility, ITS integrate synergistic technologies, computational intelligence, and engineering concepts to develop and improve transportation. Automatic Vehicle Location (AVL) systems automatically determine and communicate the geographic location of a moving vehicle [17]. The transmitted locations of a fleet of vehicles can be collected at a central server to overview and control the group of vehicles. Due to its widespread availability, low cost, and precision, the most common technology to determine the location of vehicles in AVL is GPS. AVL technology is frequently incorporated in ITS and provides a rich source of data, as it can help to monitor and control the QoS provided by the transportation system to users.

Automatic Passenger Counters (APC) are electronic devices that can be incorporated to moving vehicles to record boarding and alighting data [5]. This technology is a major improvement over traditional manual passenger counts or surveys. Several implementations of this concept have been proposed including infrared lights in doorways of vehicles, scales to measure weight changes, and CCTV cameras coupled with computer vision software. The data generated by these systems allow identifying use patterns by linking boarding and alighting data with stop or station location [8].

Automatic Fare Collection (AFC) [4] automate the ticketing system on public transportation. AFC are comprised of fare media, devices to read/write onto these media, communication technologies, and back office systems. Contactless smart cards have become the de facto technology in AFC systems. AFC systems generate highly valuable data that can be processed to extract useful metrics for both day-to-day operation and long-term planning of transportation systems.

The development of smart tools that use data gathered by ITS has risen in the past years. These tools rely on efficient and accurate data processing (even in real-time), which poses an interesting challenge from the technological perspective. The methodology for analyzing sources of urban data to gain valuable insights to describe and improve the life of citizens is described next.

2.2 Urban data analysis

Data analysis is the process of collecting and processing raw data to extract meaningful information that provides supporting evidence for conclusions and helps decision-making processes. Multiple workflows have been proposed to describe the process of data analysis, and techniques under a variety of names have emerged in different fields of knowledge at both academia and industry.

The data analysis process starts and ends in the current reality. In urban contexts, the analysis starts with collecting raw data from a city and ends with communicating findings that can help stakeholders to shape the reality of that city to improve the quality of life of its citizens. In between, the data analysis process is comprised of several phases. Firstly, raw collected data must be processed. This phase include several tasks such as placing data into structures (e.g., tables), inspecting datasets, and cleansing data to detect missing or inaccurate records. After data processing, Exploratory Data Analysis (EDA) is performed [15]. This phase may lead to detecting further inaccuracies in the data and potentially requiring further cleansing. After EDA, statistical models and algorithms are applied to identify relationships between the studied data [11]. Finally, results are interpreted and communicated, mostly using visualization. When dealing with urban data, effectively communicating results is crucial, thus, the visualization phase is described in more depth in the following paragraphs.

2.3 Related work

The advantages of using data analysis for social transportation have been studied in a thorough manner in the general review of the field developed by Zheng et al. [18]. The authors discussed the use of several sources of information, including vehicle mobility (e.g., GPS coordinates, speed data), pedestrian mobility (e.g., GPS and WiFi signals from mobile devices), incident reports, social networking (e.g., textual posts, user location), and web logs (e.g., user identification, comments). In the review, the advantages and limitations of using each source of data were discussed. Several other novel ideas to improve public transportation were also reviewed, including applying *crowdsourcing* techniques for collecting and analyzing real-time or near real-time traffic information, and using *data-based agents* for driver assistance and human behavior analysis. Finally, a data-driven social transportation system that integrates all the previous concepts and improves traffic safety and efficiency was proposed.

More related to the data analysis research included in this article, many works have studied urban mobility using smart card data from AFC in public transportation systems.

Bagchi and White [2] discussed the role of smart card data for travel behavior analysis. The transportation systems of Southport, Merseyside and Bradford in England were studied. The authors performed a simple study focused on the average number of trips and transfers made by passengers. The turnover rates were analyzed to identify the number of active users in the system. The research concluded that smart card data allow obtaining much larger samples than surveys to characterize transportation systems. However, certain information (e.g., purpose of traveling) cannot be inferred from these data. Thus, the authors conclude that smart card transactions are not an alternative to traditional data collection methods, but a useful complementary source of data.

Utsunomiya et al. [16] studied access and usage patterns of passengers in the transportation system of Chicago, US. The authors discussed the analysis using smart card sign-ups and transactions data, identifying the major issues encountered as well as general recommendations. The potential uses for smart card data were classified in categories: service planning, demand forecasting, pricing and fare policy definition, and market research. Seven days of recorded transactions were studied to analyze walking access distances, frequency of daily travel patterns, and passenger behavior by residential area. Frequent errors were due to missing transactions and incorrect bus route identification. In order to deal with these inconsistencies, the authors proposed combining smart card data with passenger counts and vehicle location from APC and AVL systems.

3 The public transportation system in Montevideo

This section presents the public transportation system in Montevideo, Uruguay, and the urban data analysis process to characterize its usage.

3.1 Overview of the city and transportation system

Montevideo is the capital city of Uruguay, and extends to an area of only 530 km^2. Despite accounting for only 0.3% of the total surface of the country, nearly 40% of the total population lives in Montevideo.

The public transportation in Montevideo (Sistema de Transporte Metropolitano, STM), is comprised of 1528 buses operated by four private companies. The bus network consists of 145 main bus lines and 4718 bus stops, Bus lines have different variants, accounting for outward and return trips, as well as shorter versions of the same line. The total amount of bus lines when considering each variant individually is 1383. Figure 1 shows the bus lines that comprise STM according to data provided by [14].

Passengers of the public transportation system in Montevideo can use smart cards to pay for their tickets without physical money. STM smart cards are contact-less cards which are linked to the identity of the owner. Two different

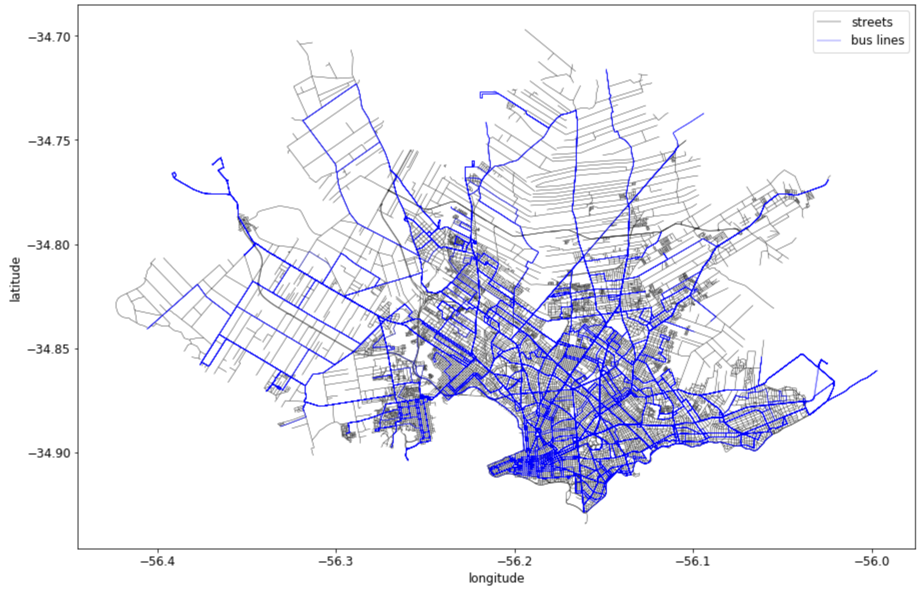

Fig. 1. Bus lines in STM

types of bus tickets are available which allow bus transfers: *one-hour* and *two-hours* tickets. One-hour tickets allow boarding up to two buses within an hour, while two-hours tickets grant unlimited bus transfers within a period of two hours. The fare scheme supports transfers between any bus line at any bus stop.

3.2 Data collection and cleansing

Data collection. Many state agencies and local governments have web interfaces for publishing open data. In this context, the most useful web interface was the geographic information site at Intendencia de Montevideo (`www.sig.montevideo.gub.uy`), which holds geographic data of Montevideo including base maps, socioeconomic indicators, and transportation network data.

Besides using open data publicly available, the analysis included data regarding STM accessed through a collaboration between our research group and IM. The sources of these data are the AVL and AFC systems integrated in buses of the STM. The data corresponding to the full set of records of GPS bus location and bus ticket sales payed with STM cards during 2015 was released for research purposes. These large datasets comprise over 150 GB of raw data.

The bus location dataset contains information about the position of each bus in STM, sampled every 10 to 30 seconds. Each location record holds the following information:

- a unique bus line identifier.
- a unique trip identifier to differentiate trips of the same bus line.
- GPS coordinates.

– instant speed of the vehicle.
– time stamp when the GPS measure was taken.

Ticket sales data contain records related to each STM transaction made, including the following fields:

– trip identifier for the sale, which allows linking to the bus location dataset.
– GPS coordinates at the moment of the STM card validation.
– bus stop identifier.
– time stamp at the moment of the STM card validation.
– unique STM card identifier, hashed for privacy purposes.
– number of passengers traveling with the same STM card.
– leg number, for multi-leg trips that include transfers.

For the sake of clarity, the reported results correspond to tickets sold during May 2015. Pre-hoc analysis of the complete dataset showed that this month is representative of the full dataset.

Data cleansing. Data cleansing is a mandatory step in data analysis that strives to detect and correct corrupt or inaccurate records [13]. Due to the lack of a backup source of information, records that appeared to be corrupted were simply filtered and deleted, according to the actions described next.

Vehicle location using GPS is prone to errors from a variety of sources. The most frequent error was records having a fixed value for both latitude and longitude, pinpointing to the middle of the Atlantic Ocean. Most likely, this was caused by an error message of the GPS unit being misinterpreted as a valid coordinate during data recording. 932.176 records suffered from this issue, accounting for nearly 4.6% of the total dataset. Additionally, 29.432 records corresponded to locations outside Montevideo. However, the dataset also holds the identifier of the boarding bus stop of each transaction, registered by the on-board GPS unit. Thus, even though the GPS measure at the moment of the transaction may fail, the boarding bus stop can be accurately determined from previous measures. Consequently, the bus stop identifier is more reliable than the raw GPS measure when defining the starting point of each trip.

Regarding time stamps of transactions, 74 sales corresponding to May 1^{st} were filtered, since they correspond to Labour Day, a public holiday in which the transportation system is mostly inoperative. Those transactions represent a clear outlier from the remainder of the dataset. Similarly, only one transaction occurring on May 31^{st} was present in the dataset. As a consequence, during the data analysis process, the month of May represents STM transactions occurring between May 2^{nd} 00:00:00 to May 30^{th} 25:59:59 of 2015.

Some transactions had trip identifiers which were not present in the GPS records. Since these records cannot be linked to their corresponding bus line, the 1634 records with this issue were discarded. Similarly, 22 transactions made with the same STM card during the same trip were detected in the original dataset. This might be explained by a synchronization problem between the bus and the centralized server where transactions are recorded.

Since the dataset corresponds to sales from 2015, some transactions refer to bus lines that were modified or no longer exist. These transactions (36.030 records) were also filtered from the dataset. Finally, 274.011 records were filtered, corresponding to transactions with identifiers of bus stops which were not part of the bus line route corresponding to the sale.

In summary, the complete data cleansing process consisted in filtering 311.772 out of a total of 20.359.835 records, accounting for 1.53% of the original dataset.

3.3 Characterizing public transport utilization

This section presents the results from the data analysis process to describe the use of the transportation system in Montevideo from several perspectives.

Cardholders. The sales dataset holds transactions made with 654.228 STM cards. The public transportation system allows several passengers to travel together using the same STM card. Table 1 reports the number of passengers traveling with the same STM card. The vast majority of passengers use their personal STM card, with over 97% of transactions corresponding to individual ticket sales.

Table 1. Number of passengers traveling with the same STM card

# passengers	*total*	*percentage*
1	19494451	97.24%
2	510043	2.54%
3	36454	0.18%
4	5468	0.03%
5+	1647	0.01%

Another interesting aspect that can be studied through data analysis is the frequency of use of the transportation system. Table 2 reports descriptive statistics of daily and monthly transactions per STM card. The *mean* number of transactions is reported, along with the standard deviation (*std*), the minimum (*min*) and the maximum (*max*) values, and the 25^{th} (*Q1*), 50^{th} (*Q2*), and 75^{th} (*Q3*) percentiles. The 50^{th} percentile corresponds to the median of the distribution of transactions per STM card. Monthly statistics consider all transactions done by each cardholder during May 2015. Daily transaction statistics only consider days for which at least one transaction was made.

When looking at monthly figures, cardholders perform over 30 transactions on average, nearly one transaction per day. However, the standard deviation is large, indicating a significant difference between regular and sporadic users of the public transportation system. The median of the monthly transactions is 22, nearly one transaction per working day in the month. Regarding daily use, the average cardholder performs 2.78 transactions each day that uses the transportation system. Figure 2 presents an histogram of daily transactions per STM card, considering only cards that made up to 10 transactions within the

Table 2. Descriptive statistics of daily and monthly use of STM cards

	STM transactions	
	daily	*monthly*
mean	2.78	30.65
std	1.53	28.14
min	1	1
Q1 (25%)	2	8
Q2 (50%)	2	22
Q3 (75%)	4	47
max	54	528

same day in order to remove outliers. Most cardholders perform two transactions per day, which probably correspond to direct trips used for commuting. It is interesting to observe that more cardholders perform four rather than three transactions. This might be explained by passengers commuting to work using a trip involving a transfer, thus, two transactions correspond to the outward trip and the remaining two transactions to the return trip.

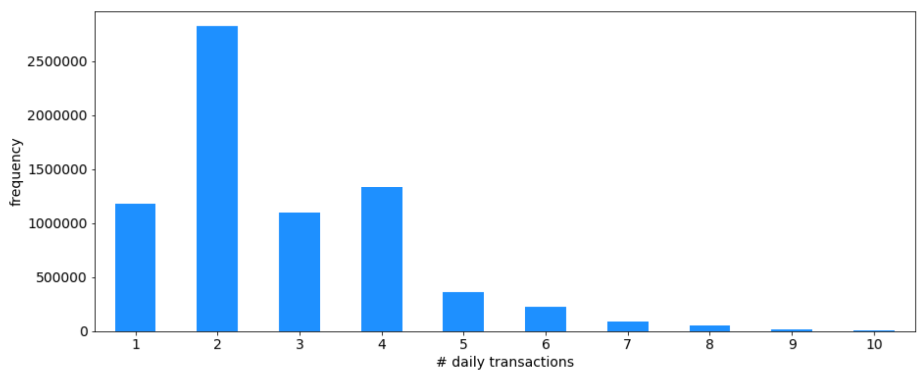

Fig. 2. Histogram of daily transactions per STM card during May 2015

A few interesting applications arise when looking at outliers within the STM use statistics. On the one hand, cardholders with very low activity can be identified by their card ID. For instance, in the studied dataset 15.440 cardholders performed only a single trip during the whole month of May 2015. Targeted marketing campaigns could be designed to encourage disengaged citizens to use the public transportation system more frequently. On the other hand, cardholders with large number of transactions can also be identified. In the studied dataset a single card was found to perform 54 transactions within the same day. Through data analysis, authorities may further investigate these situations in order to identify possible abuses to the rules of the transportation system.

Transactions per bus trip. Grouping STM transactions by their corresponding trip identifier provides a rough estimate for the number of boardings on each trip. Figure 3 presents a histogram of the number of transactions per trip. On average, 39.70 transactions are made in each bus trip (std: 28.16). The largest value encountered was a single trip with 249 transactions. It is worth noting that passengers might also board without using a STM card, so these figures represent a lower bound on the total number of boardings for each trip. Taking into account the capacity of the buses operating in Montevideo, the largest values may indicate overcrowding in some of the bus lines of the transportation system.

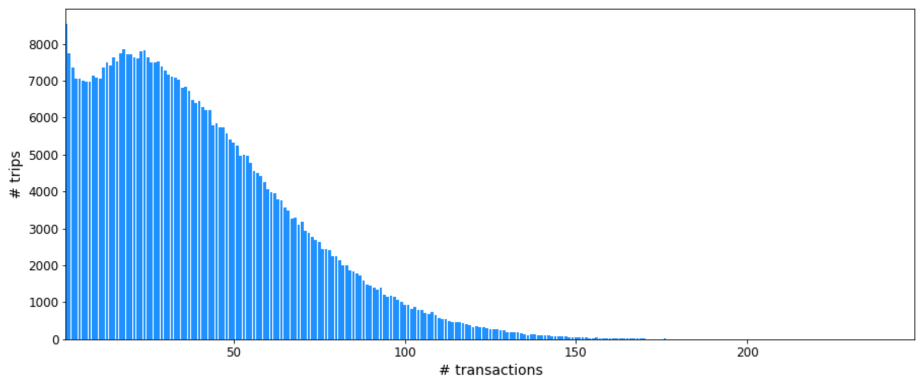

Fig. 3. Histogram of transactions per bus trip during May 2015

Most used bus lines. Data analysis over the transaction data can be used to identify the most popular as well as the most underused bus lines. Figure 4 shows the ten most used bus lines. Some of the lines overlap since they correspond to different variants of the same line (e.g., outward and return lines). For each line the regular name (i.e., the name appearing in the front of the bus) is indicated in the map, along with its variant code indicated in parenthesis. The most used bus line is 183, closely followed by 181. Both lines connect the neighborhood of La Teja, located in the west side of Montevideo, with Pocitos, located in the south by the coastline. It is interesting to notice that none of the ten most used bus lines go into the city center.

Spatiotemporal analysis of transactions. The spatial and temporal dimensions of sales data can be combined, in order to gain insights that might not be evident when studying each dimension independently. Figure 5 shows an aggregated visualization of the spatiotemporal distribution of sales in Montevideo during May 2015. In this visualization the hours of the day are used as categories. Each transaction occurring at a given pixel in the image is categorized according to its time stamp. Then, the color of the pixel is set considering the amount of transactions on each category. The color mapping, which is detailed

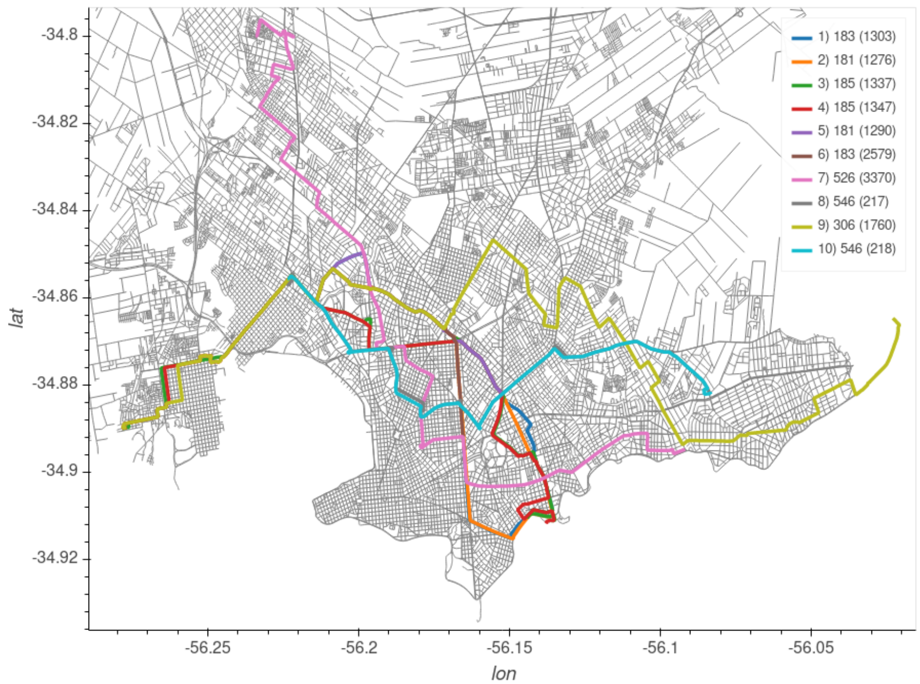

Fig. 4. Top 10 bus lines with most STM card transactions during May 2015

in the visualization, corresponds roughly to: red (12 a.m.), yellow (4 a.m.), green (8 a.m.), cyan (12 p.m.), blue (4 p.m.), purple (8 p.m.), and back to red, since hours and colors are both cyclic.

Firstly, it is observed that the city center has a prevalent blueish tone in the visualization. This corresponds to most transactions taking place between noon and the afternoon. This is consistent with the fact that many offices and public entities are located in this area of the city, thus, most transactions correspond to people commuting from the city center back to their homes by the end of the office-hours.

Another interesting fact arising from the spatiotemporal analysis of STM transactions is the clear difference between areas near the coast and areas farther away. It can be clearly observed that areas away from the coastline appear with more yellow and greener tones whereas areas closer to the coast have predominantly blue tones. This means that the majority of STM transactions in areas farther away from the coast occur earlier in the day than those near the coast. This can be explained by people commuting early in the day from these areas to workplaces located closer to the city center.

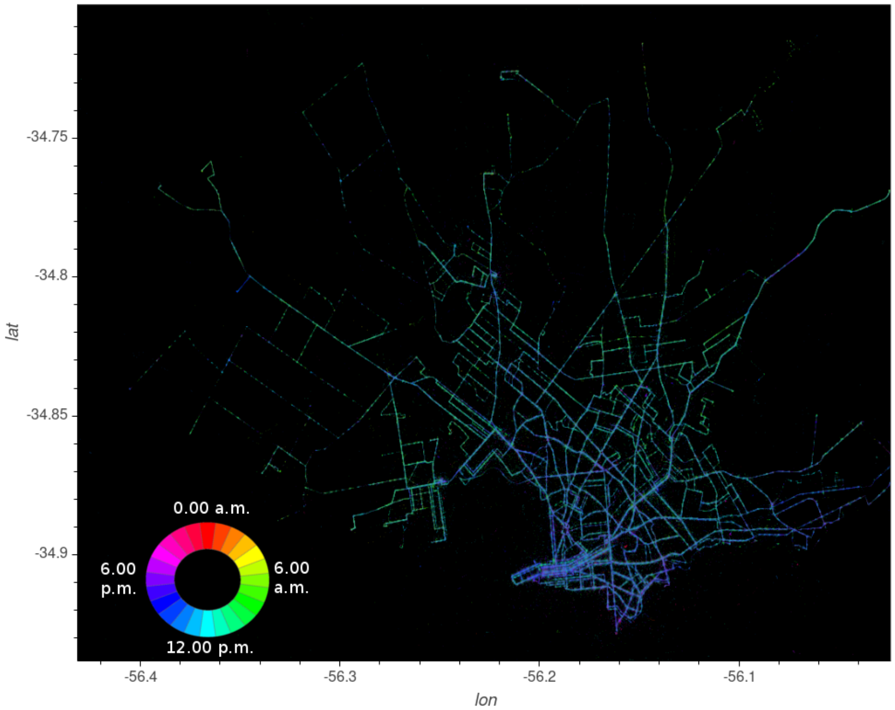

Fig. 5. Spatiotemporal distribution of trips in Montevideo during May 2015

4 Case studies: unexpected events and safety

This section outlines a series of case studies of uses of urban data analysis to detect special events and safety hazards within the public transportation system.

4.1 Anomaly detection in the spatial dimension

Geolocation data of sales transactions can be used to detect abnormal situations in the transportation system. As an example, Figure 6 shows a heatmap of transactions, along with the streets (in gray) and the bus lines (in blue). Two clusters of sales records (labeled A and B) appear in a street where no bus routes run. This represents a detour of one or more bus lines from their predefined routes. This may be due to an exceptional circumstance (e.g., road works) or due to a periodic event occurring certain days of the week (e.g., a flea market). Authorities can take advantage of this type of analysis to identify anomalies and make appropriate changes to bus routes and schedules.

4.2 Anomaly detection in the time dimension

The time stamp of sales can be used to identify abnormal use patterns in the transportation system. Figure 7 shows an aggregated visualization of combined

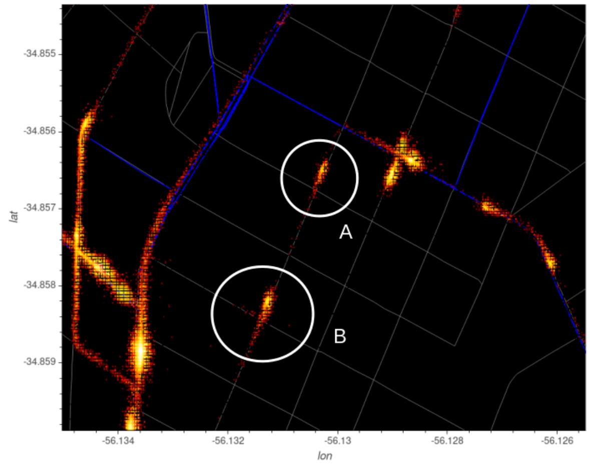

Fig. 6. Anomaly detection: example of detour. The blue lines represent bus routes. A and B are two clusters of transactions which occurred outside of the bus network.

spatial and temporal information of smart card transactions data. A small cluster of pixels in red can be observed in the map (indicated with a circle), which correspond to a group of sales occurring approximately at midnight. This pattern significantly differs from the rest of the dataset. Given the location of these records, near an outdoor venue, the transactions probably correspond to a special event (e.g., a concert) taking place at night in this venue. In these occasions, bus companies usually assign buses to allow citizens to return to their homes at the end of the event. Authorities can use urban data analysis to identify special events taking place in the city and implement strategies that improve the mobility of those attending these events.

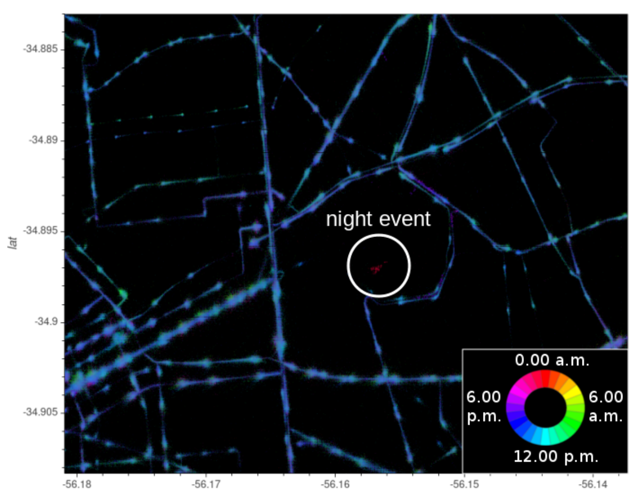

Fig. 7. Anomaly detection: example of event at midnight near an outdoor venue.

4.3 Driving behavior and safety

Another interesting use for information is to analyze the spatial distribution of sales. Figure 8 shows a heatmap of transactions occurring in one-way streets. Arrows indicate the direction of each street and bus stops are represented using blue circles. The visualization shows that the spatial distribution of sales is skewed with respect to the location of the bus stops: more transactions occur after the location of the bus stop than before. This uneven distribution is probably caused by drivers moving the bus before all the boarding passengers validate their smart cards. This might represent a safety issue, since passengers are standing while validating their cards.

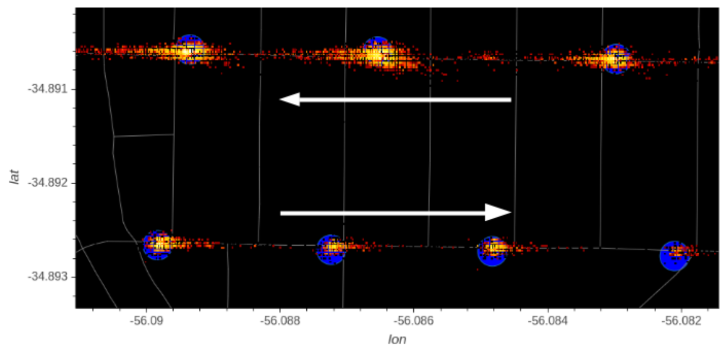

Fig. 8. Spatial distribution of transactions with regards to stop location: one-way streets

Figure 9 shows a heatmap of transactions near a roundabout, where a large number of transactions take place within the roundabout. Again, passengers are standing and validating their smart cards while the bus is moving. In fact, this might represent an even more serious issue, when drivers are also in charge of operating the smart card terminal. Driving and selling tickets at the same time is a risky behavior, frequently seen among bus drivers in Montevideo. The studied data provide evidences that support these observations. Authorities can use this type of data analysis to audit driving behavior, improving the safety of passengers and drivers of the transportation system.

5 Conclusions and future work

Under the paradigm of smart cities, ITS have emerged to take advantage of information and communication technologies to improve public transportation. ITS allow collecting massive amounts of urban data, which can be used to extract meaningful information to help decision making in cities. This article studied data from the ITS in Montevideo, Uruguay, to characterize mobility in the city.

The results reported in this article account for more than 20.4 million bus tickets sold using smart cards. During a data cleansing process, 1.53% of the records were filtered due to inconsistencies. Several insights were obtained through data analysis of the studied dataset, including: number of passengers traveling

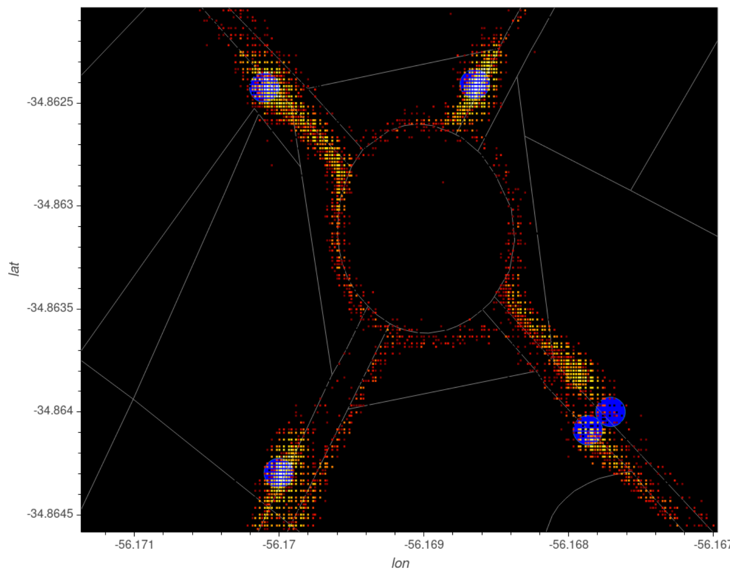

Fig. 9. Spatial distribution of transactions with regards to stop location: roundabout

with the same smart card, frequency of use of the smart cards, number of bus transfers, number of transactions per bus trip, and most used bus lines and stops. A spatiotemporal analysis was also performed which revealed that citizens from areas farther away from the coastline start trips earlier than those near the coast.

Finally, some practical case studies on the use of data analysis on ITS data were presented, including: anomaly detection in space (to identify bus detours), anomaly detection in time (to identify events in the city), and a characterization of potential safety hazards due to reckless driving. Such analysis are useful for characterizing different aspects of mobility in smart cities [10, 12].

The work reported in this article constitutes one of the first steps towards using data from the ITS in Montevideo to understand mobility in the city. As such, many lines of research remain to be explored in order to extract more and richer information that can be used to improve the public transportation system.

The data analysis reported in this article mainly focused in understanding the interaction between passengers and the transportation system. However, the available data sources offer the potential to study other very interesting aspects of mobility in the city. For instance, location data from AVL systems could be used to further study the QoS offered to citizens by the transportation system in terms of punctuality, frequency of lines, and load of passengers with regards to the bus capacity. Additionally, speed information of buses could be used to characterize the streets of the city and identify bottlenecks. This information could be used as input when designing new lines or re-designing existing ones.

Finally, it is worth noting that this work used ITS data from 2015. Since that year, the use of smart cards to pay for tickets has risen significantly. The proposed approach should be applied to recent ITS data when it becomes available publicly.

References

1. Abreu, P., Vespa, J.: Plan de Movilidad. Online: http://www.montevideo.gub.uy/sites/default/files/plan_de_movilidad.pdf (2018-12-30) (2010)
2. Bagchi, M., White, P.: The potential of public transport smart card data. Transport Policy **12**(5), 464–474 (2005)
3. Benevolo, C., Dameri, R.P., D'Auria, B.: Smart mobility in smart city. In: Torre, T., Braccini, A.M., Spinelli, R. (eds.) Empowering Organizations. pp. 13–28. Springer International Publishing (2016)
4. Blythe, P.: Improving public transport ticketing through smart cards. In: Proceedings of the Institution of Civil Engineers-Municipal Engineer. vol. 157, pp. 47–54 (2004)
5. Boyle, D.K.: Passenger counting systems. Transportation Research Board (2008)
6. Deakin, M., Waer, H.A.: From intelligent to smart cities. Intelligent Buildings International **3**(3), 140–152 (2011)
7. Figueiredo, L., Jesus, I., Tenreiro, J., Ferreira, J., Martins, J.: Towards the development of intelligent transportation systems. In: IEEE Intelligent Transportation Systems. pp. 1206–1211 (2001)
8. Furth, P., Hemily, B., Muller, T., Strathman, J.: Using archived AVL-APC data to improve transit performance and management. Tech. Rep. 113, Transit Cooperative Research program-Transportation Research Board (2006)
9. Grava, S.: Urban Transportation Systems. McGraw-Hill Professional Publishing (2000)
10. Hipogrosso, S., Nesmachnow, S.: Sustainable mobility in the public transportation of Montevideo, Uruguay. In: II Iberoamerican Congress on Smart Cities (2019)
11. Judd, C., McClelland, G., Ryan, C.: Data analysis: A model comparison approach. Routledge (2011)
12. Nesmachnow, S., Bana, S., Massobrio, R.: A distributed platform for big data analysis in smart cities: combining Intelligent Transportation Systems and socioeconomic data for Montevideo, Uruguay. EAI Endorsed Transactions on Smart Cities **2**(5), 1–18 (2017)
13. Rahm, E., Do, H.H.: Data cleaning: Problems and current approaches. IEEE Data Engineering Bulletin **23**, 3–13 (2000)
14. Servicio de Geomtica - Intendencia de Montevideo: Lneas de Transporte (2012), [Data file] Online: http://geoweb.montevideo.gub.uy/geonetwork/srv/es/metadata.show?uuid=307ffef2-7ba3-4935-815b-caa7057226ce (2018-12-30)
15. Tukey, J.: Exploratory Data Analysis. Addison-Wesley Publishing Company (1977)
16. Utsunomiya, M., Attanucci, J., Wilson, N.: Potential uses of transit smart card registration and transaction data to improve transit planning. Transportation Research Record: Journal of the Transportation Research Board **1971**, 119–126 (2006)
17. Zhao, Y.: Vehicle location and navigation systems. Artech House Publishers (1997)
18. Zheng, X., Chen, W., Wang, P., Shen, D., Chen, S., Wang, X., Zhang, Q., Yang, L.: Big data for social transportation. IEEE Transactions on Intelligent Transportation Systems **17**(3), 620–630 (2016)