Mobility and Travel Behaviour Data Collection

by

Yassine Baghoussi

Master Dissertation in Data Analysis and Decision Support Systems

Faculty of Economics, University of Porto

Internal supervisors:
Prof. Rosaldo José Fernandes Rossetti
Prof. Pedro José Ramos Moreira de Campos
External supervisors:
Prof. Fabien Leurent

September 30, 2016

Acknowledgements

"I would like to thank my supervisors, Prof. Pedro Campos, Prof. Fabien Leurent and Prof. Rosaldo Rossetti, for the patient guidance and encouragement they have provided during their supervision. I have been extremely fortunate to have a supervisors who cared so much about my work, and who replied to my queries and questions so promptly. In particular I would like to thank Zoi Christoforou ,Lydia Ladjouze and Alexis Poulhes without them Chapter 6 would not be realized.

Sincere thanks to Khalid, my father, for his continued support and encouragement. I was continually amazed by his willingness to proofread countless pages, by the support and patience of my mother, sister and brother who lived all the ups and downs of my studies abroad.

Finally, I would like to thank all the staff at LVMT, the laboratory where I have spent memorable six months of internship, not only for providing the funding which allowed me to conclude this work, but also for giving me the opportunity to attend conferences and meet so many interesting people."

Yassine Baghoussi

Abstract

For long, the pedestrian and passenger mobility behavior has been a key subject in urban researches. Many technologies were developed in the same time to fit the needs of understanding the urban mobility and many of the researchers thought about gaining the opportunity from the advances in technology. The measurement of mobility and travel behaviour is important because it is considered as an efficient way to update the current status of cities, promote behavioral changes and build a mutual satisfaction between cities and urban population. The following thesis introduces the literature review of the traditional and innovative technologies that have been already conceptualized to measure urban mobility behaviour, as a theorical part. Among these methods, we have, passenger counting using sensors, household travel surveys, Automated Fare Collection (AFC) systems and smartphone-based data collection and all of these methods provide Big Data. The present thesis was done during an internship in a Parisian Laboratory, namely, "Laboratoire Ville Mobilité et Transports" which was a good opportunity for me to improve my knowledge on Big Data analysis and a bit more on AFC and several other methods on behavioural data collection. My internship also addressed several other technologies such as data mining and agents technology. Finally, I conclude the thesis with a study that I have conducted based on automated fare collection data. The study aimed to analyze a pre-post evaluation of passengers behavioural change after the new flat fare system that took place in the greater Paris on September, 2015…

Contents

Introduction

How do mobility get smarter ? This was the first question that came to our minds before to begin our research. The answer is not as much easy as we may think, even though we all agree that technology advances are the main factor of the actual smartness of the world. To evaluate the smartness of a city, we should evaluate the urban mobility taking into account its transportation system. Passengers, vehicles and bikes traffics represents the basic topics for a mobility-based study. Many researches argued that observation is an efficient way to better understand a fact. However, the accuracy of observation is correlated positively with the accuracy of analyzing. If we are able to observe a fact, we also need to analyze it in the way we can produce conclusions according to what we do want to know. The aim of this thesis is to explore the techniques available to understand the passengers mobility behavior in urban areas or in transit systems. Based on the technologies currently present in the market, where some of them were not fully dedicated for such purpose like video cameras, the literature have categorized these technologies in two categories : traditional and innovative.

The traditional technologies represents the passenger counting methods that use tools such as camera, infra-red etc. which were used at the beginning for security purposes in the urban areas. The traditional technologies also include household travel surveys which are based on standard methods often set by the authorities specially in case of public urban areas.

Recently, new payment systems have been introduced to the market, and the transportation operators did not hesitate to adopt it. Smart cards to understand passengers usage of public transport. The idea is simple, by the mean of a smart card the transit users record the fares for future needs and the users must validate the fare every time they want to use the public transport. The validation is done by passing the smart card, in which the fares are recorded, on a reader. The reader check if there is already a paid fares to allow the access and save the validation time and several other information data that are used to understand urban mobility behavior in the future. These systems are called Automated Fare Collection (AFC) which replace the traditional ticketing systems.The data provided by AFC systems opens very strong possibilities to better and accurately understand the behavior of transit users. Among these possibilities, agents technology remain the main tool used for both modeling and simulating these data (Paul Hamilton, 2013), (X. Shi and Song, 2005). Furthermore, since the AFC is only able to provide data about transit practices, many researches have intended to find out new methods that can enlarge the study from transit systems to all the urban areas.

Today, the focus is fully dedicated to passengers and the proper idea is to track their mobility, by the way we would be able to understand real time individual behavior. The only way that permits such findings is the one related to the technology that accompanies the individuals. As a result, all the sights are now on the Smartphones and to a lesser extent smart-watches. Smartphone is a technology that accompanies us everywhere. Many individuals use it to make calls or send messages or to check emails, play games and read the news in their free time. Thanks to all the integrated devices and sensors, smartphone is now able to provide different kinds of sensing data where the most significant one is GPS data. In addition to sensors data, the smartphone's user may be also interested in providing their daily mobility information in different manners like using travel diary apps. The implementation is in its first stage, and the research argues a promising future.

This thesis was done in a period of six months of internship in LVMT, a Parisian laboratory in which I have learned and made an useful experience. The internship period was divided into two parts: First, the theorical part where I have explored all the aforementioned techniques that are available to understand mobility behavior. It includes, Chapter 1 which introduces the traditional passenger counting technologies by giving a clear details regarding the planning and implementation that are essential for counting campaigns. Chapter 2, which explores the household travel surveys based on french experience in that field, explains the theoretical and technical approaches of this technique. Chapter 3 introduces the potential of using multi-agent technologies in context of mobility and transportation, thus a clear approaches are discussed.This chapter also introduce a state of art of several multi-agent works made in response to mobility needs. In addition, Chapter 4 presents the AFC systems stating the importance of transiting from traditional ticketing systems to smart cards. Moreover, the AFC data meet all the conditions of an optimal modeling process based on agents approaches. According to the type of data that can be collected using AFC systems (passengers/vehicles location and time) and since the agents technology requires that the system includes entities that interact between each other, the AFC can be a proper example for the implementation of such technology. This chapter focuses on the theoretical approach of the AFC systems, thus it is an introduction of Chapter 7. Chapter 5 and 6 both discuss the smartphone-based data collection. Chapter 5 focuses on the data collection process in which all the essential information are highlighted. Finally, Chapter 6, which is a complement of chapter 5, it analyzes the acceptance theories related to mobile services and with respect to the new innovative ideas. In the second part of my internship in LVMT, I have realized Chapter 7 which describe several steps that were followed to achieve significant results based on a case of study. Using AFC data of greater Paris (also called Ile-de-France) and after the new flat fare system that took place in this region, my mission was to make a pre-post evaluation of the mobility of transit users in Paris. All the achievements and findings are discussed in details from the first stage to the last findings (See Giraldo, 2016). The data was provided by the organization authority of transport in Greater Paris, namely, STIF and has been explored and analyzed using R language and SQL which are very efficient for Big data. As for other analysis, I have used SPSS and Microsoft Excel.

Chapter 1

Passenger Counting

1.1 Summary

If the data are big in terms of tracking and analyzing of vehicles traffic - and to a lesser extent of the bike, the situation is much more problematic with regard to passenger mobility. Indeed, few data are available on passenger mobility, despite the presence of several passenger counting methods and permanent emergence of new technologies. Some initiatives are often set by local authorities or organizations in favor of active modes, but they remain insufficient to sustainably and significantly incorporate walking into urban mobility policies. This lack of precise quantification can be explained either by an ignorance of global mobility and passenger behavior or by the difficulties we face during the planning and implementation of counting campaigns (Boyle, 1998). Nonetheless, there is a clear intention from public authorities and stakeholders in urban mobility to better understand the issues related to passenger mobility and monitoring the tools that help on interpreting it. In Europe, projects such as Walk21 and Passengers' QualityNeeds (COST, 2010) designed to identify the needs of passenger mobility and develop systemic models that deal with these needs and help to improve the knowledge of the sector in this regard. The purpose of this chapter is not only to introduce the traditional passenger counting methods to identify, describe and illustrate various tools capable of supporting a measure of mobility, but it also explores the theoretical approach of such methods. Thus, it introduces the traditional and innovative instruments that are useful for a counting campaign and an introduction of their technical approach.

1.2 Planning and Implementation of Counting

Passenger counting is a survey used to measure the attendance of passengers in a specific location of study. Counting is designed to learn more about attending a specific place, or even a complete network (Boyle, 1998). Thus, it is an unavoidable tool for mobility managers who need to enhance their network or just know it better. The effective integration of passengers in urban mobility policies requires a good comprehension of their mobility. The passenger must be regarded as a full-fledged user, with his behavior and specific needs. As with other types of users, it is important to measure the use of public space by the passengers from encrypted data and statistical reports. To do so, hereafter several issues are needed, particularly include:

- Passengers' flows are a crucial business requirement, a strong indicator of the market potential;
- Passengers' reserved space must be adapted to the passengers' flows and their behavior (counts and observations). i.e. the width of sidewalks must indeed take into account the current passenger flows;
- Each project and impact assessment should include an essential analysis of passengers' needs. Passengers' flow counting and behavior observing will be a basic requirement;
- Measuring stipulates that passengers should no longer represent an unknown value among numbers and statistics. It is also necessary to measure the impact, follow and correct the actions emanating from the passenger Plan.

As a part of project restructuring, it has become necessary to collect such data to objectively assess the impacts (positive or negative) on passengers' mobility (Boyle, 1998). For instance, to assess the quality of a public space, we address the transit functions of the passengers' movements inside their living districts and a number of optional features. The measure of mobility can be done on the basis of several levels, where the most important levels are as follows:

- **Level 1.** Data are collected when security issues appear in a specific location;
- **Level 2.** Passenger-related surveys are conducted (counting, etc.) on the key routes and are used to improve the network.
- **Level 3.** Systematic counting campaigns on the other hand provide data (exact date, time and location) for the whole network and are used to improve the network particularly at the level of districts. Mobility representative household surveys take place at least every 10 years and are used to formalize the broad guidelines of the passenger scheme. Before-after assessments are conducted for projects of a certain size.
- **Level 4.** Permanent counting stations (automatic) provide continuous data for all the entire network and all over the year. Mobility representative household surveys and special analyses of passengers' movements are made regularly. Some passengers profiles are deducted (demographic, functional, leisure, etc.).

Achieving passenger counting allows a partial response to these needs and provide policy makers with new evaluation criteria for the establishment of coherent urban mobility policies (Boyle, 1998). Indeed, to achieve counting campaigns is not limited to a simple statement of passengers flow on a given location. The variety and the amount of data that can be collected during such a study can power many reflections on urban mobility, and respond quantitatively to several objectives, including:

- Measures of urban space occupation;
- measuring the evolution of movements under the active modes across a zone or a given territory (e.g. mobility observatory)

- Measurement of the presence in the transit
- Reflection and justification of financial measures or/and urban policies
- Evaluation of travel density before-after constructions
- Evaluation of the impact of a campaign to promote active modes
- Analysis of safety, exposure to the risk of passengers and cyclists at a given location
- Give priority to the projects of development
- Development of predictive models of passenger density
- Evaluation of the commercial potential of a district
- Reflection and adaptation of passenger accessibility standards

TABLE 1.1: Objectives of the research

Assessment of the situation (Assessment)	Controlling	Comparative analysis (Benchmarking)
Analysis of the current situation. Provides a description of the current status and new ideas. • Analysis of a specific situation or a project • Useful information for planning and development • Answer to specific research questions	Comparison of the current situation or the one intending to achieve, with regard to certain references. Provides information about the degree of success or compliance of the project with respect to the strategic objectives. • Set of objectives, goals e.g. urban strategy • Standards / references or other admitted principles • People expectations	Comparison between different places, or scales. Provides a ranking or a systematic description of best practices. • Horizontally: among different cities or locations of the same city • Vertically: between different geographical scales, different statues of the project or in a long term

In order to be effective and in line with the objective, the analysis must be pre-planned and properly implemented. In practice, an analysis of travel planning program starts with the (1) A prior definition of the objectives of the program. We also need to (2) The identification of the available resources (both economic and technical), then (3) The selection of the locations, time and duration of counting campaign. Finally, (4) An inventory training of the existing counting techniques that will allow the selection of the most appropriate ones for the desired objectives.

1.2.1 Planning

Planning counting is essential to ensure the proper implementation of the objectives of the analysis. Indeed, it is possible to quickly perform manual counting of passenger or bike traffic flows or installing automated equipment. The analysis of the planning will also ensure more relevant and complete results in the long term (Boyle, 1998). As mentioned in section 2.1, the planning of a counting program can be decomposed into four major phases. Each of these phases is essential to ensure the good proceedings of counting campaign and the quality of the results with regard to its objectives.

(1) Definition of the program objectives: why do we do a passenger counting campaign and for which purpose (s)?

The definition of the objectives of the program should enable the Organization in charge of counting to clearly identify the purpose of the analysis, and consequently to identify the locations and periods of counting, the type of materials used, etc. A clear prior definition of the counting program objectives will encourage relevant results. On the other hand, it is important to identify the data we want to collect because all the counting techniques do not measure the same parameters (Boyle, 1998). When conducting a study of passenger counting, we can extract the following information (Karim Ismail, 2009):

- Flow and intensity of passenger travel;
- Flow Direction of the passenger's mobility;
- Characteristics of the users;
- Behavior of the users (e.g. dangerous crossing);
- The usage of urban space - Heatmaps.
- Passenger routes;
- Detection of other travel modes (e.g. cycling);
- ...

(2) Identification of the available resources for the counting data collection: What are the available economic and technical resources?

It is important to determine the available resources upstream of the implementation of a passenger counting program. In one hand, it is necessary to pay attention to the availability of qualified persons who can achieve the counting, and in the other hand the availability of budget for the collection and analysis of data, or even to train the persons to using materials if necessary. Each counting method requires the use of temporary and financial resources, and there may be significant variation between these methods that should be taken into account during the benchmarking.

(3) Selection of the locations and the counting period: one or more locations of counting; short or extended period; one-off or repeated measures (Schneider, 2005).

Passenger counting surveys can be carried out over short periods of time i.e., from a few hours to a few weeks or longer i.e., several months and in more or less limited number of places. Short-term counting reflects the spatial variation of passenger activity at a particular location (e.g. flow per hour) while continuous counting allows a more comprehensive estimation of passenger activity on a network or within a given territory (e.g. daily or seasonal flow). The use of data provided by continuous counting can give the possibility to adjust the results of counting rapidly.

Furthermore, it is also important to define first the frequency of counting. For permanent and automated counting, the frequency is continuous.

In other cases, the frequency can vary from one to several times per year. There is no recommended passenger counting frequency in the literature, but on the basis of the program objectives, it is appropriate to perform 2 to 3 counting per year (e.g. measure of the mobility evolution by active modes, measurement of the impact of a promotional campaign, evaluation pre-post development, etc.).

Regarding the location of the counting process, four types of approaches can be distinguished: random location, representative location of a territory or a community, targeted location, and controlled location. The choice of the location depends enormously on the program objectives.

As for the **Random location**, locations are selected randomly in a specific area of the network (or even sometimes among a sample of locations whose characteristics are compatible with the counting that we intend to achieve). This approach is sometimes used to achieve the counting which will be used for the development of predictive models of passenger flow. The **Representative location** selects locations whose characteristics and passenger traffic are representative of the territory where they are located. Thus, the implementation of passenger counting at representative locations gives more general conclusions on the whole territory, for example, to measure the evolution of the walking practices, or the risks that passengers face. In general, for any given counting campaigns, several representative locations should be selected so that all the characteristics as well as types of the users and different environments that make up the study area can be illustrated. Concerning **targeted location**, there is no need to represent a territory since the aim just to study a particular project, or locations whose characteristics or functions are particularly sought. For example, during the pre and post evaluation of a new urban project or the study made in high accidents concentration area. Finally, the **controlled location** is suitable for measuring the impact of a project on a given location by comparing it with other locations that have never been studied in a project. For example, we can use the controlled location to measure the impact of a promotion campaign of active modes in a distinct/selected territory, comparing the collected data of that territory with the data of locations outside of it.
(4) Study of the available methods for counting: what are the available methods for the organization authority and those existing on the market?

Several counting techniques are available in the market, and each of them provides services and different results. It is therefore important to select the well-adapted technique to reach the aim of the target counting program, the suitable budget, the available time as well to the urban environment in the areas where we want to conduct the counting campaign. There are manual and automated counting methods. Among these, there is the counting by means of active and passive infra-red, Video, laser, thermal or the radio counts, etc.

1.2.2 Implementation

After the planning phase, the implementation of an analysis is needed. Depending on the counting method used, the implementation procedure is more or less uninteresting. It, typically, involves a series of tasks, including the obtaining of permissions, the training of persons, the installation, the calibration and maintenance of the materials, or even the recovery and analysis of data. This phase generally involves a series of tasks, including:

- Obtaining permissions;
- Purchase of materials for counting;
- Creation of an inventory and preparation of the equipment;
- Training of persons;
- Installation and test/validation of the materials;
- Calibration of the materials;
- Maintenance;
- Recovery, monitoring and processing of data.

1.3 Data Collection

Once the research question is clearly identified, kind of information to be collected is determined. The use of the counting campaign is not the sole way to collect data. Other quantitative or qualitative methods have to be considered. The choice of the counting method should be done according to the collected data, the resources available for it and the temporality of the investigation. The different methods are presented in section 1.5. The data that can be measured using counting are grouped into six categories **Figure** 1.1:

Flow/Volume	The number of passengers in a given time is recorded. This parameter is usually expressed by the number of passengers per hour (P/h). This measure, can help in computing the density of passengers. As counting covers a specific area or section (e.g. the number of passengers crossing a line), this will be expressed with passenger per square meter (P/m2) or passenger per meter per minute (P/m * min).
Direction	Some materials offer the possibility to differentiate the direction of movement of passengers. Entering/leaving a place can also be counted.
The usage of the urban space	Passengers are not always in move and it may be interesting to know the number of people who stay in public space, as well as other parameters such as the duration of stay for example.
Speed	With regard to the passengers, it may seem weird to measure the speed. However, this data can sometimes be useful. For example to differentiate between the different modes using a path or, at night, to switch on and measure public lighting on a way relatively with low frequency.
Routes / itineraries	Some systems allow the map-matching of the visited routes by passengers.

Characteristics	During manual counting, in the field or on video recording, we can also note certain individual characteristics: estimated age category (e.g. 0-14, 15-54, 55 + years old), gender, handicap (use of materials to move, such as Chair, Scooter, cane, etc.), and other modes similar to walking (includes skateboard, two wheel, scooters, etc.). These elements have to be clearly defined before starting the counting. On the field, the number of categories that can be taken into account is limited.

TABLE 1.2: List of relevant data collected

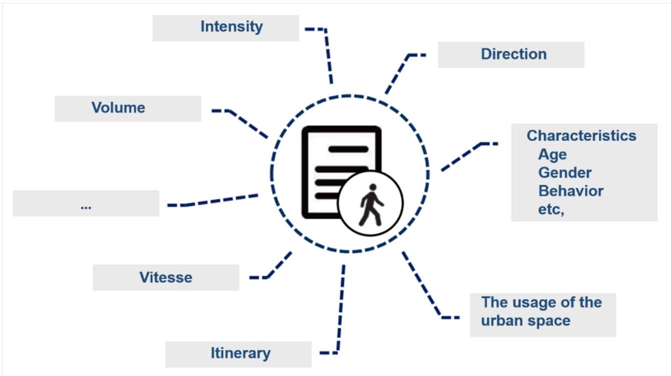


FIGURE 1.1: The data collected by the counting (Karim Ismail, 2009)

In addition to the collection of different types of information, the methods also have some differences, summarized below in terms of advantages and disadvantages.

Distinction between users	Not all the methods offer the possibility to differentiate between the users: passengers, cyclists, materials linked to vehicles, or even motorized vehicles.
Discretion	Some systems are hidden. They prevent the vandalism and do not influence the behavior of passengers. However, it is recommended to always ensure their discretion when they are installed.
Sensitivity to external conditions	The sensitivity to external conditions: located at two levels. On one hand, the measures may be distorted by certain conditions (e.g. sunlight, darkness). On the other hand, some systems are weaker than others and this requires frequent maintenance.
Data protection	Often, the data collected do not allow identifying passengers. However, some techniques theoretically offer this possibility. As a result, we have to ensure the respect of the users' privacy.
Cost	The cost of several methods has decreased in recent years; a trend that will likely continue. In addition, there is a sort of variability in the price between the different models. That is why, only approximations are given indicatively.

The reach	The reach of the automatic systems is technically limited, but at least it can cover the width of a sidewalk. Some situations require the coverage of a large area. In this case, some methods are more appropriate than others, either due to the large reach of a sensor or the installation of multiple sensors side by side.

TABLE 1.3: Advantages and limitations of passenger counting tools

Finally, the surveys can be done according to different temporalities:

Punctual	Counts are made at a specific time of the year
Periodical	Counts are conducted at regular intervals (e.g. every year, at the beginning of spring, in summer, then in autumn).
Permanent	The automatic counting system is installed for a long time, or even permanently, in the same place.

TABLE 1.4: Advantages and limitations of passenger counting tools

1.4 Instrumentation

All the counting methods can meet the objectives of a counting program, but the specific characteristics of each of them makes each of each one of them make each method more suitable for specific studies than the other one. During the planning of a passenger counting campaign, it is important to select the most appropriate method, taking into account the collected data **Figure** 1.1, objectives, the location in which surveys are conducted, etc. Among the most commonly used methods of counting, two major categories can be identified: manual and automated counting methods. In this section, these two categories will be briefly described in order to introduce their particularities and their scope.

1.4.1 Manual counting

Manual counts are operated directly in place, by one person or more (operators) and during a limited time. Data are collected visually and recorded in a form by the operators. Due to its ease of implementation, the manual counting method is often used (Poo Kuan Hoong and Weng, 2015). This kind of methods is generally recommended for organizations with a low financial means, lacking expertise to advanced techniques, or even having been unable to receive the requested permissions to deploy automated counts on the public spaces. These methods are in every case necessary for the calibration of automated methods and the computing of any adjustment factors. In addition, by using manual counts, the characteristics of the user's own (personal) data can be determined, which is not always possible with an automated counting. For example, determination of users type, age group, or even their behavior: use of a phone, crossing outside the crosswalks, etc. Nevertheless, it is important to limit the amount of the collected data to ensure accuracy and quality (Hugh Louch and Sam Piper, 2016). The latter strongly depend on the expertise of the operators during a manual counting. Thus, during the planning phase of the counting program, organizations need to ensure pre-training for the staff properly. During this training, it is important to recall the framework and objectives of the passenger counting program, precisely define data to be collected, and if is applicable to the data collection form and how to fulfill it. The choice of the form is also important. A form must be sufficiently clear and easy to use by the operators. But it must also contain all of the necessary information to ensure the accuracy of the results during the interpretation of the collected data. Technically speaking, there are electronic tablets to easily implement a manual counting (Poo Kuan Hoong and Weng, 2015). In general, these are designed to counting over the intersections and allow the operators to pick up information about the movements in all directions without too much difficulty. These tablets are developed for the counting of vehicles traffic, but they are also suitable for passenger counting. However, it is not possible to identify data related to the users' characteristics, such as age or gender. The only data identified during the counting, is the flow of passengers. A simple mechanical counter can be used (see Figure 1.2). It has the advantage of being very cheap; however, it offers no trace-ability and no further information. In general, such counters are used for very short duration counts, or even to calibrate automated counting equipment.

FIGURE 1.2: Example of manual meter
Source: http://usinenouvelle.com/expo/comptage-pietons-o3053.html)

1.4.2 Automated Counting

Automated counting gathers all the methods and materials which perform counts using a technological assistance. It can refer to counts made using a video support (Lefloch, 2007), IR sensors, thermal or laser, etc. Compared to the manual counting methods, the automated ones generally offer a greater precision, or at least, a better consistency in the level of accuracy over time (Poo Kuan Hoong and Weng, 2015).

In addition, the automated counting methods require greater resources with regard to the material cost, the costs of planning and implementation (Poo Kuan Hoong and Weng, 2015). On the contrary, the hourly and training costs are significantly lower than manual methods. However, in terms of the measured data, the automated counting equipments are not able to identify the characteristics of passengers, such as age, gender or (sometimes) behavior. Nevertheless, they can conduct surveys of longer duration and in several periods of the day (low light conditions may affect the accuracy of the counting for a certain video materials) (Poo Kuan Hoong and Weng, 2015). But, they can translate into more accurate terms passengers' activity at a given location i.e. an estimation of the daily or seasonal flows.

Recent research has shown that we should perform manual counts continuously for at least 4-7 days to get an almost significant estimation of an annual passenger flow. Therefore, surveys of relevant annual average passenger flows are possible with automated methods (Hugh Louch and Sam Piper, 2016).

On the other hand, the influence of exceptional weather conditions or activity (or over activity) in a location in which the flows measurement is taking place should be taken into account (Poo Kuan Hoong and Weng, 2015). There are adjustment factors that correct those errors. Those adjustments depend on each city, or even each district and should be defined in advance by the Organization in charge of the counting. An adjustment factor reflects the increase or decrease percentage of the users' flow compared to "normal". Even in a normal situation, the calibration of automated counting materials is essential in order to maximize the accuracy of the collected data (Poo Kuan Hoong and Weng, 2015). As aforementioned, we can achieve this calibration can be achieved using a manual counting method which can define a basic correction factor by simply calculating the ratio between the flow measured on site using manual counting and the flow measured by automated counting.

In contrast to the manual counts that offer flexibility, all automated counting methods cannot be implemented in all types of locations (Hugh Louch and Sam Piper, 2016). For example, infrared sensors should be placed on a wall or anything else to measure across the flow of passengers who may interrupt the signal. Therefore, there will be no presence of vehicle traffic or cycle track in the background that will mislead the results. In addition, one of the most important elements to take into consideration is while selecting an automated counting method is the width of the desired detection area. It must be adapted to the urban environment in order to limit losses (passengers passing outside the area) or false detection (objects crossing the area - those who are not considered to be passengers) (Hugh Louch and Sam Piper, 2016). Each automated counting method has its own zone of detection and it is therefore essential to consider it while selecting a particular technology.

Objective	Keyword	Question	Examples	Data	Methods
Evaluation	Performance	Is there any intervention needed?	Level of street comfort Accessibility (e.g. children, disabled persons) Attractiveness of a place	Flow, Individual characteristics, Space usage	Manual counter / Tablet, Form, Video

Evaluation	Prioritization	Where should we intervene in priority?	Development of the passenger network	Number, Routes	Manual counter / Tablet, Form, Infrared, Wi-Fi, Bluetooth
Evaluation	Comparison between modes	What is the part of passenger compared with other modes?	Evaluation of the surface allocated to each mode compared to the flow of traffic	Number	Form, Video
Control	Security	To what extent is this place safe for passengers?	Level of attendance at one street by night	Number	Manual counter / Tablet, Infrared, Video
Control	Results	Has the project reached the desired goals?	Enlargement Effects of a sidewalk	Number	Manual counter /Tablet, Infrared
Analysis	Comparison between locations	What is the attendance at this place in comparison to other places?	Level of attendance at two parks	Number, Space usage	Infra-red, Video
Analysis	Comparison between time periods	How does the walk evolve in this place?	Daily profiles Annual /Weekly Profiles	Number	Infra-red, Manual counter / Tablet

TABLE 1.5: Which method for which need? (Hugh Louch and Sam Piper, 2016)

1.5 Technical approach of passenger counting

The technologies used for the measurement of passenger mobility are the same technologies dedicated to the measurement of vehicle traffic. Moreover, the methodology and the conditions are different. The location must be chosen carefully. Depending on the models, material can be fixed on the edge of the road (on one or both sides) or at higher level. None of the elements must interfere with the sensors. Some passengers' behaviors, such as extended stays, can distort the measures. It is important to carefully consider the place where the counter will be installed (Schneider, 2005).

The choice of the counter depends on what we want to measure, both the resources available and the nature of the place either at the level of its attendance or the place where the system installation could be possible (Boyle, 1998). When the location is selected, any automatic system must be calibrated to estimate a margin of error (under or overvalued). To do this, it is necessary to carry out manual counting.

Some models can easily be moved and thus allow to achieve counts in several places using a single device. Data that can be collected will vary from one system to the other. In addition to the essential information such as the date and time, some models measure also other variables such as temperature, humidity, etc., thus avoiding having to use other databases when interpreting.

The models available in the market are generally coupled with analytical software. The latter offers more or less possibilities, which will have an influence on the price. The data are, most often, systematically transmitted to a server. If not, they will be stored in the internal memory of the device.

The following **Table** 1.6 summarizes technologies available for counting and observation of passengers' travel .

Technology	Data	Temporality	Advantages	Disadvantages
Active infrared	Number, Direction, Entering/Leaving aggregated, Speed	Permanent, Periodical	Ease of installation, Low Energy Consumption, Insensitive to external conditions, High precision, Distinction of users, Discretion/Hidden, Anonymous, Low cost	Mean precision, Not always possible to differentiate between users
Passive infrared	Number, Direction, Entering / Leaving aggregated	Permanent, Periodical	Ease of installation, Low Energy Consumption, Discretion / Hidden, High precision, Low cost, Anonymous	Sensitive to external conditions, Not always possible to differentiate between users

TIME LAPSE	Occupation of space	Permanent, Periodical	Visual overview of the use of urban space, Flexible and mobile, Very simple to use, Low cost	No quantified data on passenger flows, Observation at short term, Installation in up level from the ground to avoid any risk of theft or damage, and increase the photographed surface area, Frequent location visit to change the memory cards, batteries, etc.
Radar	Number, Direction, Entering/Leaving aggregated, Speed	Permanent, Periodical	Discretion/Hidden, Low energy consumption, High precision, Anonymous, Users distinction, Low cost	More delicate installation
Video (analysis software)	Number, Direction, Occupation of space, Entering/Leaving, Characteristic, Speed, Routes	Permanent, Periodical	High precision Distinction of users	More delicate installation, Sensitive to external conditions, High Energy consumption, Mandatory Guarantee data protection, High cost
(Piezoelectric)	Number, Direction, Entering/Leaving aggregated	Permanent, Periodical	Low energy consumption, Discretion/Hidden, High precision Anonymous	No distinction between users, (Intrusive) most delicate installation, High cost

Laser scanner	Number, Direction, Occupation of space, Entering/Leaving aggregated, Speed	Permanent, Periodical	High precision, Insensitive to outdoor conditions, Distinction between users, Large area detection Anonymous	More delicate installation, High cost
Wi-Fi, Bluetooth	Number, Direction, Occupation of space, Routes, Entering/Leaving, Characteristic, Speed	Permanent, Periodical	High precision Bluetooth / Wi-Fi, Far-reaching (Wi-Fi), Distinction between users, Low cost medium	Mandatory policies for data protection, Detection only if Bluetooth / Wi-Fi is available

TABLE 1.6: Which method for which need? (Hugh Louch and Sam Piper, 2016), (Schneider, 2005)

Chapter 2

Household Travel Surveys (CERTU Standard Method)

2.1 Summary

Household surveys have become one of the main sources of data to explain social phenomena in the last 60 to 70 years. They are among the most flexible data collection methods. In theory, almost any subject in relation to the population can be analyzed through household surveys. Thus, using households as secondary sampling units is something common in the context of most geographical sampling strategies. In addition, the operators have less work to do and more time is reserved for the collection of data and most of the issues can be well discussed. Furthermore, as it is not necessary to have so many operators in place, it is possible to hire more qualified and intensive persons.

The transportation systems can be developed by means of a good and general understanding of the elements that make up the system. Many years ago, transport organizations have understood that these systems are highly correlated to their users' needs and a good comprehension of passengers' behavior can alone permit a development of such system. Travel surveys are a classic methods used to collect answers for a specific problematic matter that can be solved statistically with further treatments.

Recently with the diversification of transportation means, mobility has become a key issue from a social, economic and environmental perspective. Thus, getting more knowledge and better understanding of mobility remain essential to implement new public policies that satisfy citizens' needs. To observe the mobility behavior over a territory, several types of surveys are held. Yet, the aim of a certain kind of surveys is to understand the households' mobility and the use both public and private means of transport. They are intended to describe all the movements, regardless of the motive, distance travelled, duration, mode of transport, date and time. To understand the mobility-related behaviors, these kind of surveys may also treat the possibilities of access to transit and the available means of private transport in the households.

These surveys are used for collecting socio-demographic detailed data that varied depending on the population conditions, their well-being, the activities in which they are engaged and their demographic characteristics as well as the cultural elements that affect both behavior and socio-economic changes. With their advantage of being flexible, the household surveys are a great way to collect specific needed data for the compilation of statistics that would be otherwise difficult to obtain.

The purpose of this chapter is to provide an overview to the existent travel surveys according to a French research center about Network , Transport and Urbanism, namely, CERTU. The latter explores the methodologies used for data collection and analysis according to each type of surveys, the data exploitation and the studies conducted in France.

2.2 Household Travel Surveys Typology

In France, household surveys have shown a rapid evolution and several methodologies have been developed to fit the socio-economic changes. These methods differ according to several elements while the most important is sampling. Sampling levels concern the population of study.Some of them are national while others are regional belonging to cities/agglomerations. Based on the CERTU standard method, among the highly used surveys in France are EMD , EDVM, EDGT , EGT;(CERTU, 2004a) (CERTU, 2004b) In an international level, the concept and methodologies differ according to lifestyle.

2.2.1 The French "Enquête Ménages Déplacements" (EMD)

The household travel surveys aim to collect the travel practices of urban citizens. This type of surveys differs from surveys that focus on only one mode, such as surveys over transits or surveys set up in roadsides for traffic data collection for identifying all the movements of respondents regardless of the travel mode used, including cycling and walking (CERTU, 2004a). By the means of these global mobility practices, they used to be an essential tool for the development and evaluation of the transport policies in urban areas.

At a local level, the household travel survey concerns all the population categories and take into account all the transport modes. Compared to a number of surveys or ad hoc and sectoral studies, it has the advantage of allowing a comprehensive and coherent vision. Used alone, it can give an accurate overview of using different modes of transport (walking, two-wheelers, public transport, and cars) and their "market share" respectively. Beyond the analysis of the current use of modes, it allows to study the potential customers (CERTU, 2004a). Disclosures given by the EMD allow the development of predictive traffic models needed to evaluate the choice effects with regard to travel policy or/and to transport infrastructure. Performed periodically, The household travel survey would measure the consequences of those choices. It is one of the essential tools in the development of urban transport plans whose results can power a variety of studies such as:

- Customer analysis of a transit system for an operator, an organizing authority;
- Relationships between transportation and urbanism for a planning agency;
- Traffic flows for a local authority;
- Shopping-based travel for chamber of commerce and industry, etc.

At the national level, all the household travel surveys that were realized in French urban areas are considered as a real "Database" that permits the monitoring of the activities, the practices and habits of daily trips and the opinions of residents on their transportation system for about thirty years.

Compliance with the "standard" methodology used by CERTU in France can address two items:

- The behavior evolution over time "big trends";
- Comparison between agglomerations, distinguishing what emerges from these "big trends" and what the consequence of local transport policies is.

Thus, the household travel surveys have shown, for example:

- The relative stability of daily mobility "all modes";
- Prolonged and brutal disaffection of two-wheelers;
- The decrease of walking practices in the city;
- Continued growth of car use.

The household travel surveys are, for local organizations and research institutions, a database used for powering numerous studies and reflections in the field of sociology of transport, mobility, activities and the operating of the city. The building of such "Database" allows each agglomeration to know its position in relation to others. This kind of survey should establish a database to be used for several years for the social service. It is important for the urban mobility as the general census of population is for demographic surveys.

In 1975, SETRA has defined the practical steps for achieving household surveys (the word travel was added later); In 1989, CETUR edited a methodological file updating and completing the first version; in 1998, “Centre d’études sur les réseaux de transport et l’urbanisme” (CERTU) published in its collection "References", " L’enquête ménages déplacements (EMD) « méthode standard Certu »: the household travel survey, CERTU standard method". Being strictly regulated and controlled this method provides objective, reliable and comparable data in time or in space. It also proposes methodological adjustments according to the changes that occurred in lifestyles and enables better integration of new concerns such as intermodal (CERTU, 2008).

While SETRA focused on the use of these surveys to develop generative models of traffic used in transport infrastructure studies, their utility goes out from this single framework. Today, the planning of large urban transport infrastructure, the development of urban transport plans established by “Loi d’Orientation des Transports Intérieurs” (LOTI) in 1982 made it compulsory for cities of over 100 000 citizens by the rational use of energy (Treil, 2005). In 2000 new law was introduced namely "Loi relative à la solidarité et au renouvellement urbains", the law could not be held without reliable and accurate understanding of the population needs, current travel patterns, its recent evolution and determinants. The need for a rebalancing of different transport modes usage, followed by the law, requires the consideration of all modes of travel, including cycling and walking.

The CERTU standard method meet those requirements by offering specific techniques of data collection, ensuring reliable results and the collection of travel information.

To highlight the evolutions, the policies' evaluation required the use of a tool that is both static over time to ensure comparability of results and constantly evolving to adapt to changes in lifestyle.

Particular features

The CERTU Standard methodology, which sets the rules for conducting the household travel surveys, is based on a few key principles.

- They are conducted in the surveyed households and interviewed by trained persons for this type of data collection (CERTU, 2004b).

- All persons, of at least 5 years old, living in the household are interviewed personally (children under eight can be helped by their parents). This is one of the major constraints of the survey because all the households with at least one active person can only be surveyed in the evening, between 18 pm and around 20 pm.

- All travels, made the day before the survey day, by each surveyed person, are identified. The characteristics of these travels, their motives, the transport modes, the locations, time of departure and arrival give them a precise and complex definition that must be respected to allow comparability between surveys. Thus, all modes of transport are studied, including walking.

- They cover a representative sample of households in the surveyed study area. This sample is drawn randomly by area of residence and in a housing file. The designated housing must be the primary residence of the household. The sample size is determined to ensure a minimum reliability of results for sectoral analysis.

- The CERTU Standard questionnaire includes a form Household, a form Person, a form Travel and a form Opinion. The latter is done only with a person who is at least 16 years old randomly selected. To this questionnaire, a local party, with a maximum duration of 15 minutes. These additional questions can be inserted either in the different forms of the questionnaire or in a specific questionnaire.

Data collection

The aim of data collection is to obtain both quantitative and qualitative data about travel and mobility of the persons living in the households. Firstly, the collected data can provide reliable informative items. It will therefore be possible to conduct statistical calculations on accuracy and errors in the results. This aspect requires a rigorous sampling (mandatory use of random selection of households and individuals to be investigated) and suitable sample size. Today, it is difficult to determine the sample size (this depends on the exact details wanted and the wanted territorial divisions to be used) but for comparison, and order to ensure sufficient accuracy of the results, travel surveys are carried out on samples of at least 2000 persons in medium-sized cities (less than 100 000 citizens) and from 2000 to more than 5000 households (approximately 4600 household with more than 11,500 persons) for large territory surveys.

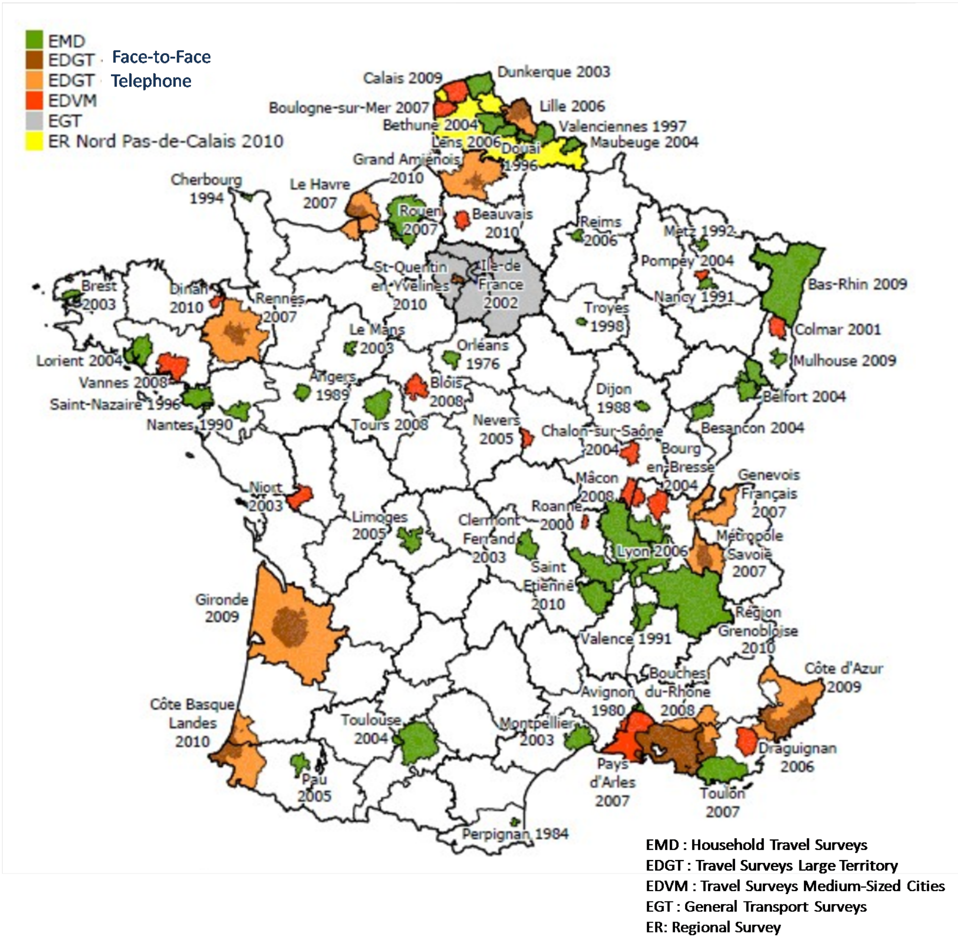


FIGURE 2.1: The "Household Travel Surveys, CERTU standard method" Source: ORT PACA 2011 - Enquête Ménages et Mobilité Urbaine

Secondly, this data collection will also provide information on other different subjects:

- Characteristics of households and the persons surveyed,
- Mobility, transport modes, reasons for travel (e.g school, work etc),
- The flow including origins and destinations.

It is therefore a tool to understand behavior but also to estimate traffic flows. Thirdly, it will will cover all trips of the area of study, including short trips made within the suburban area or the city of residence. This will provide a comprehensive and fair view of the mobility. However, it is clear that the amount of detail required for travel can vary according to each travel type. The issue of relations between the urban periphery and its urban center justifies a need for precision in locating the origins and destinations in the description or the succeeding modes used. Such description of the movements is no longer necessary in the context of internal mobility in the district of residence.

To perform the interviews of the surveys, the CERTU has implemented different methods: interviews modes, face-to-face, telephone, mail, etc. The method imposed for each survey follows specific criteria as well as a sampling manner. In statistics, the interviews by sampling could produce significant errors if not all the population but only a sample is studied. These errors namely, sampling errors, should be reduced.

The household travel surveys EMD are performed through face-to-face interviews. However, the travel surveys in Medium-sized cities (EDVM) and Large territory (EDGT) are done via telephone. The reason behind the use of face-to-face interviews in the case of EMD is because it is not possible to collect all the information existing in the forms and questionnaires. In addition, if the contract file of land-line phone members is not used in urban zones with high density for sampling only because it is incomplete. In the recent surveys, the non-subscribers to landline phones constitute more than 20% of households. For example, in Lille a northern city of France, the number of non-subscribers went from 6% in 1998 to 22% in 2006. The process of data collection can be performed via face-to-face, telephone or Email method. Each method has its advantages and drawbacks (See Table 2.1).

Face-to-Face Interviews

As aforementioned, the face-to-face surveys are carried out by means of 4 forms/questionnaires. The "household" questionnaire is given to one person only who will provide information about the housing and the private vehicle of the surveyed household, if exist. The "person" questionnaire is given to each individual -aged at least 5 years old- in the household. The latter identifies all the resident of the household and their demographic characteristics. It also allows having information on occupations/jobs and mobility practices of the surveyed persons (CERTU, 2004b). The "travel" questionnaire collect all the trips that have been done by the interviewed individuals, including the origins and destinations information, the duration, the mode used and the location. In this questionnaire, movements are also divided into trips. A route corresponds to the use of a mechanized mode. As for the "opinion" questionnaire, any person aged 16 and more per household is randomly selected to respond to this questionnaire. This person gives an opinion on the life in the city and the different modes of transport. Questionnaires can evolve at the level of the formulation issues or the terms of responses from one year to the other or from one territory to another.

Telephone and Mail Interviews

To proceed, the operators use the land-line phone book to randomly choose a sample. The mobile phones can also be added to the sampling. In the near future, mobile phones will be used alone since around 90% of the populations currently have a mobile phone (CERTU, 2010)(CERTU, 2004b).

As for the mailing method, the surveyed persons are asked to fill in the forms at any time and send them back to a specific address. This method is more efficient in case of a mutual interest. The surveyed person should accept replying to a huge number of questions (CERTU, 2010)(CERTU, 2004b)
.

TABLE 2.1: Advantages and limitations of interviews processing methods

	Face-to-face	Telephone	Mail
Advantages	• A long and complete questionnaires • Multiple feedback for the same interview from other persons. • Confirming the information via any available devices • Full control of the interview • Computer Assisted Personal Interviewing • Possibility of outdoor interviews	• Rapid • Low-cost • Possibility of CATI (computer assisted telephony) • Time optimization	• Weak logistics • More liberty given to surveyed persons • Possibility of complement of information • Possibility of long forms • Appreciated in case of mutual interest with regard to the study.
Limitations	• With appointment only • The absence of persons • Interview duration • High-cost	• Simple and short questionnaires (5 to 20 minutes/interview) • Low-control of the surveyed person (The control is required specially for the short trips • High non-response rate (Decline of the calls, absence of the person, Black list)	• Low responses rate • Attractive form required to prevent missing Information due to miss-understandings • Long durations • Unknown surveyed person

2.2.2 The French "Enquêtes Déplacements Villes Moyennes" (EDVM)

The French travel surveys of medium-sized cities are an indispensable tool to know reliably and accurately the needs of the population, travel patterns, recent trends and determinants, and assess the development of new transport policies. They concern all the population categories and take into account all the transport modes. They provide a comprehensive and coherent vision of the practice of different modes of travel and a study of potential customers. Disclosures given by these surveys allow the development of predictive traffic models , basically essential to evaluate "a priori" the effects of choice in term of travel policy or (and) transport infrastructure. Temporarily carried out, they measure "a posteriori" the consequences of those choices (CERTU, 2004b) (CERTU, 2004a).

The results of a travel survey of medium-sized cities can also contribute to the analysis of the following:

- The user of a transit system (transit organization authority)
- The relations between transport and urban planning (planning agency)
- The aspirations of people for transport (elected)
- The flow of traffic (local community)
- The shopping-related travel (Chamber of Commerce and Industry)
- etc.

As the household travel surveys "CERTU standard" are for large cities, these travel surveys of medium-sized cities are made according to a standardized method which allows to establish "Memory" to monitor activities, practices and habits of residents' daily trips and the views on travel in their respective city.

Similarly to the household travel surveys " CERTU standard", designed for the needs of large cities, these travel surveys "medium-sized cities", are also made according to a standardized method which allow the creation of a real "Memory" to monitor activities, practices and habits of daily trips and the views of residents on transportation system in their city (CERTU, 2004a).

Compliance with a standardized methodology offer possibility to evaluate behavior changes over time and compare in-between agglomerations, distinguishing what emerges from the "big trends" of what is the consequence of local transport policies.

These surveys are, for the central agencies of the Department of Transportation and research institutions, a single database for powering numerous studies and reflections in the field of transport sociology, movements, activities and, in a word, the operation of the city. Finally, the constitution of such "memory" allows each agglomeration to position itself in relation to others. The standardized methodology is based on the following few key principles:

- The surveys are conducted mainly by telephone (when possible) or face-to-face in the household.
- One or two persons (depending on the ousehold size), aged at least 11 years, living in the house are personally interviewed.
- All travels, made the day prior to the survey day, by each surveyed person are identified. The characteristics of these travels, their motives, the transport modes, the locations, time of departure and arrival give them a precise and complex definition that must be respected in order to allow comparability between surveys. Thus, all modes of transport are studied, including walking.
- They cover a representative sample of households in the surveyed study area. This sample is selected randomly by area of residence, in house file. The designated house must be the primary residence of the household. The sample size is determined to ensure a minimum reliability of results for sectoral analysis.

- The "CERTU Standard" questionnaire includes the following forms: household, person, travel and opinion. (The latter is given to only a person who is at least 16 years old)

The "CERTU standard" questionnaire forms are described as follows:

- The "household" form is related to the household and house characteristics.
- The "person" form essentially includes questions on socio-economic characteristics of each interviewee.
- The "travel" form is designed for 11 years old persons designated in the household. It lists all the movements of the day prior to the survey day
- The "opinion" form concerns the randomly selected persons in the household who are at least 16 years old.

2.2.3 The French "Enquêtes Déplacements Grands Territoires" (EDGT)

The travel surveys "large territories" methods deal with the peripheral areas (suburbs) of large cities where the " CERTU standard» method becomes inappropriate because of the logistical difficulties (extent of the territories), the challenges and needs it encounters- often non-existent in larger urban centers (CERTU, 2004b).

The characteristics of these travel surveys are the same as those of "medium-sized cities" surveys (CERTU, 2004b). The only major difference is that the interviews of this type of surveys are exclusively made via telephone. The unreachable households by phone are replaced, using a statistically correct method. This choice does not affect the efficiency of the survey. It takes into account the issues of such territory, different from those in more dense urban areas. Indeed, it was concluded - following the study of household travel survey results which were carried out on large areas, urban and sub-urban - that the omission of the unreachable populations altered somewhat the overall results. First, the land-line telephone subscribers who are not present in the phone book have similar mobility behaviors as the rest of the subscribers (CERTU, 2004a). On the other hand, the number of households that are not equipped with landline telephones and residing in the peri-urban area is low. These households, given their socio-demographic characteristics and their mobility behavior, do not constitute a key issue of transport policies: they are less mobile, and are mostly in their commune of residence. The issues are rather common in relation with peri-urban centers, primary or secondary, of the surveyed area (CERTU, 2004a).

2.2.4 The French “Enquêtes Global du Transport“ (EGT)

Global Transport Surveys are a large-scale surveys. It is about forty years that the Parisians mobility is studied via these surveys. Five surveys are already conducted after the population censuses: 1976, 1983, 1991, 2001 and 2010. As part of the urban transport plans evaluation in greater Paris, led by the transport organization authority of greater Paris (STIF), EGT 2010, co-led by the regional and inter-departmental directorate of materials and development (DRIEA, ex Dreif) and STIF, was conducted: face-to-face interviews were made with 18,000 Parisian households. The sample was extracted from a housing file census databases, completed by the households built after the census file. Ended in 2011, the survey was made according to a random drawing in the file of the General Directorate of building taxes. People living in mobile homes or in community (workers' hostels, university residences, retirement houses, detention centers, etc.) are excluded from the scope. Based on the sample, all Parisians aged 6 years or more are asked about their mobility. The questionnaire focuses on three types of information:

- General household characteristics: its location, income, number of residents and assets that constitute it; its engine (number of passenger vehicles, two-wheelers, non-motorized, night parking of these vehicles...);
- The individual's characteristics inside the household: age, sex, profession, job location, work position, etc. Children under six were less considered and were not interviewed because they generally make few moves and are always accompanied;
- The movements of individuals: each household member, aged six and more are surveyed about all the trips he/she has makes the day of the survey. For each trip, a number of important features are collected: Trip motive, mode (s) of transport, origin, destination, etc.

2.2.5 Household Travel Surveys Worldwide

The household travel surveys are a worldwide matter. In this section, Belgian and Canadian experiences are added to the French expertise in the treated field to briefly introduce the key differences among the countries (CERTU, 2004b).

The Belgian "National Travel Survey"

To study the household mobility, beligum states, the Walloon region, Flanders and Brussels-the Capital- have conducted several national travel surveys such as Beldam, Mobel. The methodology, co-led by the National Institute of Statistics, the Transporation Research Group of Notre-Dame-de-la-Paix - Namur university and the Antwerp university - was as following:

The sample is drawn from the National Register which list all the residences, the communes as well as the household composition. Once the housing sample is made (by random selection with equal probabilities stratified by geographic area), a search of telephone numbers in directories is conducted.

Traditionally, the basic method for collecting data is the use of a self-administered postal questionnaire which gives possibility to set longer questionnaires compared to the phone, and it can be sent to all sampled houses (having or not a landline phone). Postal questionnaires are sent by successive packets. In fact, the day for travel is given in the questionnaire (the respondent knows the reference day in advance, which may motivates him to pay more attention to daily activities). To ensure a sufficient feedback rate and control the postal interactions, important phone calls are implemented as well as a hotline to help respondents in case of problems. For all sample households, a first contact is made by sending a letter announcing the survey. For households without phone line, the feedbacks are made through mails. For households reachable by phone, first a mail is sent including the questionnaire and then it is followed by a phone call. The call is made before the survey day to check if the questionnaires were received and motivated the household. A call after the survey day is made to revive the operation and another call is made after one week to boost the operation. Among those who answer the calls, a telephone survey is done to validate their responses (the returned back questionnaires are read to check for errors. Thus, validation by phone will be necessary). Among those who do not respond, phone surveys of "no response" are made, i.e. a partial questionnaire to identify non-respondents is designed.

The questionnaire of the survey is divided into two parts: a "household" questionnaire and an "person" questionnaire about the mobility of people aged 6 years or more. These questionnaires are constructed according to the same logic as the household travel survey questionnaires CERTU standard method. In addition, to usual questions related to travel survey (as done in France) , other questions are set about the private vehicles (precise description starting from the brand to each component), residential mobility , long-distance trips made over the previous month, the physical difficulty to move and to travel to work or to school. These are treated separately from other movements. Another feature is that the movements are described with precision of in-between locations and not by origin or destination.

In 2001, the Walloon Region decided to launch a new survey on household mobility (Walloon regional survey, ERMM), project management ensured by the Research Group on Transport (FUNDP). Compared to Mobel, the main methodological innovation is the permanency of the survey : each year, a sample of households in the Walloon region is surveyed. The advantages of such permanent survey are multiple:

- Financially speaking, it can reduce expenses by dividing it through several years;
- It can, within a few years (3-5 years), reach significant samples of households surveyed;
- It can set up a precise monitoring evolution of the mobility behavior over time.

In contrary to Mobel, which focuses on the number of households contacted, the methodology attaches a big importance to the number of completed and validated questionnaires. Besides, the Walloon regional survey equally follows the methodological principles of Mobel:

- A sample randomly drawn with geographical stratification from the National Register,
- The interviews based on a self-administered postal questionnaire;
- The establishment of a mixed postal method and a phone questionnaire (motivation, re-launch, validation) for households reachable by

phone; for other households, the survey is made only by post (questionnaire and relaunch by post)

- A questionnaire divided into two parts: a "household" questionnaire and a "individual" questionnaire about the trips of persons aged 6 or more; the questionnaires used are very similar to those used for MOBEL.

One last feature to highlight on the Walloon regional survey is that the Walloon region intends to join this permanent survey to each of the five Walloon provinces: each year, a province would provide additional funding which would then oversample this province in the regional survey.

The Canadian "Montreal Travel Survey"

The survey covers a very large area (3,500 kilometers, more than 3 millions citizens) that corresponds to labor and residence of the Greater Montreal and aims to provide the data necessary for the management and planning of transport networks of the agglomeration. Such management is done using a particular model of movement referred to as MADITUC.

This model is a phone-based survey. The sampling rate is significant, reaching 5% of households. The sample of households is derived from a random sample which is stratified by a specific geographic area from a household billing file of Bell Canada, a Canadian telecommunication company. The random selection is used to remove 7.14% of households; given the refusal and rejection rates, this yields a final sample investigated of 5% of households given 65 000 households. Information and awareness campaigns are carried out by means of media, radio, television and advertising campaigns. All the selected households of the survey area are informed by mail about the investigation start. In addition, each selected household receives a letter a week before the phone call warning about the phone-based survey. A phone number is sent within the mail in case of any additional information and also an overview of the questions typologies. To achieve the 65 000 inquiries, a group of 45 to 50 operators are set in addition to the supervisor.

The survey is simple and it includes the following information:

2.3 Data Processing in Household Surveys

This section, dedicated to data processing in household surveys, first describes the regular survey cycle and analysis that can be carried on for household travel surveys. The recent and rapid development of the information systems has had a profound impact on methods design and application for statistical data treatment systems. Regarding the hardware, the main new element is the transition from central systems to personal computers which are increasingly powerful especially in term of memory speed. Personal computers can now perform statistical operations that vary from small-scale surveys to large operations like censuses of the population and surveys of many household samples. Along with the evolution of hardware, software quality used for processing, analysis and dissemination of statistical data has improved as well; as a result, demographers can

TABLE 2.2: Objectives of the research

Household	Person	Travel
• Survey-Day • Number of persons in the household, • Number of private cars in the household , • Household address.	• Rank of the person in the household, • Age, • Gender, • Characteristics of the car.	• Departure time, motive, • Destination, • Return home time, • Transport mode • If multimodal travel: address of the junction point, • If TC: TC lines used, up to 5 lines (which gives the TC route),and title of transport used, • If Private Car: • Nature of parking (free, paid by the employer, paid by the respondent, etc.) • Car Route (main roads).

now perform tasks that were once reserved to computer scientists. In recent years, several software dedicated to survey statistical data processing tasks have been developed. The advantages of each of these programs depend on the work at hand.

2.3.1 Regular Survey Cycle

Basically, all the surveys follow the same cycle; the steps are usually as follows:

(a) Survey planning: The survey designers should have clear decisions related to the main objectives and results usage, expected information to be collected, the procedures for collecting data (design and preparation of the questionnaire and related survey instruments like sheets) and data transformation into products as well as the design of the data processing and documentation system (DAES, 2010).

(b) Survey operation: to create the sampling frame, to design and select the sample, to collect data (measure), to prepare the data (input,coding, editing and imputation), to create the observation file (raw database), to make the estimates and in particular to compute the weights, to create the derived variables, analyze data, and present disseminate results (DAES, 2010).

(c) Survey evaluation: At this point, the data are checked and validated, see if the products were appropriately published and distributed and if the meta-data was documented and stored, etc (DAES, 2010).

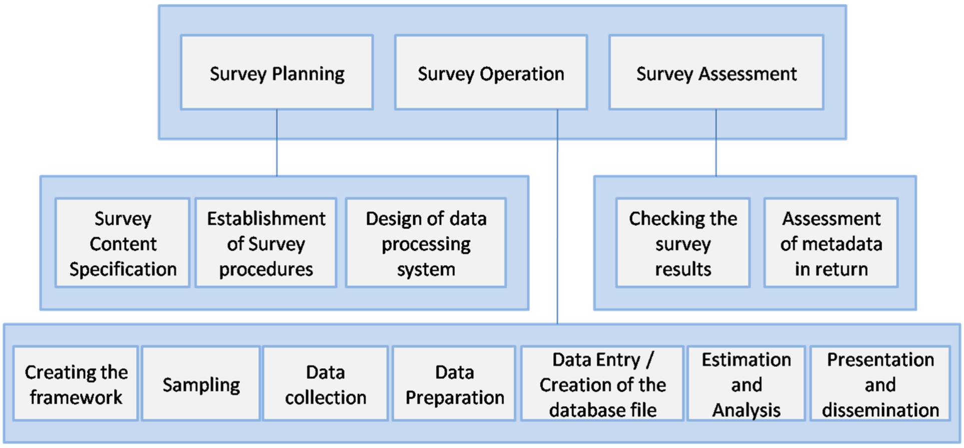


FIGURE 2.2: Household Survey Cycle" Source: Sundgren (1991).

Before designing and implementing the data processing system, it is important to imagine what will be the overall system. It is also important to spread properly the operations and processes. Thus, the objectives of the survey should determine the product design (e.g. tabulation and databases models), which in turn will determine the activities to be performed by the survey design, collection , treatment and processing of data and, ultimately, analysis and dissemination of results (DAES, 2010).

The data processing is considered as the process of transforming the answers obtained from the interviews when collecting data to a desired information which are suitable for tabulation and analysis. Data processing gathers automated as well as manual activities. The process can be long and it requires many resources, and its processing of data directly affects the quality and cost of the final product (DAES, 2010).

2.4 The scope : Particular Studies and Research

In France, several studies were conducted to measure mobility. The household surveys are mainly used to build a general overview of a subject. However, these kinds of surveys can also focus on specific questions: Parking issues , Oil consumption, etc. To study the distance traveled by car, the oil consumption is also a key item to the final results. As discussed in In (CGDD, 2010). For a clear demonstration, an example of France-located studies is resumed.

2.4.1 Travel to work: traveled distances and time lengthening in France

In (CGDD, 2010), by combining surveys' data and a specific statistical analysis, the results have shown considerable changes while comparing two years of study 1994 and 2008. These results were precise and gather very profound details about socio-mobility. Half of the active population are women who have a regular work location. Consequently , transit usage has increased. Besides, the number of active population showed an increase over the peri-urban areas. The use of private cars has also changed and in contrary to the greater Paris region and the big cities center, population favored the private car traveling than transit by achieving 72% in 2008. On the one hand, the traveled distance has also taken considerable part in the changes with an increase of 2.7 Km compared to 1994. On the other hand, traveled time has decreased with 2 minutes. Population avoids more and more the use of several modes of transport and especially in urban areas where the interest for two-wheelers has grown. Finally, the use of transport means differs according to the professions. While the simple employees are using more and more the private car to go to work, the executives and the more intellectual professions usually living in the urban area, use increasingly the transits In (CGDD, 2010).

The results of such study are interesting; the household surveys are mainly used for general conclusions about a specific and particular field. However, according to the data collected, further research in this study can be done. It is predominant to mention that household surveys are limited to people eager to provide information.

Chapter 3

Multi-agent Systems in Context of Mobility

3.1 Summary

Multi-agent approach offers a well suited framework for modeling and analyzing the complex systems. It considers systems as organizations made up of autonomous and independent entities, called agents, which interact to solve a problem or perform a task together. It has been successfully used in many areas and in particular: Robotics, distributed problem solving and modeling and simulation of complex systems. Despite the advantages these systems have demonstrated, multi-agent systems remain weakened and limited by the formalization and modeling methodology. Indeed, the design of such systems is in many cases empirical approach, or on an ad hoc basis. The development of methods of formal specification and design is an important issue of today research on multi-agent systems. The present chapter presents a state of art concerning the use of multi-agent systems technology in the field of transportation.

3.2 Theorical approach

Usually, the two terms 'complex' and 'complicated' are often used without a real distinction. However, in our context, they have different acceptances and it is good to clarify it.

3.2.1 What is a complex system?

The word complexity comes from the Latin plexus (interlacing), complexus (connection) and perplexus (tangled) (Le Moigne, 1999). These three words reflect the importance of the relationship when it comes to the paradigm of complexity. A complex system is : "a set of interacting elements, organized system, more or less open on the environment" (Dauphine, 2003) or even "set of dynamically interacting elements, organized to achieve a particular goal" (De Rosnay, 1997). Usually, we think that the whole of a complex system is greater than the sum of its parts. In other words, the behavior of the system as a whole cannot be determined "by dividing it and by understanding the behavior of each of its parts, when taken separately, which is the classic reductionist strategy of physical sciences" (Amblard, 2006). The complexity is the result of a large number of items in interactions. We consider three types of interactions:

- The internal interactions, i.e. between the various entities of the system;
- Interactions between components of a system and its initial conditions. These conditions can be formalized in the form of states, amount of resources, etc.
- The interactions between a system and its environment. The environment gathers all the elements that encompass a system which influences directly or indirectly, its evolution. Other interaction systems can also be included in the environment.

3.2.2 The complexity in the mobility systems

Since the present thesis focus on the mobility systems, it is more appropriate to define what a mobility system is, and explain why it should be considered to be complex. This will allow us to justify the efficiency of agent-based modeling. A mobility system is made up of several entities interacting between each other, more or less organized on the basis of realizing a goal. It often works in a well-defined environment: the serving territory, it is in constant interaction with other societal systems such as the political system, economic, social system, etc. These systems play a relatively important role and we strive not to overlook their impacts. However, it is very difficult to take into account all the parameters that influence a mobility system. In this chapter, we will explore the different components of the mobility system, we distinguish in three subsystems: the demand for mobility, the supply of transport, each of them being more or less complex (see Figure 3.1).

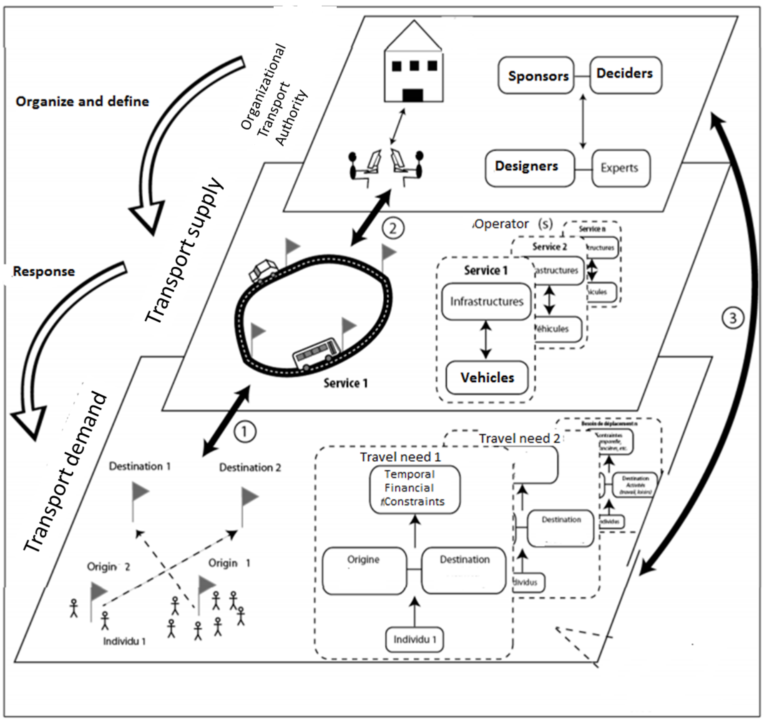


FIGURE 3.1: Levels of organization and interactions in a system of mobility.

The demand for mobility

The demand for mobility includes all individuals living on a territory which have travel needs. These needs will be based on lifestyles, behaviors, financial resources, etc. From a technical point of view, we can characterize these needs to travel by three key variables: the starting point (origin), the point of arrival (destination) and decision-making constraints of individuals. These constraints are varied in nature. Certainly, we should not neglect other parameters such as comfort, safety, environmental impact, etc. they can play an important role in the decision-making process. The demand for mobility is certainly the most complex subsystem, as it involves a very large number of individuals interacting between each other. The demand for mobility is the most unpredictable part of the system, forcing the companies to invest more time and money to investigate the characteristics and behaviour of their citizens. Household Travel Surveys are a good example of such studies implemented by communities to better understand the needs of travel on a territory. More frequent and precise, these surveys are an indispensable tool for the planning of urban transport (CERTU, 2004b).

Transport supply

Transport supply includes all services and operators. It is simply a sort of an answer to the demand for mobility. A transport service is composed of vehicles (cars, buses, metro, etc.), infrastructure (roads, lanes, rail, stops, stations, etc.) and operating principles that are going to be more or less complicated and strict according to the types of service. In France for instance, authorities entrust the management of transport services to operators according to the principle of public service delegation. Each operator is responsible for the smooth running of one or several services. These are usually large companies of transport which manage public transport on one, or even several agglomerations. Transport services are not all necessarily complex in their operation. This is the case of the subway that runs on totally fixed lines and is clocked on schedules relatively precise. Except the case of a particular incident on a line and the drivers respect speed limits, it is possible to accurately determine the location of each vehicle based on the schedules. Which consists the benefits of such a service.

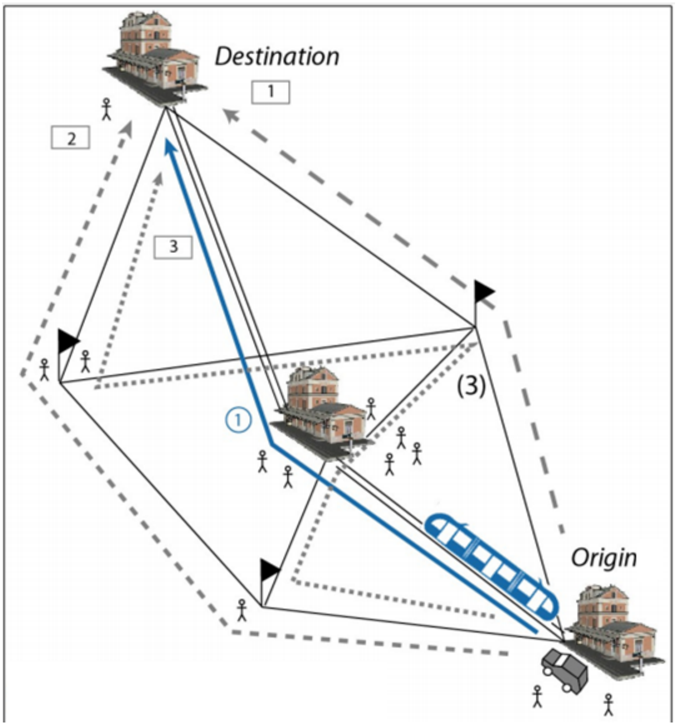


FIGURE 3.2: Levels of complexity that differ according to transport services.

3.3 Systematical approach

Directly related to the paradigm of complexity, systems theory has emerged in the United States after the Second World War through the biologist (Bertalanffy, 1986). However, Jean Brunhes employs the term system from 1910 (Dauphine, 2003). In his book, Bertalanffy sees the traditional science as a reductionist, i.e. it is splitting life in pieces while the living being is complex.

3.3.1 Systems modelling

Systems engineering is a field that "gathers the theorical steps, practical and methodological , related to the study of what is recognized as too complex to be tackled in a reductionist manner and posing problems of borders, of internal and external relationships, of structures, of laws or of emerging properties characterizing the system, or problems of observation mode , of modeling or simulation of a complex totality' (Donnadieu, 2003). Tools used are often systemic charts for which the components of a system are prioritized and represented according to codified forms. Figure 3.3 is an example of a generic model describing the transport system in its environment. The focus is placed on the representation of interactions, whether in the form of self-reinforcing (repeated emphasis of a phenomenon), of regulatory (compensation of different interactions), feedback (actions received in input and dependent from the output), etc. These models are generally static, but they can be represented in a dynamic way on software like Stella (Richmond, 1985) or using the second way of representing the systems which is based on differential equations which are used to calculate the evolution of variables based on time, and therefore to study the dynamics of the system in one time spot. It is also possible to combine these two modes of representation.

Systems modeling and multi-agents paradigm are much related. The multi-agent modeling offers a methodological framework, oriented towards the simulation of systems. In a SMA, the dynamic aspect is intuitively represented and analyzed. However, other methods of modeling such as graphics and digital models are fundamental. These models are even a preliminary step in the design of software development models such as UML (Unified Modeling Language).

The theory of systems

The aim of the general theory of systems "is to formulate valid principles for systems in general regardless of the nature of the elements that compose them and the relationships, forces, that link them" (Bertalanffy, 1993). The goal is not to present all the rules of the theory of systems, because there is too much. We briefly present only those that seem the most relevant to characterize and describe the functioning of observed mobility systems, namely the principles of organization and self-organization, and those of balance and stability.

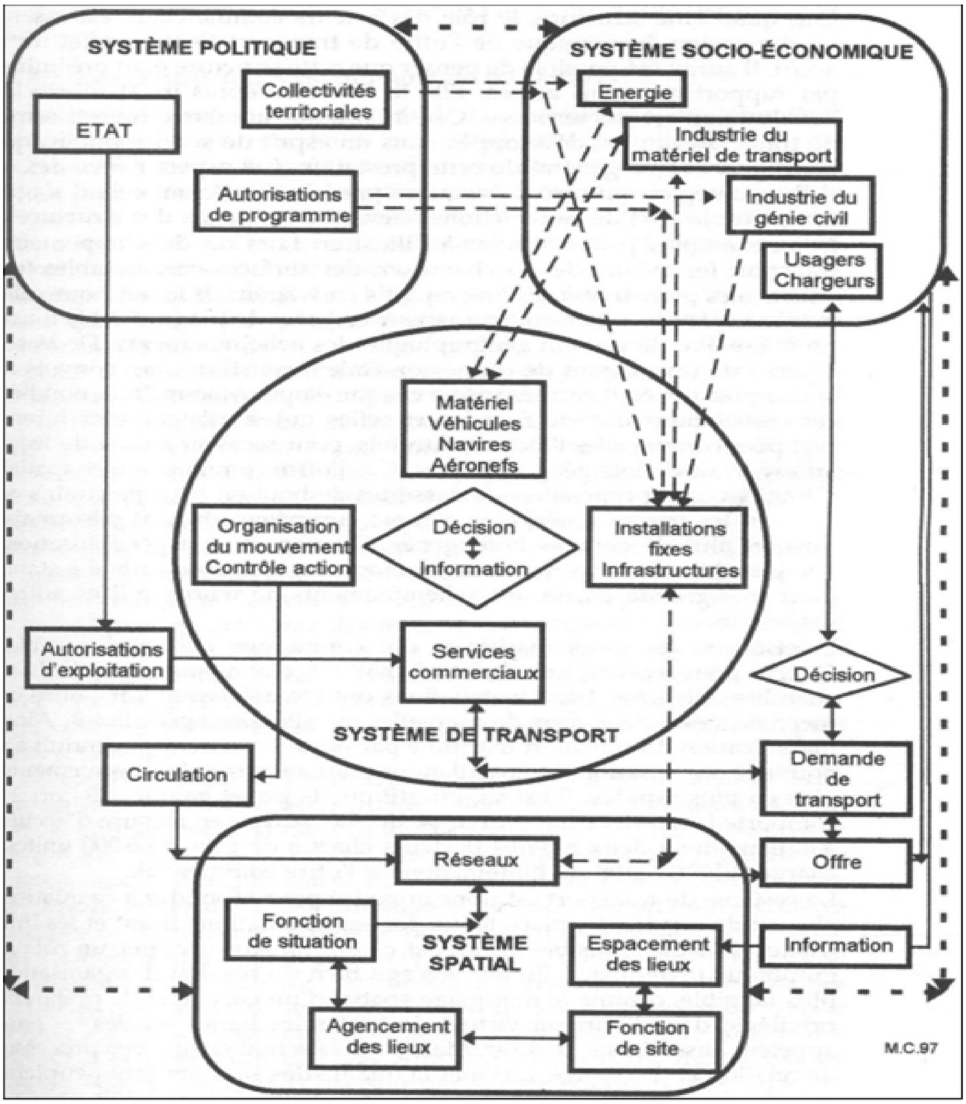


FIGURE 3.3: Modelling of a system of transport and the environment (Marcadon, 1997)

Organization and self-organization

"If the agent symbolizes the individual, the Organization symbolizes the collective" (Langlois, 2010). By definition, a system is a set of elements organized on the basis of a more or less identified goal. However, it is possible to distinguish different forms of organization. "A compartmentalized organization consists of components individualized with specific functions, such as the organs of a human [...] In this case, the creation of a new component or disappearance, changes and calls into question the entire organization. " (Langlois, 2010). Transport services controlled by an entity of highest level are all organized and compartmentalized, systems that each of its component is essential for the functioning of the system. They are hierarchical, interacting, and dependent on each other. For example, a transport on demand with vehicles making optimized tours. If one of the components of the system (booking service, or bus) stops working, all the service is blocked.

In a self-organizing system, each of the elements is autonomous and more or less conscious of the phenomena emerging at the global level (Dupuy, 1994). They may know whether or not there are other entities, similar to them, more or less in interaction. In the field of transport, the autonomous entities can have three modes of functioning. In the first case (i), they operate completely independently. They have a very limited knowledge of the composition of the system, their functioning lowly impact the other entities. This case remains relatively rare. For example, we can imagine a situation when every taxi operates within a well-defined territory. The targeted customers are limited to people wishing to move in the area occupied by the taxi. Taxis may not have a relationship between them, yet global service is organized and each person can move on its territory. In a second case (ii), the entities are autonomous and operate in competition. The functioning of entities has an influence on that of other entities. Examples of competition are frequent in third world countries. In some countries, illegal taxis are a very good example of service in competition yet for self-organizing global operation. They look up to the most profitable areas to serve, regardless if other transport services are already operating in that area. If the competition is too tough, they deploy business strategies or even find another territory. Finally (iii), entities may be self-organizing and cooperatively. Not only they have knowledge of the presence of other entities, but the functioning of an entity facilitates or enables the functioning of another entity. Cooperation is often done by exchanging information between the elements. For example, some taxi services do not hesitate to communicate with each other to organize their service.

Balance and stability

A balanced system maintains its state over time. "The state variables that describe the system are not growing and their derivative with respect to time is zero. This conservation can be static if the relationship between the system and environment are constant, which is exceptional in geography or dynamic when the transformation between the inputs and the outputs is constant" (Dauphine, 2003). Ideally, a transport mode such as bus or subway service is a balanced system. It is designed in the way that every day (based on periods of mobility), it will be able to respond consistently to the demand for mobility. In other words, it is calibrated to satisfy a certain amount of customers in a territory given a time.

3.4 Example of Agent-based Models for the Study of Mobility Systems

3.4.1 The multi-agent modeling

As shown in Figure 3.4, agent-based modeling comes of the microsimulation, modeling by cellular automaton and artificial intelligence (Troitzsch, 1997). These fields of research 'have in common the fact to represent explicitly entities and their behaviors in a computerized form and to study, thanks to the power of today's computers, the emerging structures that are born from virtual interactions" (Treuil, 2008)

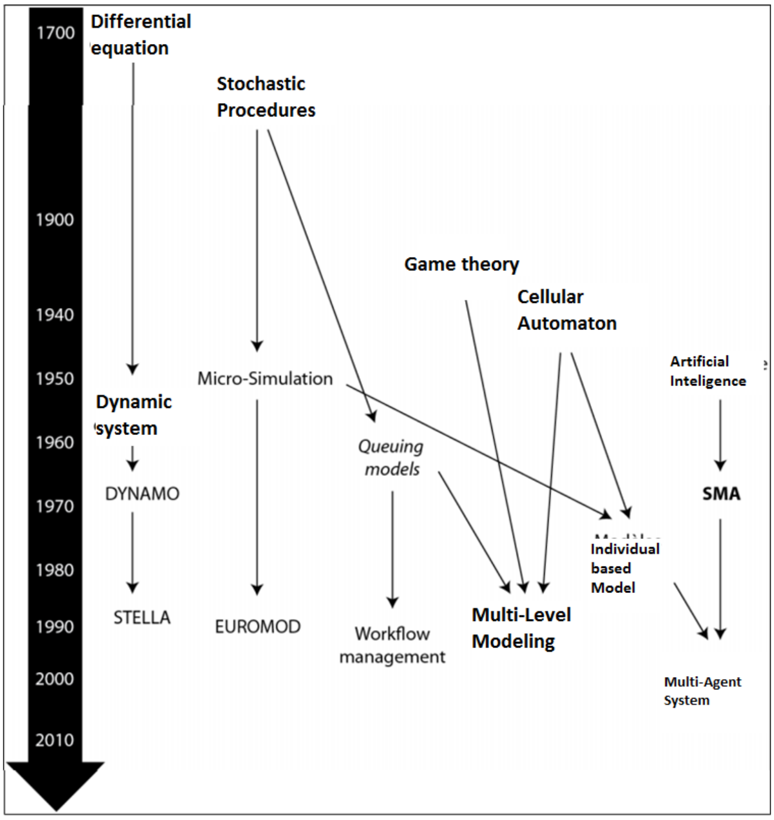


FIGURE 3.4: chronology and development of modeling methods (adapted from Troitzsch, 1997).

Although the current presence of many tools that combine a cellular automata with a multi-agent system, the two methods, even very close, must be distinguished. John Von Neumann introduced the first examples of cellular automata in his book "Theory of Self Reproducing Automata" in 1966 (Von Neumann, 1966). Originally, these, very simple models, consist of computer programs (cells), arranged in sort of grid. Automatons have computing capabilities and they are subject to states which vary during the simulation. It is the combined functioning of a large number of automatons that allow the modeling of phenomena that are considered as complex. The cycle of life that allows simulating very complex forms from simple rules is one of the oldest and most well-known examples (Gardner, 1970). Today, Thomas Schelling segregation model (Schelling 1971) is very popular by geographers and sociologists. It helps in demonstrating that the high levels of residential segregation can appear, even when individuals are willing to have a majority of people of different ethnicity living in their neighborhood. Many works are always based on this model. It's a relatively simple model, however it possesses a real heuristic and teaching qualities. A model is "an abstract construction that allows understanding the functioning of a system in response to a question concerning it. Simplified representation of this system, a model is based on a general theory and it is expressed in a specific language called modeling language". Agent-based modeling aims to understand the functioning of a system by studying the behavior of individuals (agents) and by analyzing the relationship between these individuals, both among themselves, and with the environment in which they operate. Agents are autonomous virtual entities (computer programs) able to be kept or to be maintained (Ferber, 1995). They have attributes, more or less sophisticated behavior and are able to interact directly or indirectly with other entities of the system (see Figure 3.5). They always have a more or less limited knowledge of their environment (capacity of perception) and of their proper resources (Amblard, 2006). Each agent is located and operates in an environment itself defined by its metric, its rules of functioning and by all kinds of attributes that can evolve during the simulation. According to their function within the system and so according to their behaviour, agents may be mobile, responsive to the different signals, or again, they can communicate and learn from their interactions. They are then called cognitive agent.

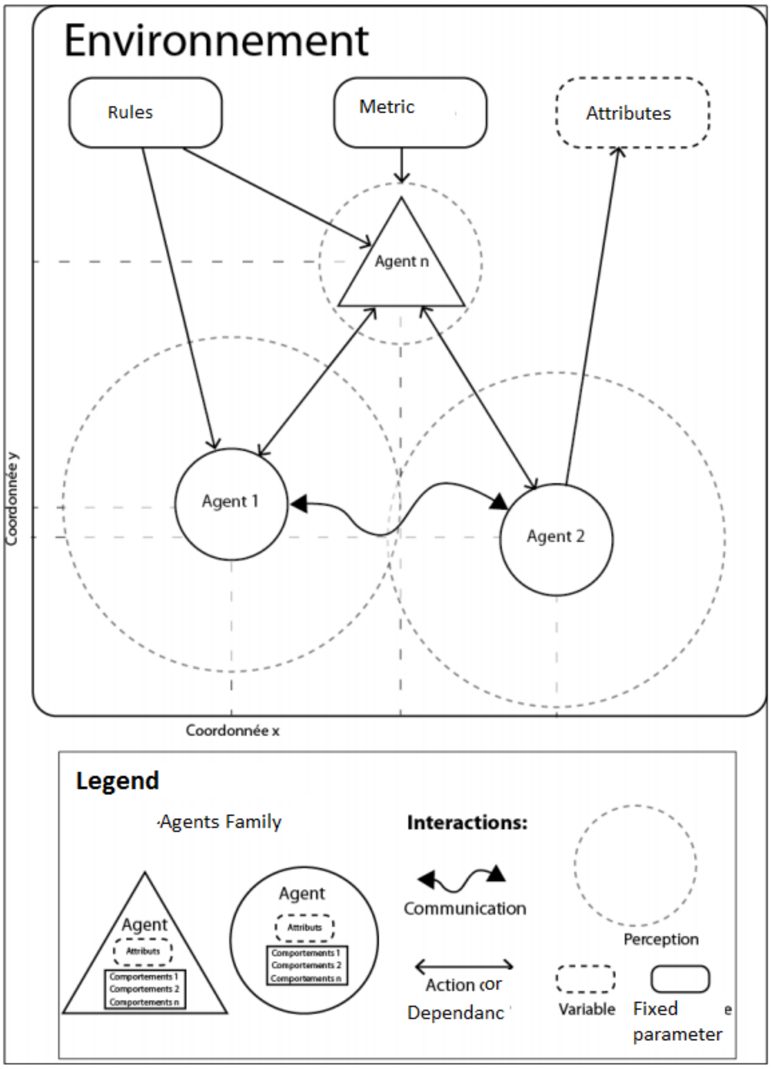


FIGURE 3.5: Scheme of agent-based Model

The multi-agent systems allow to study at a global level, the operation of one (or more) subsystem (s), known at a local level. It is the operation of a large number of agents in interaction leading to the complexity of the process at a global level and the relative unpredictability of the system over a long time. In other words, by modeling behaviors or phenomena at a certain level, we can analyze the organization and structuring of one (or more) phenomenon (s) at a more aggregated observation level (Daude, 2006). The behavior of agents will determine the operation of the system and the simulation allows observing the emerging forms at the macro level. However, an MAS doesn't have to be a black box. In other words, the simulation is not minimized to "the implementation of a model and to the analysis of its response based on the input parameters" (Amblard, 2006). Thus, the importance relies on the understanding of the phenomena. The Modeler must always focus on the relationship between local behaviour and the overall system. Agent-based simulation is much developed in Europe over the past 15 years. Researchers from diverse disciplines are now around the science of complexity. There are a large number of tools available to model and explain systems, and those that we use are also well used in physics, biology, sociology, etc.

Multi-agent systems are today very popular in the Science of human and society. On the one hand, because the two variables which are the (discrete) time and space can be advantageously represented. On the other hand, because the interactions between several entities as well as the overall dynamics of a system can there be intuitively analyzed. The MAS, being widely used in several applications, are particularly well fit to the simulation of mobility. Movement, i.e. the 'change of position in space based on time, with respect to a particular system" (Robert, 2001), is an intrinsic property of the agents (Marilleau, 2006).

3.4.2 Various modeling objectives

Multi-agent systems can be used in different ways and in very different research settings. As we shall see, the complexity levels differ according to the approaches and the models finalties as well. Nevertheless, the logic remains the same: reasoning at the level of agents to analyze structures or emerging solutions at a more aggregated level. Here we present three objectives of multi-agent modeling with models applied to transportation and mobility. The first goal is the prediction of the evolution of the systems, with descriptive models for simulating realistic mobility systems at small and large scales. The second category of models is intended to explain the dynamics of systems. It does not consist in providing a 'valid' forecast, but rather to ask 'why' and 'how' the system evolves (Amouroux, 2011). At a more local level, the last goal of modeling is the optimization of the systems. This category of model is widely used in operational research, particularly in the field of intelligent transport (Intelligent Transport System) systems (Tao, 2007).

Simulate realistic mobility systems

In the cognitive and social sciences, a first category of descriptive models is designed for the prediction of the evolution of the systems. These models may be attached to the principle of KIDS 'Keep It Descriptive Stupid' (Edmonds, 2005). In spatial analysis, for example, the geo-simulation multi-agent (GSMA) now allows to create geographic virtual worlds closest possible of reality, using GIS and relatively complete geographical databases . These works also revealed new types of MAS as the located multi-agent systems. Within the theme of transport and mobility, we find that today several models capable of simulating systems mobility in very precise and close to comprehensiveness exist. Simulation platforms deployed are indeed very sophisticated and researchers work not only as a team, but especially in multidisciplinary (geographer, computer scientist, economists, etc.). The goal is to be able to present these simulators-looking tools and decision-making for transportation planning and urban design. In the worldwide, a multidisciplinary research team has developed the platform simulation MATSim (Balmer, 2009). The tool allows simulating the mobility of a population both at the micro (City) and macro (country) level. The model is composed of several modules that can be combined or used separately. It was originally designed and calibrated using a Swiss census databases, but it seems to be able to adapt to any territory. It has for example been tested in South Africa (Fourie, 2010), on the city of Tel Aviv (Bekhor, 2010), on the greater Toronto (Gao, 2010)), etc. The implemented transport services may be modified, only with a good knowledge in programming and clear vision of the architecture of the model (Balmer, 2009). Recently, an algorithm of optimization of type Dynamic Vehicle Routing Problem (Maciejewski, 2012) was integrated, allowing simulating Transport on demand-type of services. The Simulator turns powerful and relatively accurate for the simulation of traffic Figure 3.6 , or to set scenarios of evolution of mobility in a territory.

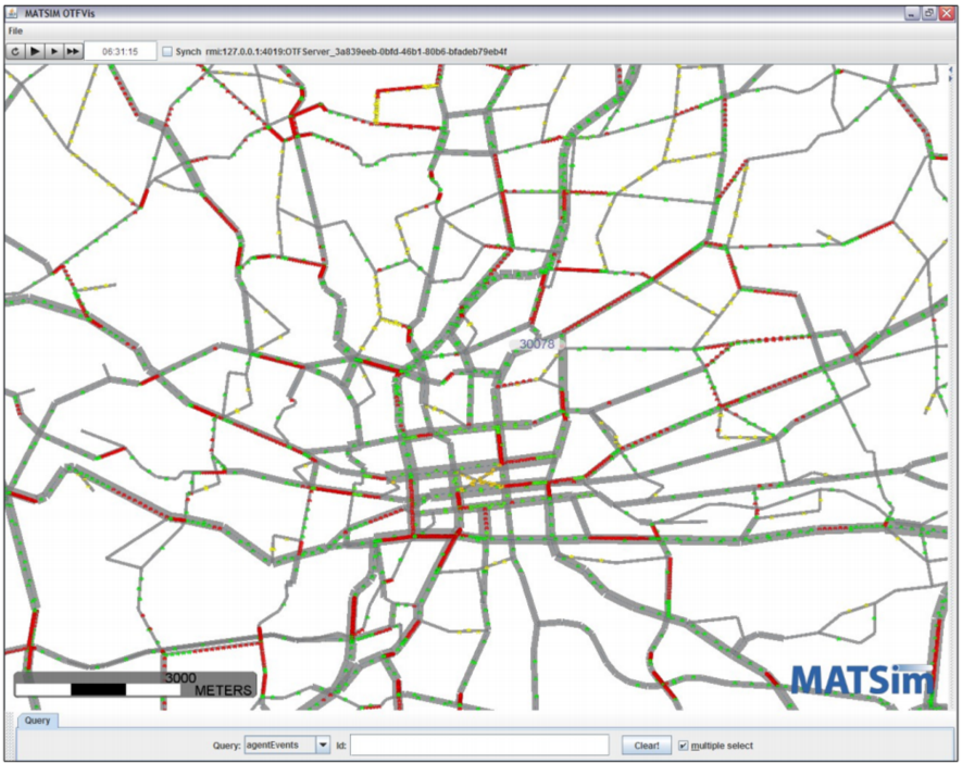


FIGURE 3.6: Simulation of traffic (congestion in red) on the network of Johannesburg (Fourie, 2010)

Understanding and explaining the systems

What characterizes a good model is once its implementation is realized and its behavior is already studied, it is no more useful and the Modeler does not need it again [...] The Modeler, by building a model of its object of study, involves cognitive abilities by selecting the key points to take into account to explain the phenomenon in hand. By designing the model, the modeler discusses some qualitative assumptions about generating mechanisms of the observations on the target system. By simulating this model, the modeler can test these assumptions and discuss them. “The modeler retrieves the properties of its model and the consequences of its assumptions" (Amblard, 2006). " A second category of models is intended, not to the description of realistic scenarios, but in exploring and explaining the dynamics of systems. "Simulation is used to understand the functioning of a particular system, considering the model as a miniature replica that can be studied more easily" (Treuil, 2008). These models are attached to the KISS "Keep It Simple Stupid" approach (Amouroux, 2011). Thus, the aim is to design relatively simple models and so manageable from a point of view architectural and algorithmic to avoid malfunctioning (Axelrod, 1997). "There is no need to design models that we would not seriously studying the properties and thus forget the internal validation, defined as the existence of good properties of the model in the formal setting of the latter" (Amblard, 2006). In other words, we’re looking to keep control of the models to better understand the results of simulation and finely explain the behavior of the system in its environment, as well as its reaction to the variation of some parameters. In such approach, models are designed from theories, rules and assumptions (Dauphine, 2003) and not from comprehensive databases. The challenge is to design simple and controlled operating rules, to follow and understand their impact on the overall functioning of the system. Knowing all the components and all the initial settings of the system, the tool allows simply observing the artificial world and intuitively understanding the overall behavior with structuring effects, of transition and emerging.

Today, several MAS, relatively accessible, offer the possibility of programming this type of model. The NetLogo development environment that we present below (cf. 3.4.3) is a good example. In recent years, this tool became highly used by geographers. The training and thematic schools have multiplied and many research projects are adopting this software. The first model appeared in NetLogo addressing the theme of transport is a simulation model of urban traffic known as the Gridlock (Wilensky, 1999). This model is integrated into the library of models, provided with NetLogo. It is therefore free to access and use. Over the past ten years, this model has largely evolved. It was first adapted to optimize the management of traffic on a road network of type (Gershenson, 2005) in summary, the originality of this model is to offer a self-organizing traffic lights management. These are modeled as agents able to count the number of cars waiting, change state depending on the queue and interact with each other to synchronize. Following these works, a dynamic carpooling service has recently implemented by the GTI group of the University of Vigo (Spain) and the CRP Henri Tudor in Luxembourg (Armendariz, 2010). They therefore retained the original structure of the Gridlock model incorporating a number of new features. The first originality of this work is the integration of an algorithm of optimization for dynamic carpool service. The goal is to combine both models, i.e. to analyse the impact of implementing a car-pooling service on the density of road traffic Figure 3.7. To do this, they have enriched the model by completing the attributes of various agents and with sensory capacities and communication. Thus, drivers of cars and carpool service users (modeled as agents) can exchange on their location of origin and their destination. Depending on the number of places available, the time constraints and their initial journeys, the drivers are then able to choose optimally the passengers to take along the way.

Several models addressing issues of urban mobility, and accessibility exists already in NetLogo. At the micro level, we find the AcceSim model that can simulate the accessibility to shops by the citizens on the scale of a residential area. The principle is simple: individuals move into the city to reach a shops and shops need a regular attendance of individuals to survive. The goal is to show with an educational tool that the spatial distribution of the services or equipment is often uneven and accessibility varies from one territory to the other. To do this, indicators allow to assess continuously and at the level of each district the rate of satisfaction of residents, the average distance to access businesses and finally total sales. Developers have also set up a system of dynamic Figure 3.8 map which shows the satisfaction of various agents in real-time.

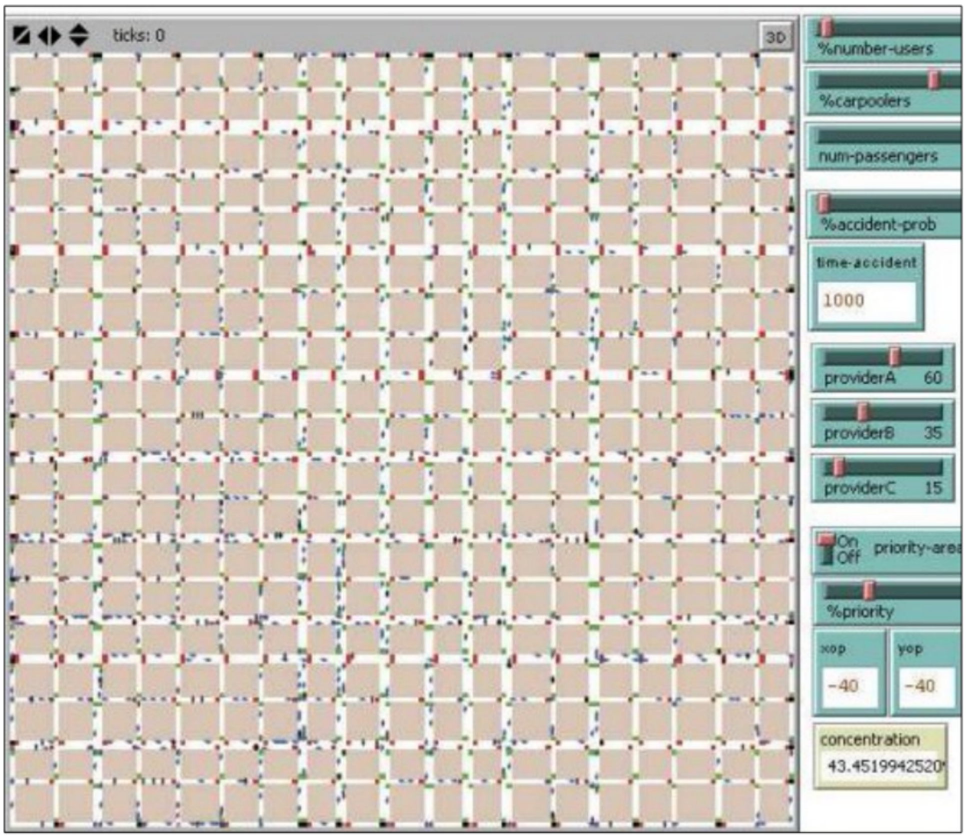


FIGURE 3.7: Screenshot of a simulation with the Dynamic Instant Service Simulator.

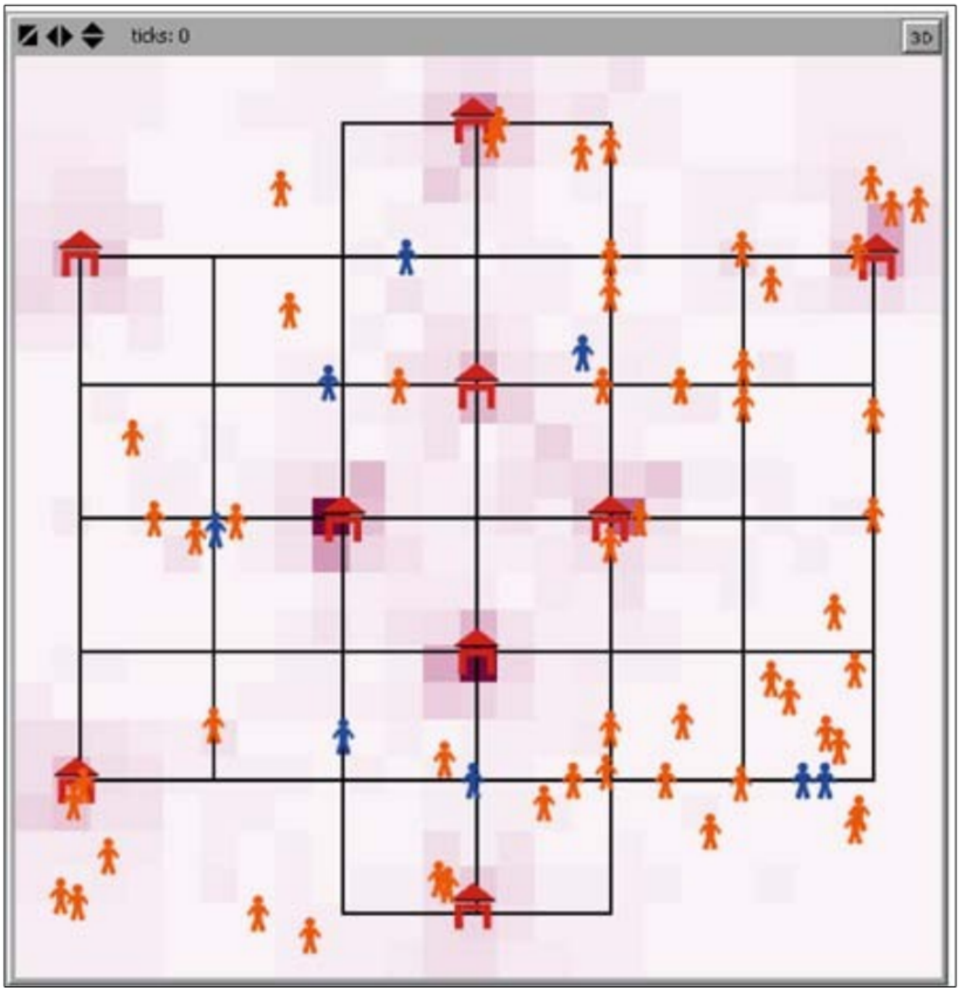


FIGURE 3.8: Example of simulation with model AccesSim (Davis, 2008)

Finally, other models give possibility to address the issue of the transport of people on a small scale. The SimpopNet model aims, for example, to simulate the causal links between transport system and system of cities. The challenge is to simulate a loop of retro-action between the size and accessibility of a city and observe the consequences of this interaction to the higher geographical level. During the simulation Figure 3.9, innovations in transport networks thrive on the territory, resulting in a reduction of the time to transport on the relevant axes (represented by color links). Transportation times are calculated from the distance measured on a real network. This structuring shows of new axes of transport and new nodes that aren't necessarily cities (represented then by crosses). The population growth of cities being dependent of its relative position with respect to other cities in the network, the entire hierarchy of the city system that evolves. The aim of combining these mechanisms is to emerge properties to the system. ' At the level of population growth in the city system, it is expected to see a correlation between the size of a city and its accessibility, or more exactly between its size and its potential for interaction in the system, resulting finally in a strengthening of the hierarchy of city system "(Schmitt 2013).

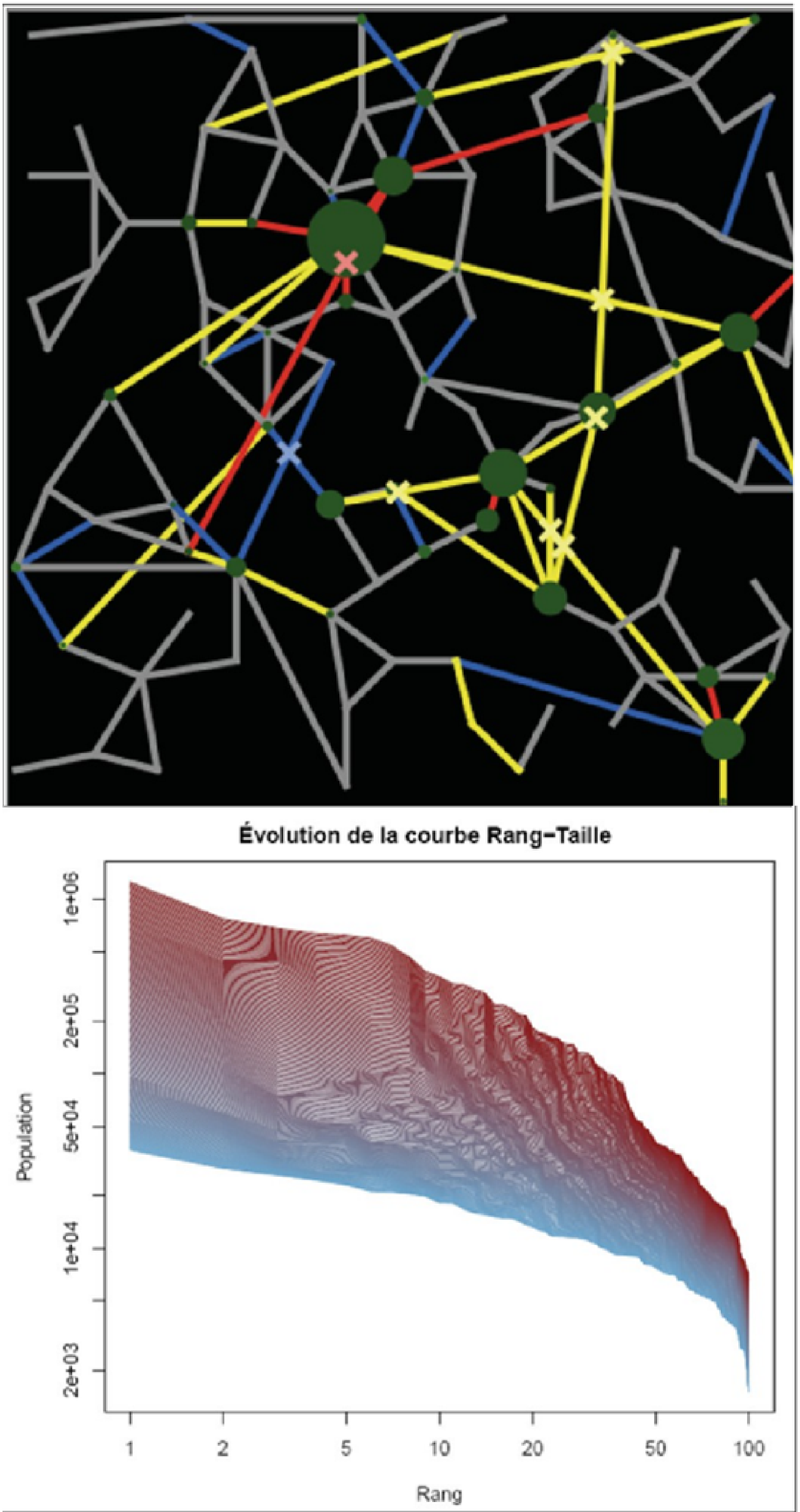


FIGURE 3.9: result of a simulation on 200 years (200 software iterations) with the model SimpopNet (Schmitt, 2013)

Solving problems and optimizing tours

In computer science and robotics, the MAS approach is used to design computing algorithms to solve complex problems. For instance, solving problems (Gechter, 2005) or the resolution of distributed constraints (Vidal, 2006). This field of research has resulted in new forms of MAS such as reactive MAS or the adaptive multi-agent systems (AMAS). Agents, with their cooperation and the coordination of their activities, permit to realize high-level tasks. To optimize transport, many teams in operational research are now using agent-based programming. The principle of distributed computing allows developing innovative and efficient transportation system. Within the theme of intelligent transport, Stephan Winter for example carried out a feasibility study for the creation of a self-organized modern transport system (Winter, 2008). The vehicles are represented in the form of agents. They have sensory abilities, and they can communicate with each other to organize their service. In the same area, Wu et. al used the MAS to develop an optimized carpool services (Wu, 2008). The goal is to offer an effective connection via radio system, between people intending to travel and the vehicles available. The various entities of the system were represented as reactive agents. Always within intelligent transport, several agent-based works have been carried out by French laboratory 'Systèmes et Transports'. Their goal is to set up a traffic control system to optimize intermodality of the transit company in Montbéliard (Fougeres, 2000). Finally, MAS are also widely used to optimize flexible transport systems, including those dynamically optimized based on demand. In Latin America, researchers were able to develop a distributed agent-based system dedicated to the resolution of the "Dial-a-Ride Problem" (Cubillos, 2007). The goal was to calculate optimal vehicle tours based on the constraints given to agents. These methods have been well adapted to transport on demand such as (Zhao, 2011), (Nabaa, 2007), (Zargayouna, 2012). In these articles, the authors offer dynamic optimization algorithms of vehicle tours, booking and vehicle fleet management.

3.4.3 NetLogo simulation platform

There are many platforms for multi-agent modeling and simulation of complex systems such as Swarm (Hiebeler, 1994), Cormas (Bousquet, 1998), Repast (North, 2006), GAMA (Amouroux, 2009), etc. NetLogo was created by Uri Wilensky (Wilensky, 1999) and continues to be developed within the Center for Connected Learning and Computer-Based Modeling at Northwestern University. The software is dedicated for the simulation of complex systems for natural, human and social sciences. Models can be of type cellular automata, multi-agent, or both combined. The platform is composed of patches (cells arranged on a grid), Turtles and Links (types of agents) specified by some Breeds (agent family) and an Observer (equivalent to a high level agent) who can read and act on all the variables in the model. NetLogo is particularly suitable for models based on "mobile agents that evolve in parallel inside a grid environment, and whose behavior is based on short interactions" (Treuil, 2008). While the software allows simulating models with thousands of agents, the limits in terms of computing capacity can be quickly achieved with relatively complex models. This was the case during the simulation of some of the models I have developed during my master course about climate change. The main objective of NetLogo is to provide an educational and ease of use tool. The software is free, available for all operating systems and patterns run identically. However, it is not open source. It has a substantial online help and documentation. It is also accompanied by a body of free access patterns, which can serve as conceptual and algorithmic basis for the users. The tool has a simple and intuitive interface. Its programming language dedicated to handle high level and structures of primitives that reduce the modeling effort (Treuil, 2008). Moreover, It has relatively complete functions to run simulations (export of images, videos, data tables, etc.) and a function of batch simulation. NetLogo allows practicing individual-based simulations with a spatial dimension. Even if it does not have all the functions of a Geographic Information System, the tool leaves very appreciable flexibility including through its extension GIS. It is now possible to import geographical layers and integrate them into the simulations which make it quite adapted to simulations in geography.

Chapter 4

Automated Fare Collection System

4.1 Summary

While the Automated Fare Collection Systems' prevalence is expanding out in the transit systems, a substantial amount of information is accumulated day by day, user by user and trip by trip over these systems. In spite of the fact that they are for the most part expected to meet financial purposes, AFC platforms may be a solution for transit systems planners to have a clear view of the users practices under the system, thus it promotes the development of services. The main advantages of those systems are the information that are able to record every tap-in / tap-out in an open transport system. The information gathers the time and location which are the fundamental information for any origin-destination analysis or for travel tracking in general. The datasets collected by AFC systems are low-time requiring and it only requires a good data mining tool and a high computing power machines. We believe that data mining with all its methods may be certainly helpful for exploration and analysis of transportation data.

FIGURE 4.1: Simple AFC system - Paris (France).

4.2 Smart Card Collection System

AFC systems are being adopted by a large number of transit organization authorities throughout the world. The following section aims to present the smart card used by the user to access the AFC systems and why it is useful for collecting data. Likewise, it argues the fundamental reasons about its emerging in the transportation, and points out the advantages and limitations in using AFC in this area. Briefly, AFC uses a smart card that takes a form and a size of credit-card , a rectangular made with plastic, being extremely convenient and solid (Lu, 2007). It has a built-in microchip that represents the head of the card, it stores, processes and writes data. Thus with all these abilities in treating information, the card can exceptionally be used in large scale of areas, however in this chapter we only consider their presence in transportation systems as a source of collecting fares but also as user behavior data collection.

Traditionally, time and location are collected when the user tap-in the card on the reader. The reader is setup either inside the buses or in the stations at the entrance of the metro/train platforms. Although, these cards basic idea was to allow or not the user to travel according to payment issues, they can now be used to understand the user behavior and how the network is going on.

This technology was introduced more than 30 years ago. Whereas, it was not developed to follow a pull strategy but a push strategy only. The strategy brought several failures which did not promote its usage in the transportation industry before 2000 (Barry, 2002). Hong Kong was the first introducing Octopus card, a card that opened a large possibilities to transportation sector and in which it was increasingly implemented.

There are two types of Octopus card: ordinary and customized. An ordinary Octopus Figure 4.2-left does not require presentation of ID card or passport for purchase. If the owner loses it, there are only the recorded value that is lost and no personal information or bank account number or credit card is retrieved.

Customized Octopus or "Nominative" Figure 4.2-right can be used as a key to access some residential buildings and offices. Generally, it is a monthly pass which gathers personal information including a list of permissions. The card can be blocked if lost.

FIGURE 4.2: Andante - ordinary and customized card - Porto (Portugal).

4.3 Comparing Smart Card with Other Fare Collection Systems

Many methodologies are used for fares collection, which is one of the most relevant aspects in determining users' behavior over the public transport.

This section argues the advantages and limitations of the use of smart cards and try to find out , by giving an outline of smart card usage compared to other forms of fares for public transport (e.g. cash, prepayment tickets, monthly passes , etc.), even if each contributes in increasing the standard of performance, especially in the view of passengers experience. This increase in the standard performance is mainly computed or treated in two ways: customers' satisfaction, and effectiveness amelioration with respect to traditional methods. Traditional methods of fares collection can reveal a real insecurity and often cause delays in boarding. The reasons are much, where the two later reasons are the most considerable ones arguing why transit organization authorities began to demonstrate interest in smart card technologies, and many of them are now adopting it to replace the traditional cards or tickets as a primary payment option (Blythe, 2004). Smart cards are seen as a safe technique for passenger identity approval and fare payment and offer another techniques to enhance the collection and validate the fares' payment (Blythe, 2004). It also decrease the delay since the driver is no longer responsible in collecting the fares. Moreover, smart card increases the quality of the data collected with respect to traditional ones which require a particular service for its validation, improve the quality of the service, and provides new opportunities such as user behavior analysis (I.Gokasar, 2015).

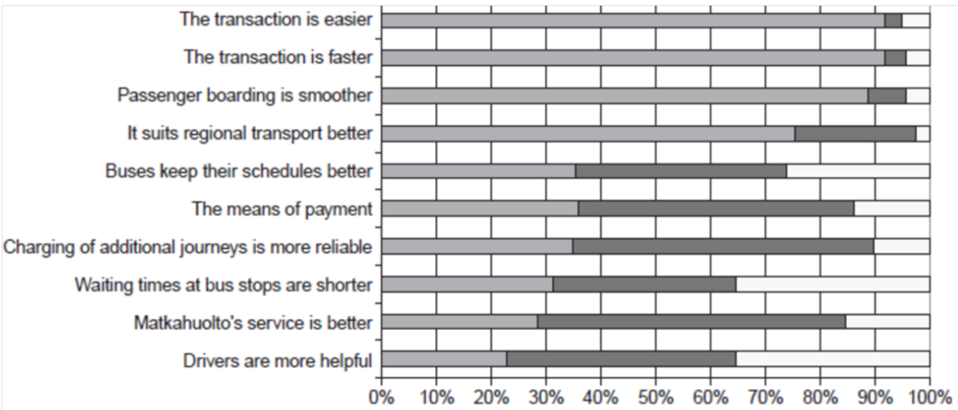


FIGURE 4.3: Advantages and disadvantages of using the smart card (sample of 160 persons, in Tampere, Finland on May 2000) (Blythe, 2004).

In Finland, a survey was made at the beginning of smart card introduction in transit systems treating the usage of smart card used to determine the opinion of the passengers towards this way of accessing the network instead of the traditional ones. Some characteristics of AFC systems are detailed below, based on (Pelletier, 2009), and using the criteria presented by (Vuchic, 2005) for the evaluation of fare collection systems :

- Flexibility and lifetime : Due to its small size which does not disturb the passenger and being lifelong, it appears as an advantageous system compared to other kind of cards, such as magnetic cards (MC) because they are not as durable as the smart cards, and of course to regular tickets. Moreover and in contrary to MC, the smart cards do not require the user to enter the card inside any device (Blythe, 1998).

- Avoid waiting time : The process of identifying the smart card to a machine capable to read the information inside is relatively faster than other traditional methods , thus we can avoid delays before boarding. However the smart card is not the fastest way in the majority of the cases except on the case of identifying the pass (weekly, monthly...) inside the vehicles. This is not usually the single choice, since the readers can also be located at some point in the platforms, instead of being setup inside the vehicles.

- Ease to use : Instead of using magnetic cards to pay the fares, smart card transactions are more appropriate for financial tracking in a specific sector as the transit.

- Costly system : To implement a smart card system in the public transport, a costly equipment are required and should be installed at the vehicles doors or at the stations platforms. Cost is one of the major complications of such system, whereas several other machines can be integrated as banking devices, map displayers and driver roles management.

- Security : smart card is a way to decrease fraudulent transactions since each validation give possibility to travel or to enter a station. Compared to tickets and cash payments, it is a less time-consuming way and requires more human presence.

- Multi payment modes : To proceed to smart card recharge we can do it using several of payment modes , thus the system allow the boarding according to the fares availability. Furthermore , in some systems the fare structures have the possibility to be modified by reprogramming the reading services. Traditionally, one of the hard fare structures is the case of multiple zones where the system requires a validation at both stations of entry and exit (what we call tap-in and tap-out).

Figure 4.4 demonstrates a study done by (Pelletier, 2009) where they group the advantages and limitations for general use of a smart card in the purpose of fares collection as well as the performance analysis.

4.4 Data Acquisition

Concerning the public transport planning purposes, many researches was recently carried out. An overview of some works conducted about this purpose have been grouped here into three categories. The first level is strategic, it concerns the network planning in long-term, behavior analysis of the user, and expectation of the demand. The next level is Tactical, the key issue is time tables adjustments and individual trip patterns. Finally we explore the operational level which relates to supply and demand indicators, as well as to smart card system operations.

4.4.1 Strategic-Level

In this section, long-term planning works are explored (see Figure 4.4). While many studies have argued the possibility of using smart card data collected from transit system for this purpose, smart cards are not the only systems available in transit network and other users , even if sometimes are a few portion from the population, use other ways to collect fares, thus some adjustments are required to build a general view of the network usage. The majority of works conducted have given a possibility of accessing data about user typology and classification while privacy concerns still a priority and access to personal information is not allowed. Even though, some works have used such information.

Author(s)	Data	Analysis/use	Benefits
Agard et al. (2006)	Boarding date, time, and location. Card type.	Define typical user type and measure their trip habits. Analyze use variability according to the day, the week, or season.	Better understand user behavior.
Bagchi & White (2005)	Time, space, and structure. Personal and travel data.	Turnover analysis. Marketing.	Analyze consistency of users travel behavior over time. Produce targeted marketing campaigns to retain users in specific groups.
Blythe (2004)	Route load profiles.	Manage the demand through the network.	Make public transport more attractive.
Chu & Chapleau (2008), Chu et al. (2009)	Boarding date, time, and location. Estimated alighting point. Card type.	Runtime estimation. Itinerary reconstruction. Spatio-temporal portrait on the network. Concept of Driver Assisted Bus Interview (DABI).	Make adjustments to network geometry and schedules. Richer information than that from travel survey. Adapt network to user needs.
Deakin & Kim (2001)	Information from '*Travel Information Services*'.	Provide information in real time on actual and planned conditions of the network.	Allow users to follow an alternative itinerary or to make a well-informed choice before journey.
Utsunomiya et al. (2006)	Personal information. User address, boarding point, frequency of use. Trip information, demand elasticity.	Development of a mailing list for service change announcements. Fare policy analysis. Marketing analysis: identification of market segments with low penetration, conduct targeted surveys. Analysis of demographic profile of riders by route or station.	Allow users to follow an alternative itinerary. Improve user trust in service. Fare adjustment according to user needs. Demand forecasting.
Park & Kim (2008)	Historical data.	Future trend estimation. Creation of a future demand matrix.	Service adjustment (long term). Network extension and adaptation.
Trépanier et al. (2004)	Boarding date, time, and location.	Transportation Object-Oriented Modeling. Planning of the public transport network.	Anticipate network extensions.
Trépanier et al. (2009)	Boarding date, time and location. Estimated alighting point. Card type.	Comparison of smart card data with household survey data (bus use, temporal and spatial distribution of trip).	Improve accuracy of data from both sides. Complete survey with smart card data.

FIGURE 4.4: Example of strategic transit planning works

Today the use of AFC data is being an ultimate choice for any strategic planning over all the data collection methods currently used in the transit sector. The more these data are sufficient the more we are able to understand the mobility behavior since it provides an accurate information about space and time in contrary to other means of data collection (Bagchi, 2005). In each validation of the card, the system saves the time, date and card number, thus can help on statistics at a precise times series studies. Furthermore and since the locations are systematically recorded, the spatial analysis are also possible. In case of ordinary smart cards, the user personal information are missing which means the socio-data, thus these data are used to conduct particular analysis which does not require personal and socio-demographic information. Other data can complement the analysis such as the household survey data (Pelletier, 2009). In chapter 6, we focus on individual itineraries a research previously conducted by (Utsunomiya, 2006) and (Utsunomiya, 2001).

4.4.2 Tactical-Level

Service adjustment is the aim of many researches (see Figure 4.5). Similar time tables are carried out by transport public authorities in the weekends. While, a different tables are carried out for each day of the working days due to the ridership variation. To solve time table issues, many studies were based on smart card data to analyze the network route by route.

Author(s)	Data used	Analysis/use	Benefits
Bagchi & White (2004)	Trip data and personal data.	Reconstruction of user trips, by identifying bus used and transfer point.	Transport offering adjustment. Improved data quality. Statistics available. Schedule adjustment.
Blythe (2004)	Pattern behavior, profile, and preference.	Customer-loyalty scheme.	Better understanding of customer needs.
Bagchi & White (2005)	Origin and destination.	Construct trips made and examine travel pattern.	Service adjustment.
Chapleau & Chu (2007)	Boarding date, time, and location. Estimated alighting point. Card type.	Analyze the boarding passengers variability on a specific route. Detect transfer coincidences. Analyze transfer activities. Identify linked and no-linked itinerary.	Detection of maximum boarding point and return runs. Schedule coordination between bus and metro.
Hoffman et al. (2009)	Magnetic card data on entry points.	Iterative classification algorithm to obtain more information on transfer journeys.	Better view of transfer journeys. Could be applied on smart card data.
Utsunomiya et al. (2006)	History of use.	Determine frequency and consistency in user patterns.	Service adjustment.
Morency et al. (2006, 2007)	Boarding date, time, and location. Estimated alighting point. Card type.	Longitudinal analysis. Spatio-temporal variability. Frequency of use of the bus stops. Temporal variability.	Classification of cards according to boarding patterns. Better knowledge of user behaviors.
Seaborn et al. (2009)	Oyster smart card data.	Method for identifying complete journeys.	Identification of direct links or reroutes to minimize transfers.
Trépanier et al. (2007)	Boarding date, time, and location. Card type.	Algorithm to estimate the most probable alighting point by looking at card journeys and historical data.	Route load profile for each run available.

FIGURE 4.5: Example of tactical transit planning works

Individual trips inference is not a topic that is treated in all the studies, however (Trepanier M., 2007) introduced an algorithm to estimate the most probable routes. First, it can be generated by collecting all the next boarding points of the same day or using historical data of the same card to extract similarities from other taken routes. Thus the itinerary can be estimated due to transaction available data. When the data are missed about the exit, it is theoretically possible to predict the correct trip using the information about the first boarding following it up with time table ,information about the next stops and historical data if necessary, however the literature have not explored this purpose yet. Analyzing transfer journeys as proposed by (Hoffman, 2009) is another research subject, by using the MC platforms which collect identical information as the AFC. A deep understanding of mobility practices by analyzing spacial and temporal data allows a good planning of the network and ultimately meet the travelers' needs. Finally, the data provided by smart cards also allow transit planners to update or add new policies on ridership.

4.4.3 Operational-Level

The performance of the network is the major aim in this level, smart card systems are able to meet a precise indicators of the network performance, like respect of time tables, value of vehicle/kilometers and person/kilometers for each transaction, route, or day (Pelletier, 2009). Time table respect coherency can be estimated by comparing the smart card validation times at a given vehicle along the route with the theorical time table. To do so, the data should be carefully explored since the difference maybe few seconds. Subsequently, smart cards data indirectly refer to data collected in Automated Vehicle Location systems (AVL) (Hickman, 2002); Whereas, AFC can provide the results by fare type.

Author	Data used	Analysis/use	Benefits/Potential
Attoh-Okine & Shen (1995)	Personal data, fares.	Payment.	Reduces perception cost.
Chapleau & Chu (2007)	Boarding date, time, and location.	Detect, quantify, and analyze errors and inconsistencies in transaction data.	Improvements can be proposed to system, and corrected data can be obtained.
Deakin & Kim (2001)	History of use and personal data.	Implementation of fare structure permitting time-of-day pricing.	Better flexibility in fare perception.
Morency et al. (2007), Trépanier et al. (2009)	Boarding date, time, and location. Estimated alighting point. Card type.	Analyze transit user behavior with data mining techniques. Follow buses.	Performance indicators. Schedule adherence.
Park & Kim (2008)	Personal data and bus run data.	Describe user characteristics (transfer point, boarding time, bus run by mode, type of user).	Better understanding of user habits.
Reddy et al. (2009)	Entry-only automated fare collection system data. Magnetic fare cards.	Several operational statistics available at individual level.	Reduces costs previously needed to calculate performance indicators.
Trépanier & Vassivière (2008)	Boarding date, time, and location. Estimated alighting point. Card type.	Several operational statistics available at bus run level. Intranet for dissemination.	Service can be adjusted at micro level. Defective equipment can be detected if data are incorrect or missing. Improves use of embedded system.

FIGURE 4.6: Example of operational transit planning works

Errors recognition can also be detected using the smart card systems itself, this may allow the identification of fraudulent transactions, defective equipment or employee errors. Some errors can mistake the data analysis such as the desynchronization in the machines, changes in initial itinerary (only for the bus), problems in the networks and vehicles (case of unplanned vehicle re-assignement). Such errors can be corrected using methods and techniques which validate the data with a precise error information (Chapleau, 2007).

4.5 Data Mining Tools and Applications

Data mining was definied by (Anand and büchner, 1998) as a tool able to discover implicit , unsolved, previously unexplored and possibly valuable and understandable patterns from big data sets. (Westphal and Blaxton, 1998) categorized data mining tasks as estimation, classification, segmentation, and description:

- Classification is a task that based on the information data extracted from historical data, it assigns labels to a new data.
- Estimation involves completing the missing values.
- Segmentation usually called clustering splits a population into smaller sub-populations based on the behavior similarity and a predefined

metric. It maximizes homogeneity within a group and maximizes heterogeneity between the groups.

- Visualization and description provide a descriptive analysis about the relationships among variables and data in general.

Data mining techniques are dedicated to meet several areas, for example the transit sector has taken its part by (Atizaz Ali, 2016) who published a report of their study about travel behavior analysis using smart card data as well as (FangEmail and Zhan, 2015) who applied data mining to analyze sentiment adding to it the healthcare sector, bio informatics etc.

4.6 AFC and Data Mining

Data mining is a collection of techniques and tools that matches very well the automated fare collection system data due to its large scale amount of data (millions of transactions daily in the case of large cities). Furthermore, based on several dimensions such as date and time, routes choices and fare categories, the travel behavior of passengers can be modeled in different patterns. According to its large scale of data that provides, AFC data cannot be analyzed using traditional statistical methods, especially in case of real-time data collection. (Morency, 2006) used data mining clustering techniques (essentially the K-means algorithm) to identify and analyze groups of passengers according to behavioral patterns of their daily travel. They explored the variability of transit users' patterns in a subsequent work (Morency, 2006). In their analysis, (Ma, 2013b) and (Ma, 2013a) used k-means technique to categorize travel patterns and extract information about trips.

4.7 Applying smart cards and Multi-Agent Technologies in Transit Systems

Today, smart cards are made in response to RFID platforms requirements. This technology is well adopted and is introduced in several areas. In this section, the integration of RFID platforms in transit systems is treated. The key defiance issue comes from the alternatives of incorporating RFID in transport practices, particularly when the researches and studies in this topic are almost rare. Considering, the RFID labels set to individuals (passengers or transit staff) and equipment (transport machines, payment services, vehicles, etc.) the RFID technology allows the authentication, tracking and tracing of vehicles/passengers, security, and other transportation specific issues (Hamilton and Sankaranarayanan, 2013).

In computer science, a software agent or agent is a software component that acts autonomously. It is a program that performs tasks in the manner of an automaton and depending on what its designer defined as a role. In addition, an agent is designed to interact with other agents or human users. As for the multi-agent systems (MAS), it is a series of such agents each in relation with the other.

Agent-Based technology is one of the most advanced and effective techniques for the design of distributed complex systems (D. Guo and Zhang, 2003). In the transit area, multi-agent systems are able to offer alternatives that facilitate the scheduling, decision-making process, presenting more details about a certain situations and decreasing the number of processes carried out by the human operator. Moreover, several multi-agent systems have been introduced to the transportation field of study. These systems address the scheduling issues, real-time information extraction from travel databases, vehicles and staff management and several other facilities.

In addition, the passengers' spatio-temporal density information was made available using several analytical and regression models (L. Sun, 2012), the embarked vehicles (T. Kusakabe, 2011), journeys routes (A. Chakirov, 2011), and public transport modes (C. Morency, 2007). Even though, the information mined using smart cards data has limitations (Pelletier, 2010); for instance, in most of time routing information are missing since the smart card systems only capture information at the entrance and exit from stations or vehicles in case of on board smart card reader machines.

Agent-based models, in opposition to analytical and regression models, aim at modeling each entity agent the most similarly to the real system (Bonabeau, 2002). In particular, the agent acts dependently , produce decisions and communicates with other agents directly. Thus agent-based models allow the monitoring of real-time problems that are resulted from the interactions of the agents with each other.

4.7.1 Urban traffic Monitoring

As mentioned in Figure 4.7 , several multi-agent systems were introduced to facilitate urban traffic control, improve the time precision and prevent unexpected breaks of the systems. MASs provide new possibilities for using a large centralized databases containing travel information to analyze and create solutions (M. V. Belmonte and Fernandez, 2005). Today, in most situations, passenger travel information is recorded in databases of the transport organization authorities, including the personal passenger information (e.g. residence) or the historical travel locations.

Due to the increasing number of vehicles which made the transportation system more complex, many prior studies in the last decade have addressed the issue on how to control and manage the traffic with an effective strategies. For this purpose, a urban traffic control model (Roozemond, 1999), (D. Guo and Zhang, 2003), signalized intersections management (Roozemond, 2001), (M. C. Choy and Cheu, 2003), bus journeys management (M. V. Belmonte and Fernandez, 2005), bus-holding control (J. M. Zhao and Dessouky, 2003), the dynamic travel pattern-directing system (X. Shi and Song, 2005), introduction of urban traffic control and travel pattern-directing system (Li and Shi, 2003), and distributed traffic data process and management (H. S. Zhang and Hu, 2004), (Chen and Bell, 2002).

The first urban traffic control (UTC) model were designed with a centralization concept, thus many difficulties faced the implementation and time adaptation especially when the traffic grow (Bazzan, 2005). In (Bazzan, 2005) and (S. Ossowski and Garcia-Serrano, 1998), the authors have introduced a non centralized traffic control and a distributed decision support system (S. Ossowski and Garcia-Serrano, 1998) to solve problems of the traditional UTC.

Project name	Research group	Application domain	Key Features
CARTESIUS	Logi and Ritchie [23], Technische Universität München, Germany	Traffic congestion management	Two interacting knowledge-based agents for inter-jurisdictional traffic congestion management
IDTMIS/UTC	Roozemond [27], [29], Civil Engineering and Geoscience, Delft University of Technology, Netherlands	Urban Intersection Control	A theoretical UTC model based on several Intelligent intersection Traffic Signaling Agents (ITSA), Road Segment Agents (RSA), and some authority agents
	Guo et al. [28], Tsinghua University, China	Urban traffic control	A conceptual three-layer urban traffic control model
	Choy et al. [30], National University of Singapore, Singapore	Traffic signal control	Hierarchical multi-agent architecture, embedded fuzzy-neural decision-making modules, on-line learning capability
	Zhao et al. [33], University of Southern California, U.S.A.	Bus-holding control	Achieve dynamic coordination of bus dispatching through the negotiation between a bus agent and a stop agent
CORBA-A	Zhang et al. [36], Tsinghua University, China	Traffic data management	Distributed MAS architecture, the software implementation is based on the combination of CORBA and agents
TRYSA2	Ossowski et al. [39], Rey Juan Carlos University of Madrid, Spain	Decision support for urban traffic control	An application of TRYSA2 system for distributed decision support of urban traffic control
	Srinivasan and Choy [45], National University of Singapore, Singapore	Traffic signal control	A single layer cooperative multi-agent architecture, the cooperative zones of individual agents can be updated and allocated dynamically
aDAPTS	Wang [4], [5], Chinese Academy of Sciences, Beijing, China	Urban traffic control and management	A hierarchical three-level agent platform for urban traffic control. The platform integrates mobile agent concept with multi-agent systems
ACTAM	Chen et al. [46], National Chiao-Tung University, Taiwan	Traffic signal control	The modular design of an adaptive and cooperative traffic signal agent
HUTSIG	Kosonen [48], Helsinki University of Technology, Finland	Traffic signal control	Allow on-line simulation that takes real-time detector data to model traffic situation
	Moreno et al. [53], DEIM, URV	Taxi dispatch	Customers can request a taxi from his/her PDA through a personal agent running in a JADE-LEAP agent system
NTuCab	Seow et al. [54] – [56], National University of Singapore	Taxi dispatch	Propose a taxi dispatch system that attempts to minimize the total waiting time of a group of customers
	Ossowski et al [59], [60]. University of Rey Juan Carlos, Spain	Decision support systems	Define an abstract architecture for multi-agent decision support system. Instantiate each type of agents in the abstract architecture for two case studies in the domain of traffic management
SATIR	Balbo and Pinson [61], Paris Dauphine University, France	Decision support systems	Propose an interaction model called ESAC (Environment as Active Communication Support) to allow agents to interact directly, sending and receiving messages through logical filters of emission, reception, and interception
SATIR	Balbo and Pinson [62], Paris Dauphine University, France	Disturbance modeling and decision support systems	Knowledge related to the network dynamics that is stored in vehicle sensors is incorporated into disturbance modeling and decision support system

FIGURE 4.7: MAS applications in Transportation.

4.7.2 Multi-agent Traffic Modeling And Simulation

Today, the transit systems contains several smart component such as signalized intersections and human drivers etc. that belong to one complex network and communicate between each other to realize a particular function. The multi-agent systems allow modeling and simulating this network based on its advantages in defining each component role separately i.e. each component role is considered as an agent. Since a cooperation exists between the agents, multi-agent technology has recently found its place on the routes modeling, railway transportation systems (Bel and Stokkum, 1999), and optimization and congestion (Nasri Bin Othman, 2015) etc. Paul et al. in (Paul Hamilton, 2013), have proposed a RFID-based transit system and Multi-Agent technology in bus trips scheduling and an enhanced smart card payment system was also implemented. This implementation introduced a full readjustment of the task in which transit system are being examined, to consequently provide an ease of routes tracing and time management of trips based on the monitoring of vehicle/passenger interaction supply and demand. The possibility of getting the vehicles' location in a specific time as well as the passengers flow at stops by the means of RFID, highly reinforce the scheduling capabilities. The aforementioned system in (Paul Hamilton, 2013) was based on Google Map to determine the geolocation coordinates of bus stations/stops, stations indicators and populous places. Their application was combined with an Android package based on JADE-Leap agent development kit, which allowed the user to consult a map providing information about the his/her actual location as well of a desired bus including the bus schedules. Thus, the system would also be able to inform the scheduling updates, fares tables and suggest other bus stops with an estimated time of arrival.

Chapter 5

Behavioural Data Collection Using Smartphones

5.1 Summary

Research has been increasingly in need of real-time or close to real time monitoring tools to understand human behavior. The mobile phone has become one of such tools widely used in such studies. The reason for its popularity is easy to explain. Exclusively, at least three good reasons explain the fact that mobile devices are irreplaceable for this type of study.

(a) First, mobile phones are essential companions to all of us ; always in hand. This gives a unique opportunity to monitor human almost continuously. The ability to monitor people non-intrusively in real time without requiring them to carry additional hardware is fundamental.

(b) Second, modern mobile phones are equipped with several types of sensors. These sensors can provide information about the location, acceleration, activities, sound, and with high importance, social interactions. Continuous information provided by these sensors, and the correlation of these data with other phone data is an asset for social monitoring. Thus an intensive information about daily activities of people is recorded.

(c) Third, mobile devices are now the link between us and our social networks. They can be used to make calls or send text messages, as well as to remain connected to our virtual networks, organize our daily lives and search, create, consume information. This huge amount of social information we carry in our mobile phones is a gold mine for researchers to better understand the interactions, context and social behavior of people.

The purpose of this chapter is to illustrate the counting methods based on mobile phones. It presents the prospects for the foreseen counting methods, whether it is emerging Smartphone-based technologies which are considered to be a new potential means of measuring travel behavior. Either they convert pre-existing devices, or they rely on other sources of data through all the personal data they can recover. Big Data, geolocation (voluntary or not), traces etc. are all potential ways to measure the mobility and its evolution. So, what do these new tools provide to measure mobility? What are their specific advantages and limitations? In what these digital methods (Rogers, 2013) can be used or are still under design to identify practices of mobility? How can other methodological devices (out from maintenance or

major travel surveys) be mobilised? We think as well to the networks and social applications (Facebook, Twitter, Instagram, Tindr, Grindr, etc.) which give other tools or means to think, visualize the mobility, the ways in which it is set (or not) on stage or in which individuals represent it.

5.2 Data Collection

The data collection process consists in collecting the information saved systematically by the Smartphone system. Since the market- availability Smartphone contain a database system: where user's logs are recorded, the exploitation of these data maybe significant to study the user behavior. Moreover, to make it easily to understand, we can classify the data that can be collected into three categories Table 5.1,5.2,5.3 , where:

1. Continuous data (Sensors data).
2. Historical data (Saved in the system).
3. Request-based data (Current Data).

Request-based data stands for the data that we need to ask for to get the present values/information from the smart-phone system such as the current position, battery autonomy, nearby Bluetooth etc. Historical data refers to information that are previously saved in the system.For instance, the historical data may be the contact list, call log, SMS log, etc. Continuous data are all the data that we can extract continuously such as sensor data (accelerometer, gyroscope, magnetic field, and etc.).

TABLE 5.1: Sensors data

Sensors data	Description
Light	Ambient light level
Proximity	Proximity of an object in cm relative to the view screen of a device
Temperature	Temperature of the device in degrees Celsius (C)
Magnetic field	Ambient geomagnetic field (x, y, z) in uT
Pressure	Ambient air pressure in hPa or mbar.
Screen display	Screen of the phone (on and off)
Softwares	List of running applications
Activities	User activity log based on motion sensors

TABLE 5.2: Historical data

Historical data	Description
Call	User call log
SMS	User SMS log
Applications	List of installed application
Hardware	User's Smartphone hardware info
Browser Bookmarks	User Bookmarks
Browser Searches	User Browser log
Contact	User contact (phonebook)

TABLE 5.3: Request-based data

Request-based data	Description
Position	GPS data (user location)
Wifi	Nearby Wi-Fi signals
Bluetooth	Nearby Bluetooth signals
Battery autonomy	Battery status
IP adress	Phone's internet IP address
SSID	Service Set Identifier (phone network provider)
Operating System	Windows, Android, iOS, Blackberry
3G/4G	Type of connection through the mobile network

5.3 Smartphone sensors

Smartphones contain different types of sensors capable of providing information about the mobility. These sensors can be categorized into three groups according to their application to travel data collection:

- Motion sensors:
  - Accelerometer, which measures the device linear acceleration;
  - Gyroscope, which measures the angular rate of change (i.e., rotation velocity); and
  - Magnetometer (i.e., compass), which measures the magnetic field strength;
- Location sensors:
  - GPS, which is commonly used in outdoor settings, and
  - Network-based location services, which use cellular network and wireless fidelity (Wi-Fi) to determine the location (i.e., via triangulation);
- Ambient sensors:
  - Light sensor,

- Microphone, and
- Proximity sensor, which detects nearby objects and can indicate when the phone is near the user's ear (e.g., during a call).

Technically, two types of sensors exist. In one hand, hardware components, which means the presence of a physical component on the device. This type of sensors provide data of an exercised physical movement. In the other hand, we have sensors that are a developed software. They are based on physical sensors output data to calculate new data.

Generally, all the devices have at least one sensor. For example, the majority of devices have an accelerometer or a magnetometer, but few of them have a thermometer. Moreover, it is possible that a device has several copies of a sensor, but calibrated differently so as to have different results.

5.3.1 Motion Sensors

Today, mentions sensors are embedded in the smartphones Figure 5.1, such as the accelerometer, gyroscope, and magnetometer sensors (Rachuri, 2012), these tools offer to us a side channel into a user's mobile device, from which we can deduct several types of information. Specifically, we point out to the onboard accelerometer and magnetometer sensors, which are two of the most commonly supplied motion sensors on smartphones. Up to the recent version of major platforms, e.g., iOS , Android , and Blackberry, no security permissions are required for accessing these two sensors , and enables considerable accesses by a third party application to the underlying resources in the mobile device due to the assumption that data collected by motion sensors is not sensitive.

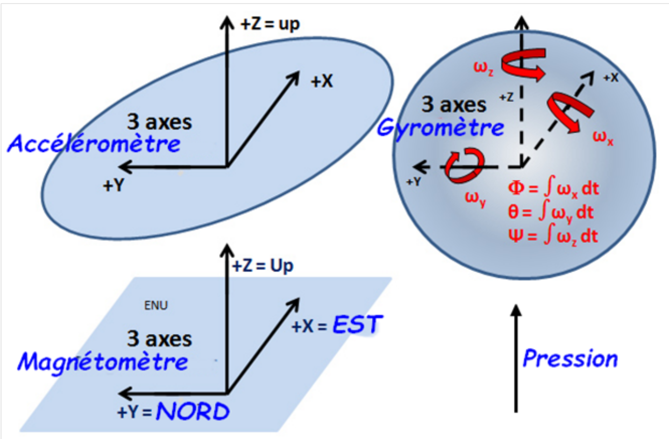
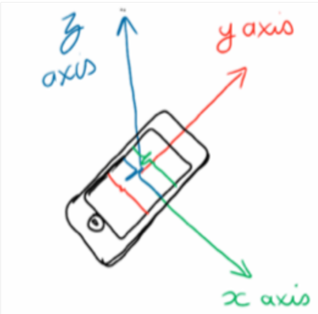


FIGURE 5.1: Motion sensors https://sites.google.com/site/wikismartphone/techno/

The Accelerometer

The accelerometer measures the non-gravitational acceleration based in three axes Figure 5.1 It is an electronic component of a few millimeters implemented inside the device. It measures all the accelerations experienced by the Smartphone, including in particular the Earth's gravity (Akram Bayat, 2014). The accelerometer is rarely set in the center of Smartphone. Acceleration is the rate at which the user motion is changing — if the user is changing direction or speeding up or slowing down, he /she is accelerating. Our perception of motion naturally rely on three possible orientations : left/right, up/down, and forward/back, and the phone has 3 sensors ("accelerometers"), one for each of these directions. If the phone is in the user's pocket, and the user start walking, often the user's phone is moving down, but then changes direction and is moving up. The gait of the walk means the body turns, so sometimes the phone is moving left and sometimes right. The lines depicted in Figure 5.2-Left represent the three changing motions of the phone.



deg	X	Y	Z
0/360°	0.5g	0.5g	1.0g
90°	0.5g	1.0g	0.5g
180°	0.5g	0.5g	0.0g
270°	0.5g	0.0g	0.5g

FIGURE 5.2: The axes of the smartphone sensors. Source: http://bildr.org/2011/04/sensing-orientation-with-the-adxl335-arduino/

The raw data of the accelerometer are represented by the values of the standard acceleration due to gravity as depicted in Figure 5.2-Right

The Gyroscope

The 3-axis gyro present in our smartphones and our tablets allows to our smartphones to determine their orientation. It measures three degrees of freedom. While the accelerometers measure the translational movements, gyroscopes measure the rotational movements (Zhang, 2012). That is to say, the gyroscope does not detect linear movement along an axis, but a rotational acceleration around an axis. The rotation around the X, Y, Z axes may respectively be refereed to as roll, pitch and yaw as in boats or planes or in terms of degrees of beta, gamma and alpha rotation. (Figure 5.3)

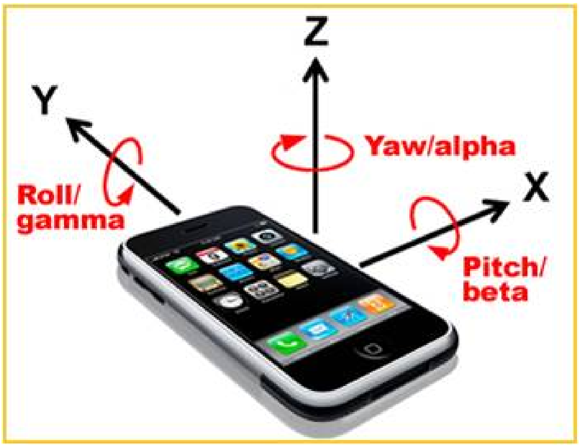


FIGURE 5.3: Rotation around the X, Y, Z, axes of a mobile phone. Source: http://hillcrestlabs.com

The gyroscope gives only relative data (e.g. Rotation of 30 degree clockwise) (Zhang, 2012). The smartphone can determine the orientation of the phone according to the vertical using the accelerometer and gyroscope to get the absolute value of pitch and roll. As for Yaw value, we can use the compass and gyroscope. But most of the games use only the angles' variations. E.g. Car racing games in which the users use the Smartphone as a steering wheel, or shaken the device to turn off the music.

Important: There are several technologies of gyroscope (the vibrating beam, the fiber optic gyroscope, etc.), however, the most common is called vibrating structure gyroscope. These gyroscopes transmit data 200 times per second, a frequency of 200 Hz. The vibrations of our voices are within a range of 80 to 250 Hz (Zhang, 2012). Consequently, the gyroscope can also be used to "listen", namely, 'gyrophone'.

The Magnetometer

The magnetometer sensor based on the modern solid state technology produce a miniature Hall-effect sensor that senses the Earth's magnetic field along 3 perpendicular axes X, Y and Z (Chiang, 2013). a Voltage is produced by the Hall-effect, corresponding to the force and polarity of the magnetic field along the axis each sensor is directed. A digital signal is produced from the detected voltage which represents the magnetic field intensity. Magneto-resistive devices, one technology from others that may be used for magnetometer (Chiang, 2013). It modifies the measured resistance based on modifications in the magnetic field.

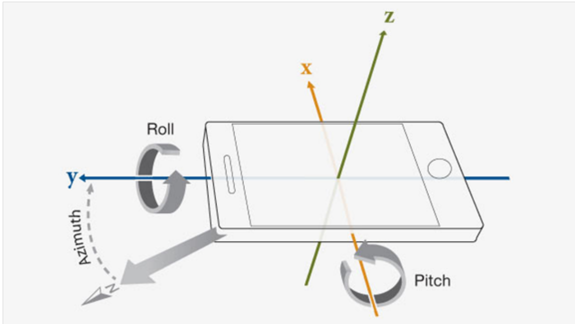


FIGURE 5.4: Smartphone position related to the North Source: http://fr.mathworks.com/hardware-support/android-sensor.html

Concerning the magnetometer design, it is often integrated with another sensor (e.g. accelerometer) both installed in a small electronic chip. Typically, it incorporates with another built in sensor to correct the raw magnetic measurements using the data from this sensor. Additionally to the rotational information, the magnetometer allows the detection of orientation relativity of the device with respect to the earth's magnetic north.

Important: Apps made to detect metal and to rotate maps/interfaces/graphics use this sensor as well.

5.3.2 Location sensors

Based in positioning technologies the smartphone location can be deducted. Such technologies A-GPS, GPS and Wi-Fi are usually embedded in all the smartphones and are mainly used for geolocation task. However, these technologies are different in term of the usage and methodology in travel data collection. Figure 5.5 briefly compare them.

	GPS	Wifi
Accuracy	👍	👍
Energy consumption	👎	👍
Start-up	👎	👍
Indoor functioning	👎	👍
Urban area	👍	👍
Rural area	👍	👎

FIGURE 5.5: Comparison between GPS and Wi-Fi

GPS and A-GPS

A-GPS and GPS are two different navigational systems which use satellites' data to determine their exact location on Earth.

GPS stands for Global Positioning System. Being available anywhere on Earth, this successful communication tool communicates with 4 or more satellites to find its exact location usually represented by two values : the latitude and longitude. The communications can be held at any weather situation, if a signal exists (see: www.diffen.com).

A-GPS stands for Assisted Global Positioning System. The communication basic is equal to GPS, whereas the data reception becomes quicker since the communication with the satellites is done by the mean of network resources e.g. mobile network, also called assistant servers.

Traditionally, using GPS data we are able to apply Transport mode inference and path detection algorithms (a.k.a map-matching, MM), based in two steps (Schuessler, 2009): first, divide the data into several unimodal segments, than deduce the transport mode of each segment; Second, perform a map-matching for each segment independently. **Important:** A-GPS is rapid in determining the location however GPS determine a more precise location information.

Transport mode inference

GPS devices provide high accuracy and density of location data. Based on the coordinates communicated, acceleration and deceleration as well as the speed can be computed. The modes are distinguished by a deterministic rule-based methods conditioned by a predefined criteria. Considering the average speed, (Bohte, 2009) could determine the transport mode of a particular journey. Other criterias may be used, such as 85th percentile of speed , acceleration and deceleration , in addition to possibility of using rail and ferry networks to recognize the proper modes (Stopher, 2008). Moreover, (Chung, 2005) in their side , they have used a method based on a set of predefined mode change chains to determine new sequences of modes. The research has demonstrated that the deterministic rule based methods are sensible to data outliers, thus different approaches are proposed: possibilistic and probabilistic approaches. Using the speed and acceleration data, (Schuessler, 2009) introduced the usage of fuzzy logic in the modes identification. In addition and since the transport mode inference is a sort of standard classification problem, machine learning methods may fit the research needs. Consequently and according to (Zheng, 2008), the decision trees were the best and the most convenient model in performing the classification to speed data with respect to Baysian networks , conditional random fields and support vector machine.

Smartphones includes a built in GPS, however the data are less accurate compared to GPS devices' data. Despite of this, the smartphone records additional sensors data such as accelerometer data ACCEL (Akram Bayat, 2014). Traditionally, multiplicity in data create a good conditions for machine learning methods. This explains their large usage in this research area. Many features of the brute and raw ACCEL data can be explored: mean, variance, time between peaks ,etc. (Ravi, 2005), (Nham, 2008), (Kwapisz, 2010), (Ding, 2010). Besides, several other types of data can infer the phone holder status. For instance, (Reddy, 2009) conducted a classification of transport modes using an application of naive Bayes, decision trees, K-nearest neighbor, support vector machine and hidden markov model. The naive Bayes, decision trees, Bayesian network and random forest were also tested by (Stenneth, 2011) using the proximity data to bus stops and train stations. The problem of data outliers create some complication while using rule based methods. The machine learning methods aim at rapid inference of modes. However, standardization of the problems is required and generally are simplified to fits the usual frameworks.

Unimodal map-matching of a segment using GPS

Generally , the raw GPS data are complemented by the transportation network data in order to provide significant results. The map matching techniques are efficient using data from navigation systems. A good literature review is made by (Quddus, 2007) which involve 35 Map-matching algorithms. This latter can be validated using validation strategy, according to (Quddus, 2005). Thanks to its navigation-based design , the actual map-matching algorithms produce an online deterministic identification of the real route from a single GPS point. Even if some of it assume arcs connectivity and contiguity, they still do not validate the roads detected and path final combination (e.g., (Greenfeld, 2002), (Ochieng, 2003)). The historical data , also call experienced data, can be used to extract whole trip path. For instance, (e.g., (Yuan, 2010)) used these data type to create new navigation techniques dedicated to routing strategies which uses the experienced road network users data (e.g. data recorded from floating cars in the transportation network) to estimate travel time. (e.g., (Ebendt, 2010)). According to (Bierlaire, 2008), the input for route choice models are the path observations.

Network-based location services (Wi-Fi) in context of mobility

Wi-Fi geolocation operates the known position of close by wireless networks to determine the location of a device. A smartphone equipped with Wi-Fi technology can estimate the location based on the nearby Wi-Fi access points. The accuracy depends on the power of Wi-Fi access points, that is to say, a few dozen meters (network, 2012). Wi-Fi technology is able to provide real-time positioning data, even indoors where GPS cannot be used. It does not explicitly ask users for their permission, as it does not require any user intervention. However, the only way out for user, is to completely disable Wi-Fi on the mobile device (Antonio J. Ruiz-Ruiz and Teruel, 2012).

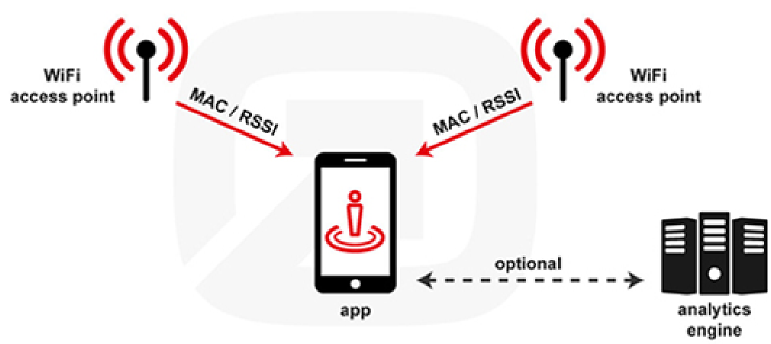


FIGURE 5.6: Wi-Fi positioning system scheme Source: https://www.infsoft.com/blog-en/articleid/40/indoor-navigation-using-wifi-as-a-positioning-technology

Important: Google uses cars that roam the cities to take pictures to make it available in Google Street View and to record all detectable wireless connections. This data is combined with geolocation of the car and serve to reinforce the power of Wi-Fi geolocation.

5.3.3 Ambient sensors and Indoor localization

The location characteristics such as light, proximity and noise etc. can be detected using ambient sensors. This latter are interesting due to their low consumption of energy compared to location sensors surch as GPS (Carroll and Heiser, 2010). It was confirmed based on several experiments that collecting data from all four ambient sensors from Samsung Galaxy S4 showed a 180 times less power than location sensors (S. Mazilu and Troster, 2013). Thanks to its low-consumption and without any other extra-infrastructure required at home, these sensors embedded in our personal phones allow a continuous inferring of home's ambient properties. Moreover, the ambient sensors can be adjusted with respect to energy-accuracy variation. For an accurate location data, we can use power intensive sensors. Thus, the battery gets empty faster. Or in case of long-term positioning, we use low-consumption ambient sensors. However, the accuracy is low (S. Mazilu and Troster, 2013). Determining the indoor location inside the residence is a good example to infer daily-life indoor and users' activities.

Temperature, humidity, pressure and light sensor

Temperature, humidity and pressure sensors have been recently joined to the existent light sensor in the Smartphone. These sensors are not well exploited by the recent researches except to detect phone's environment (Barometer-based weather information). The question is now linked with the future of these sensors and the phone producers are hesitating if they should keep/drop these sensors from the next devices due to its rare application. Thus, several studies are now challenging the issue and explore a new useful techniques of using such ambient sensors. Many researches argued that the data provided by ambient sensors are able to sense phone's context and localization: In one hand, a study explored a technique which uses acceleration and sound data (power-hungry sensors modalities) with other ambient data such as light , room color and fingerprints to detect the phones' context SurroundSense (M. Azizyan and Choudhury, 2009). In the other hand, a research applied light sensor to determine the light properties of the locations (e.g. office building) to detect the indoor location (N.Ravi and Iftode, 2007). In addition, the pressure sensor also took its part from the research and (T. Watanabe, 2012) have explored the use of these sensors in a case study of subway stations' identification. The localization context is not well determined by pressure sensors, however they are very efficient in floor changes and ascension modes (elevator, stairs etc) (K. Muralidharan and Agarwal, 2014). Moreover, the indoor localization is quite impossible to be determined using GPS due to the lack of signal, thus the ambient sensors were tested by (S. Mazilu and Troster, 2013) in terms of indoor localization. The goal of their research was to sense a particular user's phone assuming that the rooms inside a house each have different ambient properties (orientation, light , temperature and humidity). The position (geographical orientation) of a room as well as the heating devices available inside the rooms to determine its temperature. For instance, the rooms containing heaters are warmer than the others and the rooms fronting the sun all the day too. The rooms are also influenced by the user activities inside it, for example taking a shower increases the humidity inside the bathroom and cooking increases the temperature inside the kitchen. Using the barometer, the floors take different pressure levels Figure 5.7. Finally, the light inside a house is different from room other and the equipment are unlikely similar.

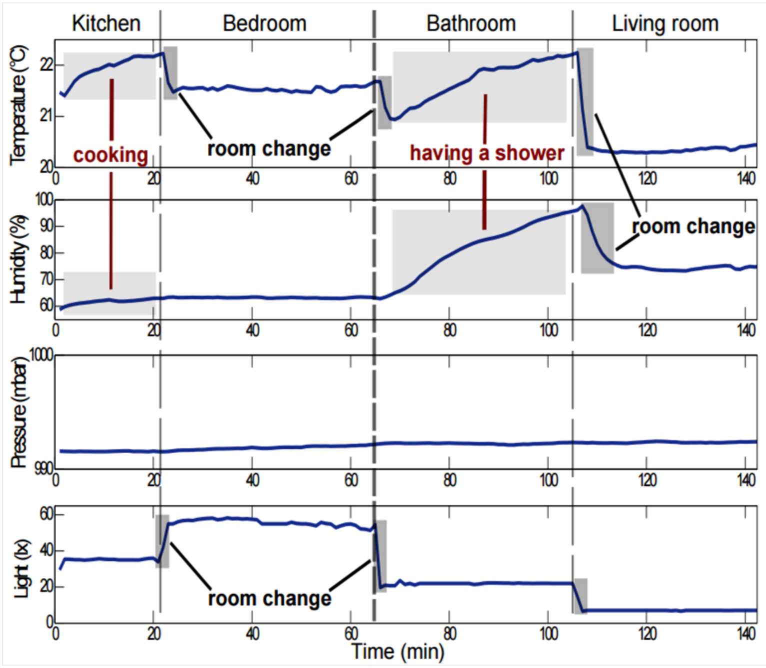


FIGURE 5.7: Temperature, humidity, pressure and light collected from the internal phone's sensors

During approximate 2 hours at night. A user was asked to hold the phone in his/her proximity during the movements inside the house and to perform the usual daily-routine activities.

Noise Sensing Microphone

Sound sensors made Individual's context in daily-life possible to infer. It provides a rich information data, thus the context may be recognized because each activity has its characteristics (speaking, eating, bathing , walking etc.). The sound changes according to the location type, for example restaurants provides different sound patterns than streets. Determining real-time sound information is possible using mobile applications. For instance, It is usually interesting to change the phone mode during a meeting, or give information about the user location. The researches argued that sound recognition is a quite good way to infer user activities (J. Chen and Shue, 2005), social events (M. Rossi and Troster, 2012) or environments (A. J. Eronen and Huopaniemi, 2006), (A. Mesaros and Virtanen, 2010). Using sound sensors has many advantages compared to others sensors usage.

They are cheap and available in all the Smartphones as well as in the ordinary phones, noting that they work efficiently even hidden (T. Franke and Bannach, 2009). Today, the majority of the researches applied the sound-based inference data in computer simulations and has been analyzed in an offline mode. However, due to the advances on Smartphone technologies and the high computational power combined with internet connectivity , the real-time analysis and applications made possible. M. Rossi et al. in [20] explored the these real-time applications implementation to confirm the system design parameters, while other researches focuses on the classification purpose of sound data. This sensor modality have shown a real efficiency in recognizing the locations context , in (J. Chen and Shue, 2005) for bathroom, (M. Stager and Troster, 2004) for workshop, (E. Miluzzo and Campbell, 2008) for kitchen , (A. J. Eronen and Huopaniemi, 2006) for public places, and (A. Mesaros and Virtanen, 2010) to evaluate an classification system for a set of classes around 60. Whereas, the real-time issue was not applied to the aforementioned researches. Concerning the real-time applications of sound classification, just few studies have been conducted. In (Mirco Rossi, 2013) system called AmbientSense was proposed by M. Rossi et al. useful to infer the sound properties (class) in real-time. This system was developed for smartphones use, it is available on Android or work by the mean of additional server to determine a captured sound classification. For 23 ambient sound classes Table 5.4, the recognition accuracy of the system was 58.45%, which meets result of previous efforts in auditory scene analysis that considered a similar number of ambient sound classes, e.g. (A. J. Eronen and Huopaniemi, 2006).

TABLE 5.4: The 23 ambient sound classes used to test AmbientSense system.

Beach	Crowd football	Shaver
Bird	Dishwasher	Sink
Brushing teeth	Dog	Speech
Bus	Forest	Street
Car	Phone ring	Toilet flush
Chair	Railway station	Vacuum cleaner
Coffee machine	Raining	Washing machine
Computer keyboard	Restaurant	

Proximity sensor

Today, the user ear proximity to the phone device can be detected using IR-based proximity sensor. The main aim of this sensors is to economize the power consumption each time the phone is close to human ear by turning off the LCD and deactivate the touch screens to avoid the cheeks contact. Similar to other technologies, these sensors have advantages and limitations. Concerning the latter, it includes the high consumption of power and it is costly. Other proximity sensing techniques such as Capacitive Proximity Sensing are in view of development and they can replace the IR sensor technology, thus the cost will decrease.

Proximity Sensing-Cased Applications

Several features was made possible using Capacitive proximity sensing. This technique was recently implemented in smartphone devices and brought several advantages with respect to traditional IR-based sensing. Among the advantages we note, cost reduction , integrated functionality and the most important one is power consumption (Noopuran, 2014). These features includes:

(a) **Face detection:** This technology is not a real new technology, it existed for over a decade. This technique aim at preventing the inadvertent touches from users when the device is close to user's ear. Thus, it economize the energy by shutting off the screen. Since the mobile phones use a battery so this technique is really useful (Noopuran, 2014).

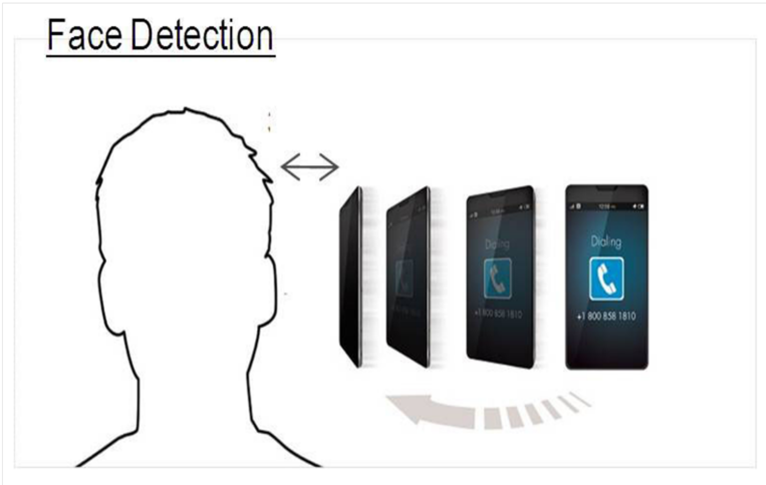


FIGURE 5.8: Face detection Source: www.wirelessdesignmag.com

(b) **SAR (Specific Absorption Rate) regulation:** This technology usually exists in tablets, however in some cases the mobiles phones integrate it as well. SAR stands for Specific Absorption Rate, which is the rate that relates to the energy absorbed by human body at the time of radio frequency / electromagnetic field exposition. The device producers are required by the Federal Communications Commission FCC to implement a system that reduce the absorption of radio frequency energy by reducing the transmission power of the device when this latter is close to human body (Noopuran, 2014).

FIGURE 5.9: SAR using capacitive sensing Source: www.wirelessdesignmag.com

(c) **Wake-on approach feature:** Another proximity feature, also linked to battery energy preservation , it shifts the screen from on to off depending on the user availability and usage. If the user is absent, the screen is in standby state and wake-up in the opposite case (user's hand approaches the touchscreen). Thus, the wake-up time is reduced and the energy is optimized (Noopuran, 2014).

FIGURE 5.10: Wake-on approach feature using capacitive sensing

5.4 Smartphone Log Data

Mobility measurement requires a real-time rental observations. In order to provide useful information for travel behavior study, the mobility logs must be recovered from the raw Smartphone data. Smartphone log gathers several type of data : contacts, call and SMS logs, sensors, GPS, Wi-Fi, Bluetooth, etc. Depending on the information we need, we can choose the appropriate data provider. For example in (Min JK, 2013), researchers tried to study the constancy , temporal tendency using contact list. The relational life type (e.g family, work) reached 90% of accuracy and it was carried out using machine learning techniques in addition to their own methods. In (Min JK, 2011), researchers proposed an artificial assistant namely SmartPhoneBook that propose a phone numbers that the user possibly would like to call in a specific situation. To achieve such objectives, they use social relations among the contacts based on the contact patterns that are constructed by the mean of the user emotional behavior which is deducted based on the mobile log. Bayesian networks were used to avoid uncertainties. Besides and out from using the contacts information, (Singh VK, 2013) smartphone log was used to know the business relationship. They intended to predict the spending behavior of clients and finally classify the clients as loyal customers or overspend. In (Meng L, 2014), researchers have explored location features to learn about positioning, proximity and profiling user to understand user behavior. Their results approved (i) relations (SMS contacts and Facebook) in proximity highly influence the traffic consumption,(ii) applications choice depends on personality, and (iii) depending on the location, its usage differs. Another location-based study was discussed in (Montoliu and Gatica-Perez, 2010) where the authors used location information obtained from different sensors such as GPS,Wi-Fi, and motion sensors) to develop a new framework that proposes in each user location a list of interesting places to visit.

5.5 Massive collection of data

The majority of travelers are holding their smartphones during their mobility, this may be the source of a massive collection of data. Smartphones also produce raw information other than GPS, and we mention: Bluetooth, Wi-Fi, photos etc.

Among the information needed during an investigation with traditional tools, we have flows computing and behavior measurement.

5.5.1 Flows computing (using geolocation services)

Computing of flows in an urban space, the rate of entering/leaving and level of attendance and the predicting development. Localization is essential tool in the management of flows: indeed, the measurement of people mobility has no importance if it isn't time-stamped and geocoded. The location also finds applications in users mapping for the location of travelers or the analysis of behavior.

If the location of stationary sources (e.g. counting stations, fixed sensors) is determined in advance and therefore requires no dynamic localization, that is not so with mobile devices. Regardless of the mobile platform (Android, iOS...), the localization APIs are based on the data from the GPS, Wi-Fi and the cellular network. Two modes are possible, the ongoing subscription to periodic updates (at a given frequency) position when a movement is detected (see Figure 5.11), or one-time recovery of the last known position (more economical for the battery). However, localization is not available in areas that have no network coverage and accuracy can vary from 2m to 200m depending on coverage (e.g. no GPS inside buildings).

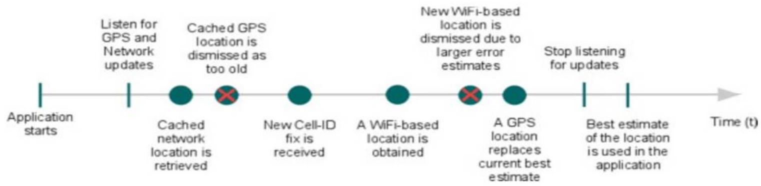


FIGURE 5.11: Google Maps API: window position over time estimate Source: https://developer.Android.com/guide/topics/location/strategies.html

5.5.2 Behavior measurement and Panel Methodology

The behavior can also be measured. For example, in a situation of earthquake, a study was described by (Lesas, 2015). The smartphone has been used as a sensor for the understanding of the behavior of the people. The results of his experiment confirm the relevance of the implementation of mobiquitous technologies applied to seismology.

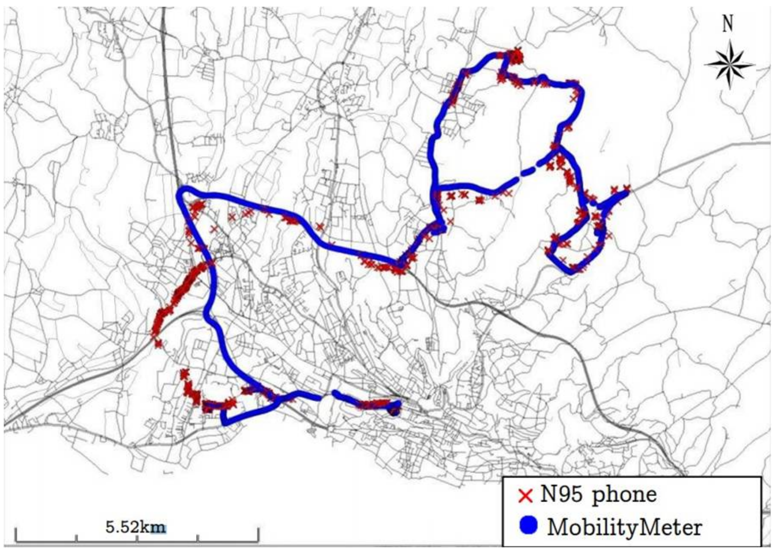


FIGURE 5.12: GPS traces from N95 and a GPS device (Source: (Chen, 2013))

Accelerometers Sensors embedded in the smartphone in its normal use when the user is moving: prototyping (on Android and Windows Phone) has demonstrated that even if there is a possibility to identify a particular user activity Figure 5.12 (i.e. change of direction, walking, running, steady state, shaking of the mobile) by comparing the signals to a pre-established repository, it is however difficult to differentiate vibratory movements caused by a 'normal' user of a movement of seismic origin activity , and this despite the implementation of a protocol for automatic corroboration of signals with devices located in the vicinity. An alternative (and we have prototyped) involves prompting the user which then triggers an alert, but our tests have shown that in underground or the crowd movements, the system fires unintentionally and therefore no interest.

Like other information that can be retrieved automatically via technologies integrated on smartphone, there are applications that enable questioning the user. According to several case studies, as in social networks, to use Facebook user must make an inscription which presents itself as a form requesting personal information. In a same way, mobile applications do. Users in accordance with the trust and importance who wins to the tool, they can share some information such as their identity, photos, name, address etc.

5.6 Data processing

GPS records are constant and ought to be segmented to movement and travel periods (i.e. trip) for travel analysis. The technique proposed by (Schuessler, 2009) is utilized for such purpose. This strategy does not generally create accurate results. It is frequently needed to manually mention the starting and the end of the trips, both concerning the space (i.e. starting point and destination) and time (i.e. begin and end time). In this way, additional data sources for trip and OD inference are currently under research. It is well known that Wi-Fi data can be used to indicate location and infer activities taking place at specific locations. Hence, they can be complementary to GPS data to learn mobility. A part from the conducted studies focuses on the inference of location (home or work as well as additional places of interest) of the users; it identify potential origins and destinations for each user; Trip time detection. Many algorithms are currently under investigation for this purpose. Clustering of Wi-Fi Records for identifying Meaningful Locations: The aim is to detect the most frequently visited places of each user and try to assign a meaning to each one of them (semantic places). Clustering of Wi-Fi access points (APs) for each user can be performed in order to identify the most frequently viewed APs, and analysis of time patterns of visiting the clusters assists in assigning a meaning to them. At the moment, the algorithm is used to infer home and work locations. A density-based spatial clustering of application with noise (DBSCAN) is tested to group APs (Montoliu and Gatica-Perez, 2010). It takes the most visited clusters and use time information to infer home and work. Assumptions with respect to the time patterns include: 1) during the night the user is supposed to be at home, not at work, 2) the user is not at work during the weekend, and 3) during the day the user can be either at home or at work. Trip Time Detection Using Wi-Fi Data: The clustering from the previous step can be used for trip detection in the following manner. The trip detection is based on change of cluster. As soon as the user remains in the same cluster for a pre-defined threshold of time, no trip is identified. If the person leaves the cluster and a new WLAN-cluster is detected, a trip is identified. A trip occurs between two meaningful clusters with time difference greater than a pre-defined threshold for trip duration. As for path and mode identification , mode inference and map-matching are in the most cases treated in two different stages, with the risks discussed above - i.e. deterministic map-matching and mode inference findings may contain mistakes in route choice models if the detection/inference is misled. The approach proposed by (Bierlaire, 2008) tackles theses issues. Nonetheless, although the methodology is applicable to GPS records, by the lack of such data the authors only presented results from data recorded from an interview. The probabilistic multimodal map-matching (PMM) (Bierlaire, 2003) capitalizes on this framework and renders it applicable to route choice data collected from Smartphone. The PMM method proposes a probabilistic measurement modeling framework. The framework takes into consideration and deals with the inaccuracy and sparseness of the data. The errors in both GPS and the transport network data are modeled probabilistically, and the unknown trajectory between the consecutive GPS points(path inference is modeled with a realistic travel model. These features of the algorithm dodge with potential biases introduced by traditional deterministic matching and shortest path assumptions. The proposed unified Smartphone measurement model, combines several sensors measurement models, and returns the likelihood of observing all the measurements from the various sensors on the multimodal path at the given time. PMM deals with the general map-matching problem when a trip is multimodal, and the modes are unknown. The algorithm identifies the physical path and modes of trip simultaneously, by utilizing various kinds of Smartphone data. As the path and mode inferred simultaneously the network attributes are exploited to contribute to the mode inference. As an example, a segment labeled with bus mode should follow bus lines. The method exploits the potential of the GPS, BT and ACCEL records. At first, probabilistic measurement models are derived for each sensors in order to capture the data generation process. Subsequently, a Smartphone measurement model is constructed merging the multiple sensor inputs in a unified framework. The output of the Smartphone measurement model is a set of candidate paths each one associated with a likelihood of observing the measurements on this path. The paths are multimodal , i.e each segment of a path is associated with a transport mode, and different segments of the same path might be labeled with different modes. The GPS measurement model has a two-dimensions latent status variable, including the true location and the mode. The BT and ACCEL measurement models are empirical and calibrated based on the data. The BT measurement model is employed to distinguish between public and private transport by monitoring whether there is a nearby BT device or not. The ACCEL measurement model assists the differentiation among walk, bike and motor modes. For details on the measurement models as well as the PMM we refer to (Chen, 2013). The PMM method is implemented as a software package in C++. It takes Smartphone and OSM network data as input, and generates probabilistic map-matching paths. Numerical analyses of the performance of the algorithm with respect to the identification of the mode by means of the different sensors are provided in (Chen, 2013). 36 data sequences that are known to have one single mode are used. The results indicate that the proposed method achieves high accuracy in mode inference, and that Bluetooth and acceleration records contribute to the identification of the mode.

5.6.1 Modeling : Route Choice Models and Smartphone Data

(Bierlaire, 2008) studied the resolution of route choice data inaccuracy using their framework.The framework namely Network-Free aimed at estimating the route choice data using a network-based model. The main idea is to link the observations collected, since they are separated, based on the paths on the real physical network. The method of using data collected from GPS point is complemented by defining physical areas in a network this is called Domains of Data Relevance DDR. In Figure 5.13, the GPS data (depicted as small circles) while DDR (depicted as dashed line circles) are presented for illustrating the idea. The authors' assumption concerns the size of DDR circle which varies depending on the data accuracy (sufficient satellites signal quality of the received GPS data). Based on measurement equations the network free data are linked with the network through the DDRs. They relates several paths to an observation while the missing observations are not taken into account by the data treatment and assumptions. The observation is a sequence of itinerary-related pieces of data (e.g. sequences of GPS records). Measuring the likelihood results the probability of sequences observation given a path.

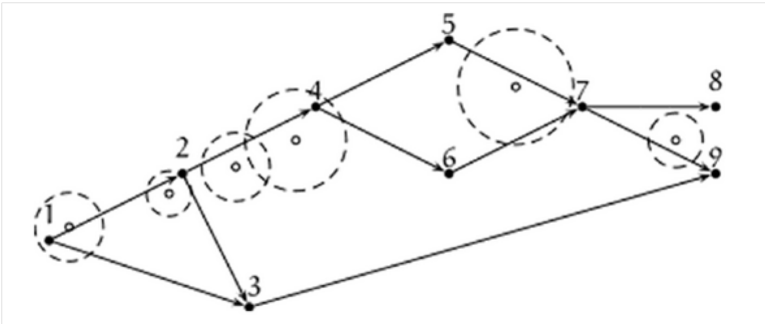


FIGURE 5.13: Example of GPS records and DDR Source : (Bierlaire, 2008)

5.6.2 Overview of map-matching algorithms

The map-matching algorithms were introduced in 90s as a simple geometric algorithms. The algorithms have known several changes over time, sometimes considering the dead reckoning technique (Kim, 1996), however the main principal of matching the location data obtained from GPS points to a closest node, road segment/arc, remained the same. The historical data previously determined such as vehicle direction (bearing, northing and easting using the gyroscope and compass), the speed (by measuring the wheel rotations or by the mean of a built-in accelerometer), are taken into account by the dead reckoning which compares this data with the newer data to estimate the new position of the vehicle. Since map-matching algorithms are used respecting several other scenarios too, and the data are not always vehicle-based which results the unavailability of input data for dead-reckoning techniques, most of the works were done on the basic geometric techniques (C. E. White and Kornhauser, 2000),(Bernstein and Kornhauser, 1996). The research have shown that the matching process is a complex task and Bernstein and Kornhauser argued that, a considerable amount of attention should be given to curve-to-curve matching usage and to topological information in order to get a significant results (Bernstein and Kornhauser, 1996). According to the usage of mathematical-based operations for the algorithms' results, especially at crossroads (junctions and U-turns), the geometric algorithms were not quite efficient. Consequently, new algorithms appeared which solved the problematic by using historical data (previously matched points), GPS observations with time and date information and digital map data (topological information). The later algorithms are called topological algorithms, the particularity of these algorithms is to label the chosen segments with weights (scores) based on the observations and the usage of physical road information to avoid the connection between unconnected road lines (Greenfeld, 2002) , (M. A. Quddus and Noland, 2003). In addition, some probabilistic algorithms were developed for this purpose. The aim is to use probability theory to identify the candidate segments, measure the weighting scores and common characteristics between segments and candidate points (Y. Lou and Huang, 2009) (Yuan, 2010). Moreover, later on another complex algorithms were introduced relying on the fuzzy logic (M. A. Quddus and Ochieng, 2006) and Kalman filter (L. Zhao and Noland, 2003) techniques. However, this kind of techniques requires larger amount of input data compared to geometric and topological algorithms. The techniques uses more computational power and slower matching times. All the aforementioned probabilistic, topological and advanced algorithms rely in the basic of map-matching geometric techniques.

5.7 Particular features of an application-based surveys

Today, several smartphone applications and applications are available in the market. Some of them are dedicated for the same subject and even the same matter but each of them has its level of popularity. This lead us to a problematic with a number of questions regarding the developing conditions for a good mobile application which has (i) a good acceptance theory (ii) well-constructed view of loyalty, (iii) less consuming with regard to energy and finally (iiii) enjoyable environment.

5.7.1 Acceptance Theory on Mobile Services and Applications

The individual acceptance of mobiles services have not increased rapidly according to many recent researches (Bouwman, 2012) or shown a sustainable evolution (Carlsson, 2005). According to service providers point of view the market of mobile services has not reach its optimal level yet and this problematic introduces many questions that occur concerning the slow acceptability degree. For instance:

- Why smartphone users only keep using one kind of application?
- How do the point of view of users with respect to a particular application changes according to each context and situations?
- Does popularity impacts the application acceptance process?
- What is the impact of application developers?
- How smartphone OS platforms are chosen by users?
- What is the role of social networks applications?

In recent years, many acceptance model have been developed. In a research conducted by (Venkatesh and Davis, 2003) a new unified model namely

UTAUT was developed based on a combination of different traditional acceptance theories. The main principal of this model evolves eight variables where the first four are the four fundamental determinants of intention and usage behavior and the other four factors relates to the key relationships. These variables are used to provide a clear comprehension of individual acceptance of Information Technology artifacts to managers and researchers . The first four determinants are presented as Performance Expectancy, Social Influence , Facilitating Conditions, and Effort Expectancy while the last four variables links to Age, Gender, Usage Voluntariness and Experience. Figure 5.14 illustrates the link between the core determinants and the moderating factors related to usage behavior.

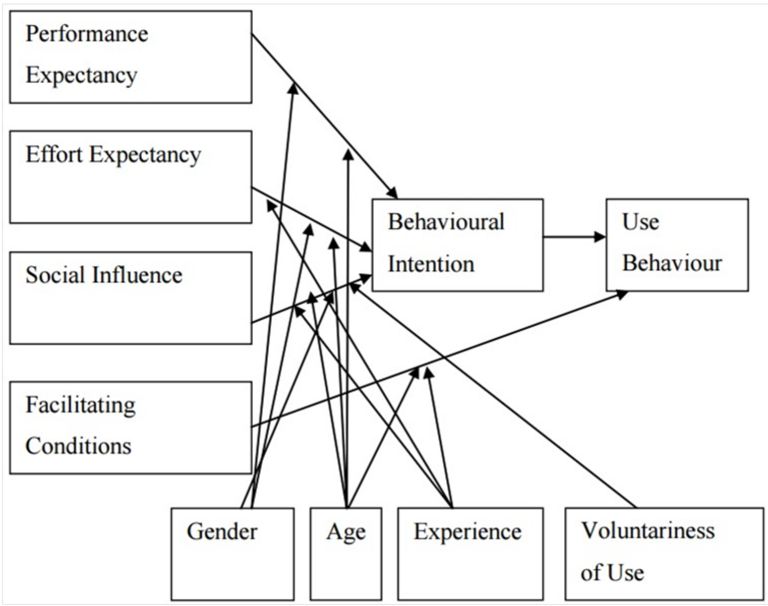


FIGURE 5.14: Unified Theory of Acceptance and Use of Technology Model

In (Venkatesh and Davis, 2003), the authors defined the Performance Expectancy as the degree of user belief with respect to the performance improvement after using a particular technology. This variable were also introduced in other technology acceptance theories such as Technology Acceptance Model (TAM) defined as Perceived Usefulness. Than we have the Effort Expectancy variable which is the degree of easiness of using a particular technology or a system, this variable is also introduced in TAM model as Perceived Ease of Use. The unified model strongly rely on the moderating factors such as Experience, Gender and Age. For instance, it was found in recent studies that the younger women are more likely to use new technologies. Additionally, another determinant is the social influence which is the degree to which the user perceives how important is point of view of others in using the technology. In Theory of Planned Behavior (TPB) and Theory of Reasoned Action (TRA) the social influence is introduced as Subjective Norm. The fourth determinant is Facilitating Conditions which is the degree in which the user perceives that infrastructures may support his/her usage of a particular technology. In the Theory of Planned Behavioral (TBP), this determinant is introduced as perceived behavioral control. Chapter 5 explores the aforementioned acceptance theory models.

5.7.2 Loyalty and Environment

Retain an existing user is 5 to 10 times less costly than find a new user. The issue of retention takes several dimensions according to the user's category. The user is loyal when he finds his interest and especially the information he needs in the application, otherwise it becomes a waste of the time. To prevent the negative feedback of the user, it is recommended to:

- Know and understand the users that the application deal with.
- Construct an honest strategy between client-researcher.
- Ensure confidentiality without having to repeat it - better than mention it.
- Make the standalone - no need to say more about the features
- Ensure the functionality without limitations
- Design a good compatibility - which provide the users with more enjoyable environment.
- Avoid the complexity.
- Know awarding the user in case of loyalty
- ...

5.7.3 Amusing

Promote user's interest and information attract less the user. The application also need to have a light and enjoyable use. From there, the design is being important. In the majority of application, the user has the ability to change the theme, mode and even the colors because this ensures the property. For example, regarding the geolocation of the user it is recommended to use personal-navigation assistant (PNA). Besides, the application must also ensure a good performance and should be used without causing a performance problem to the device. Here, we talk about energy consumption.

5.8 The existing projects of application

5.8.1 New applications oriented to the multimodal

Connected services open on a different world than the automobile. For example, we can consult the timetable of aircraft on the Garmin GPS in the Nulink website. Same approach at Audi, where this information will come to enrich the offer Connect. For its part, Citroen has designed an Internet portal on mobility. Accessible outside of the vehicle, it is named Multicity and is defined as an application « Route and transport ». On iPhone screen, a user can caompute a door-to-door route, regardless of the mode of travel. The service offers combinations of public transit in the city (bus, metro, tram) or over longer distances (regional train, TGV, plane). It takes into account passengers and goes up to calculate the price of the taxi. Multicity can also book a rental with National Citer vehicle online and deliver it at home (service Call Car) in major cities. Peugeot, with its offer Mu (available in Smartphone) can also rent vehicles, including electric Ion, and even a bike or a scooter to try other mobilities.

5.8.2 Ready for use mobility services on a Smartphone

In the urban environment, the citizen is no longer single. He can locate himself on a map (with Google Maps) and locate the metro station, or the nearby bus stop. There is generally an application to take knowledge of transit lines, in the best cases, schedules. Otherwise, the Pages Jaunes are responsible to find the phone number of the taxi companies. But in some cities, we can directly call a taxi on the fly by connecting to a server which will then seek the drivers available. The answer is given with the type and color of the car, as well as the time required to reach, appears on the screen when the taxi is found. All the bricks exist, thanks to the efforts of developers and thanks to the opportunities that provides Internet (when it is not the users themselves that feed data). The ideal would be of course when a telecom operator (or service) can integrate these services and to propose a single portal.

5.8.3 Optimod' Lyon Project

Lyon city intend to implement an innovative system to make more efficient daily mobility. If greater Lyon has many infrastructures and transport services, here as elsewhere, all these data are split between the different organizations (the Region, SNCF, councils, Sytral, the State, corporations of highways, the Greater Lyon) who manage transportation. The solution? Consolidate these info and return them in the form of a single browser. It is the objective of Optimod project ' Lyon, certified by Ademe and inscribed in the investments for the future. With the help of partners such as IBM, Orange, Renault Trucks, the project will develop a « data warehouse», which will be supplemented by data from new fixed and mobile sensors for 36 months. The objective was to propose by 2014 an innovative services on the Greater Lyon:

- A forecast of traffic (first in Europe in the urban context), which will allow to optimize the management system of 1 500 intersections of the agglomeration, which ensures priority to public transport and the fluidity of road traffic.
- The development of all modes browser, in real-time, on mobile phones, which will be a first in Europe and probably throughout the world. This service will develop a multimodal calculator giving all modal options to go from point A to point B on the territory, according to different criteria. It will integrate the notification system in real time in the event of unforeseen circumstances, and all functions associated with the GPS and will provide the user a real-time urban navigation tool.

5.8.4 SYSMO 2015

This vision is being realized also in the Paris region, both in the Saclay, Versailles-Satory and the territory of Saint-Quentin-en-Yvelines, where several competitiveness (Mov'eo, system@tic and Advancity) have implemented the SYSMO 2015 project which aims to explore the new mobility. Worn by the RATP, with the contribution of many partners including the SNCF, Renault, Valeo and Continental for automobile, it will take place for three years. Its budget is EUR 24 million. The idea is to provide real-time optimal and customized answer to the question « how to get from point A to point B? »This response gathers all offers from transport (transit, carpooling, car sharing, car or electric bike, in complementarity with the individual car) using the Smartphone.

5.8.5 Modalyzer

Modalyzer is a fully automatic mobile application capable to recognize eight different transport modes when the user moves with its Smartphone. The application allows a visual feedback on all distances. Statistically, Modalyzer allows learning of real-time personal user mobility without the calling for users' effort. Modalyzer is developed as a research tool. It supersedes the traditional, such as questionnaires, which offers the user to fill out forms in paper or digital format during each end of the day. Among the advantages of such application is that it directs the user to research in transport in a fun way, so researchers will have added value in research. Users contribute with their data to support research on mobility projects and thus actively shaping the mobility of tomorrow. No data containing information about the user will be transmitted. Contribution to the research and help shape the future of mobility may be rewarded.

5.8.6 Mobi-survey

Mobi-Survey is a revolutionary solution that allows us to create and disseminate in a few minutes and without any development, questionnaires and forms on mobile and tablets. These collection medium can operate in either connected or disconnected, for use in all situations. A central platform that allows the user to monitor deployed applications and automatically synchronize data with mobile devices. The data collected are available to be processed in real time with statistical tools or to be recovered directly in by enterprise applications.

5.8.7 Mobi-lise (TamTamTop)

An example of a social network for mobility. TamTamTop aims to observe the behavior based on the concept of sharing diary information of its users. The user provide information to start the registration such as (family and first name, full adress, telephone, and profession etc). The concept is to share the itineraries including the motives. The main advantage of this tool is the opportunity that it provides to save the mobility history as well as the planning of future trips. The drawbacks are the imperfections, the tool need more advanced techniques. The lack of the automatic tracking of the mobility using GPS/Wi-Fi which is currently a basic technique for user tracking. Besides, the validation of users' information is not possible since the application does not provide a way to confirm the results systematically (e.g. using sensors)

5.9 Advantages and Limitations of the Smartphone - Based Surveys

Compared to traditional surveys, smartphone-based surveys may only reach audiences who often use their smartphones. According to Parker, (Parker MJ, 2012), surveys on Smartphone can encourage respondents to participate. They have adopted a structure which consists of two phases:

Phase 1: sending an email to invite to participate in the study by filling out an online questionnaire.

Phase 2: sending a second email notifying those who did not respond to the phase 1 that a research assistant will move to their office during the next two weeks to proceed a survey in place.

The result of such approach was 49 % and 99% for phase 1 and 2 successively. Missing responses were quite similar and the characteristics of participants in Phase 1 and 2 differ. Most interested participants of with the topic of the survey tend to respond in greater proportion to Phase 1. Smartphone-based questionnaires can easily be fulfilled with less missing data and possibility of additional information.

The smartphone-based surveys or application-based for more precision are traditionally led easily by communicating a link for downloading the application. These kind of surveys have several advantages compared to other surveys (e.g. online surveys). Among these advantages, we find the respondents identity validation by the mean of sensors. By sensing the respondents we can know if the respondents are a real persons and not a robot program. Using this kind of surveys increase the amount of information, and many extra data can be collected. For instance, the users are surveyed and sensed in the same time. The information about user's activities or location information are easily collected using the sensors embedded inside the device. The transfer of additional documents is made easily too such as pictures, recorded voice , videos etc. Another advantage is the real-time data that can collect the Smartphone. Many researchers have led a studies using Smartphone to collect real-time data which are systematically collected without any users contribution. The surveys are also easy because they can be completed everywhere and at any time for example, in case that the respondents is walking or traveling etc. Concerning the advantages, the smartphone-based surveys are first limited due to battery limitations especially in case of high data quality which requires a high computational power (high accuracy GPS and accelerometer). In addtion, several sensors are not available in certain contexts (e.g. GPS unavailable indoors, Wi-Fi unavailable without nearby APs).

Depending on each Smartphone device, the surveys should be developed (iPhones to Android phones to Windows phones). There are several operating systems and for each one the application should be developed. Moreover, since new technologies have been introduced to the market such as tablets, most of them have different operating systems and each operating systems have several version, thus the development should be adapted to each of them. Along all of these advantages, the mobile surveys remain not suitable for long and complex questions due to respondent independency and boring issue. Since the smartphones are used in multitasks, the respondent are more likely interested in closing the survey application to do something more interesting or to play a game, send message make calls etc or they main never complete the survey. The application size and computational properties are also taken into account and the surveys application should not slow down the device performance. The developers should know that the Smartphone hold several other applications and files such as pictures videos and music , and the Smartphone are of limited hard drive disks. A good application should not have only acceptance value but also adoption which means that it is more appropriate that the respondents answer to the surveys without uninstalling it at the end. This promote the real-time data analysis. The network coverage is also a problem, thus the developers should take into account the off-line mode of making the surveys (people from the city are more likely to have a good internet connectivity). In addition the knowing his/her audience is important too, if the majority have Smartphone now others still don't have it and this decreases the smartphone-based surveys efficiency. Traditionally, the surveys could definitely be skewed to more affluent and/or younger audiences.

Compared to other data collection method, Smartphone data collection system aim at collecting more coherent data. These data are lesser incomplete data, lesser errors , consistent responses and the data are delivered. Despite of the higher cost of their development , the overall costs of implementing and lunching the survey data collection system is lesser compared to traditional data collection system. Smartphone data collection has potential to improve timeliness, data integrity and reduce costs.

Chapter 6

User Acceptance of Mobile Applications

6.1 Summary

Smartphone application is the result of technology advances of the recent years. Services and applications can be easily downloaded by the mean of the Internet, Media, IT and other advanced telecommunication tools. Several studies and researchers from the Telecommunications business lead their research in view of their own benefit. Additionally, multiple issues in mobile technology have been under study for a long time in Information Technology and Information Systems studies of numerous scientists, Telecomunication operators, Smartphone producers, Smartphone services/content suppliers and application developers. Whereas, a standout amongst the most fascinating research area is Smartphone application developments. It has been contended that the eventual fate of Smartphones is based on the services its provide (Carlsson, 2006). In the latter study, authors' forecast was possible because of The rapid advancement in telecommunication innovations. Smartphone applications can be installed on the cell phones out from traditional voice services, and it is was proven that the advantage offered by such services are profoundly differing (Bouwman, 2012) (e.g. using Viber to make calls). There are various distinctive Smartphone services and applications accessible for free or to purchase online in the market of telecommunication. However, several earlier studies reported that the individual acceptability and adoption of Smartphone applications have not been increased (Bouwman, 2012) or demonstrated an asynchronous design (Carlsson, 2006). From service suppliers' viewpoint the Smartphone applications market has not yet achieved its ideal level keeping in mind the expected goal to give back the large ventures made by Telecommunications businesses. Then again, in light of services suppliers' point of view, one can brings up numerous issues and problematic because of the moderate acceptability and adoption ratio. For instance:

- Why Smartphone application users usually adopt a kind of service and the other?
- How user's application recognition shifts in various situations and circumstances?
- Does demographic view assume a noteworthy part the acceptability process?.
- What is the contribution of Smartphone services suppliers ?

- How Smartphone services platforms are met by the users?
- What do the social network service plays as a role?

It is important to mention that the previously stated questions are only few from numerous others, and we cannot cover every one of them in the present chapter. In this manner, we attempt to address just the aforementioned questions which are more applicable to such purpose. A Smartphone application will be accepted and used by a user, if this application provide a significant value. That is to say, there are a several criteria that make an application successful or a failure.

One of the critical problem that may influence the individual application acceptability is the users' personal information. Age, gender and training, and revenue are considered to be the important for an application acceptability. For example, an application may fit to men's interest (phone gaming), while, a few application may just target females 'mobile web browsing' (Jung and Wiley P atton, 2009).

Finally, we will explore a problematic which relates to user acceptance and mobile network operators. Whereas, just recently application platforms has become a critical issue in the telecommunication industry. Over a long period, mobile network operators were the only providers of mobile applications. Even though, just recently Nokia and Apple and other Smartphone manufacturers have entered into mobile communication market. These companies have already began to use their own platforms as a mean of application supply. Phone producers such as Windows Phone with 'Microsoft-store' and Apple with 'App-store' are contributing to the application delivery.

Typically, mobile network operators takes a long time to develop a new mobile application. This is basically because of standardization issue. On, the other hand, device producers cooperating with application designers lead faster application development process. A Smartphone application platform in this situation goes about as the middle-ware between the gadget and the application developers. Smartphone application platforms are currently an essential tool due to the emerge of utilization of such kind of devices; To study the user acceptance of Information Technology (IT), many acceptance theories like Technology Acceptance Model (TAM) (Davis and Warshaw, 1989), Diffusion of Innovation (DOI) (Rogers, 2003) or Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh and Davis, 2003) have been proposed and been the subject in numerous earlier studies. Over recent studies, the acceptability and usage of Smartphone applications were the key topics of the study. Many earlier studies were keen on understanding how and why innovative and developed tool spread.

Long list of related researches in which, analysts have utilized one or a combining of the acceptance theories planning to discover a response to promote the comprehension of the acceptability and manageable utilization of the Smartphone apps. For instance, application design and characteristics (Bouwman, 2012) and (Nysveen and Thorbjorn sen, 2005) and mobile device merchandising (Mahatanankoon and Lim, 2005), acceptance of advanced mobile services (Lopez Nicolas and Bouwman, 2008), service platforms (Halteren and Pawar, 2006) and the effect of use context on the mobile ticketing (Mallat and Oorni, 2009) have been widely under research. However, none of these earlier research results have demonstrated a complete answer for the acceptability and adoption of mobile application issues.

6.2 Individual IT Acceptance Theory

Several theories have been already developed to clarify the idea of IT acceptance. The basic idea of the majority of such theories is based on evaluating individual intention and forecasting user's behavior toward the new technology. In any case, the research argue that all the achievements which have been proposed until now to anticipate individual's behavior are not adequately enough to give a complete comprehension toward this issue. A short demonstration is depicted in Figure 6.1 concerning the basic idea of Information Technology acceptance.

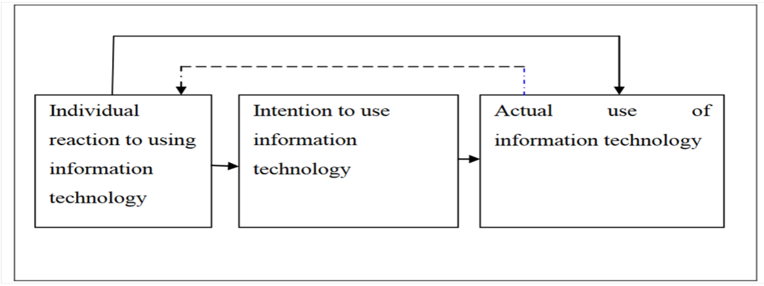


FIGURE 6.1: Basic design of user information technology acceptance Source: http://www.jstor.org/stable/10.2307/30036540

As it is depicted in Figure 6.1, the individual behavior toward the utilization of IT influences the intention of using it. As result, we can predict the present behavior of the user based on the that intention. Several IT acceptance theories will be discussed and compared in the following subsections.

6.2.1 Technology Acceptance Model

The idea of such theory is relating the ease of use and usefulness of using information technology to individual's acceptance. Many of prior research which adopted the individual's acceptance of mobile application as its main research have used this model. Which make it the most widely adopted theory. Developed by Davis and a group of researchers in 1989 (Davis, 1989), they found that there were two variables impacting the user behavior toward IT. First the most valuable variable, where basically the idea is the confidence the user has concerning the possibility of improving his performance when using such IT; the other variable is related to the ease of use, the value of how the individual find the use of technology simple and easy or the opposite. In addition, their idea assess that an individual's usage manner results their main intention to use, and their attitude of use is influenced by the main intention. Several prior studies (Bouwman, 2012) (Ha and Choi, 2007) (Hong and Kim, 2006) (Hong and Tam, 2006) which adopted TAM have shown that the model is an efficient method to predict and understand individual acceptance of new technologies by giving a clear idea of a such acceptance. Whereas, this theory were not fully accepted by many researchers since it has been proven in other researches that projects based in TAM are not the best solution to better understand the user behavior while other variables are seen to be essential for this purpose, (Kulviwat and Clark, 2007) (Fishbein and Ajzen, 1975) (Nysveen and Thorbjorn sen, 2005) (Stern and Bienstock, 2008). The authors argued that variables such as Demographic Influence, Subjective Norm, Enjoyment, and Critical Views (Markus, 1987), must be taken into close thought in individual IT acceptance. As a result, Technology Acceptance Model TAM being based on two conditional variables (i.e. Ease of use and usefulness) remain significant; whereas they are not the single variables that assess the acceptance of individual toward IT. To better predict and understand the user it is recommended, by many researchers, to take other variables into account.

6.2.2 Theory of Reasoned Action

TRA and the idea behind of it has been first proposed by Fishbein and Ajzen (Fishbein and Ajzen, 1975). It argues that the available information attract more the individual interest. In its original design, personal conviction of the individual, mentality, intentions and social feedback were considered as the determinant factors to better understand the link between behavior and attitude. This theory is used to clarify and predict an individual's decision-making attitude, is the most persuasive fundamental hypothesis about user's behavior. It gives a focus to behavioral action related to a user's relative importance of intention to carry out a behavior, which is the fundamental manner for understanding users behavior. Basically , two variables have an impact on behavioral intention. One is related to personal feeling to carry out the behavior, a user's positive or negative feeling and this depends on each individual. The other variable is subjective standards that is a sort of social elements, it argues that the individual is influenced by the social feedback to do or not such behavior. A behavior is conducted according to the individual conviction and acceptance as well as the subjective norms, such action is highly influenced by each user mentality and his strong conviction. In summary, an individual's behavior is conditioned by this individual convictions or/and subjective norms. Note that the social influence is found as Subjective Norm (SN) in the literature.

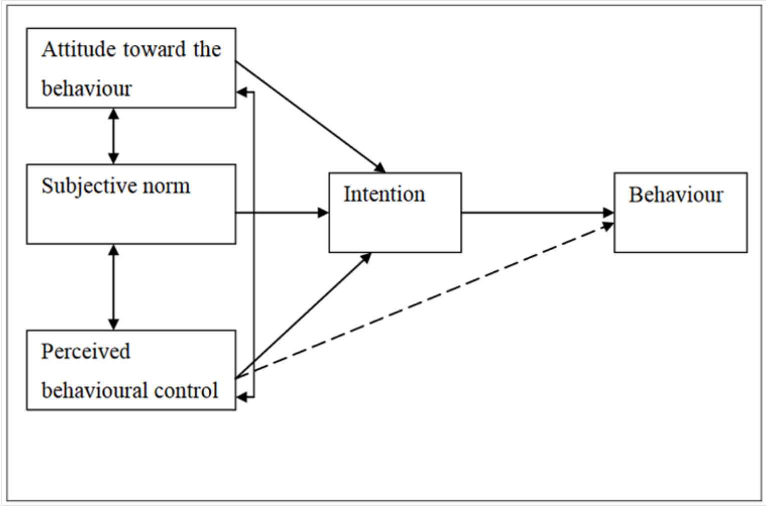


FIGURE 6.2: Theory of Planned Behaviour Model

6.2.3 Theory of Planned Behaviour

Theory of Planned Behavior (TPB) proposed by Ajzen (Ajzen, 1991) comes after the theory of reasoned action TRA which is an extension of it. The basic idea is that the understanding of relationship between behavior and attitude is less possible using the principal variables of TRA, thus, other variables such as perceived behavioral control and social elements are added to TRA to assess the comprehension of the relationship between the behavioral intention which is the attitude and usage manner which is the behavior. The author argued that TRA can only predict individual behavior when no constrains are present. However, according to TPB, the attitude and behavior of an individual is influenced by his behavioral intention and in contrary his subjective norms have an impact on his behavioral intention, which has the same perspective with TRA. Besides , TPB includes the idea of perceived behavioral control in its theory and takes into account the principle that an individual's perceived behavioral impacts the behavioral intention, consequently this influences the behavior. The perceived convenience and control decide the perceived behavioral control, which tend to an individual's ease or difficulty when carrying out the behavior. Additionally, the perceived control of the individual gathers the opportunities and resources needed to carry out a specific behavior, while the perceived convenience is the evaluation of such factors. To conclude , TPB has as a main hypothesis that performing a particular task can be explained due to motivational variables that have an influence over the behavior. Thus, it is the person's willingness to make an efforts in carrying out a particular behavior. Consequently, performing a particular behavior has a positive correlation with the individual intention (Ajzen, 1991).

6.2.4 Unified Theory of Acceptance and Use of Technology

In addition to others theories discussed above, authors in (Venkatesh and Davis, 2003) advanced the thought and innovated a theory UTAUT that focus on the information technology acceptance. Many of prior researches which their main subject were the information technology acceptance with several prediction variables have been confirmed.

Following their theory, different traditional IT acceptance theories were joint to build a final unified model. The basic design considered four principal factors that assess the usage and intention where other four variables are the key relations. The point of conducting this theory is to promote a more profound comprehension of individual and organizational IT artifacts acceptance to researchers and operators. Whereas, new keywords were introduced such as Encouraging Conditions , Execution Expectancy, Social Influence and finally Exertion Expectancy as the principal and absolute factors of individuals acceptance and usage behavior. Moreover, age, gender and experiences as the key elements for understanding the acceptance and usage behavior. Figure 6.3 explores the links the between these determinants and additionally the key element linked with usage behavior.

As explained by Venkatesh in (Venkatesh and Davis, 2003) the degree an individual is confident that utilizing a specific framework or innovation will boost his/her performance is called Performance Expectancy. As you may confirm in (Venkatesh and Davis, 2003), comparing traditional acceptance model such as perceived usefulness in Technology Acceptance Model to performance expectancy, they appear quite similar. Another important determinant according to UTAUT is the level of simplicity (Ease) of utilizing a framework or a technology, which refers to the perceived ease of use that we find in TAM model. The model treating behavioral intention is highly influenced by age, gender and experience, a conceivable reason is that more youthful women as an example in most of time prefer to try new technology. In addition to this, another important factor is related to entourage beliefs on his/her usage. This impact is mentioned in the traditional theories as Subjective Norm. To conclude we have the Facilitating Conditions which are the amount on how a user feels that the design promote the usage of a technology or a framework. Such determinant relates to Theory of Planned Behavioral.

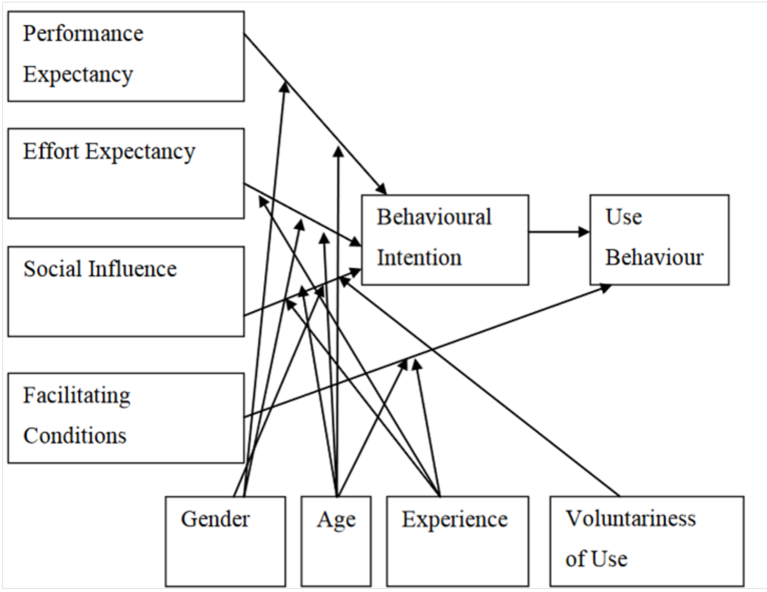


FIGURE 6.3: Unified Theory of Acceptance and Use of Technology Model Source: http://www.jstor.org/stable/10.2307/30036540

6.2.5 The spread of Innovations

The idea behind this theory DOI is based on the manner and the reason that a technology or an information spread among network of individuals and societies. Rogers in (Venkatesh and Davis, 2003) compared the diffusion of information and technology as the IT communication process by the mean of certain ways among the people in a social system over the time. Moreover, the author argued that idea , method of communicating a message between people, the period required for an individual from a social gathering to complete a development and organizational system where units having shared objective are the key components in diffusion research. Theory of Diffusion of Innovations assess that the diffusion procedure is based on five essential phases which are as follows:

1. Acquisition: is the first relation of the user with a new innovation, the user still in his/her first level.

2. Persuasion: is a phase that the individual really gets comfortable with the innovation and starts to be occupied with, consequently looking to get more experience about the innovation.

3. Decision: is the most complicated phase where the user decides in keeping the innovation or leaving it. The advantages and drawbacks are evaluated and the individual analyzes the results.

4. Implementation: in this phase the innovation is accepted and the individual look to find more advantages and ways to explore it.

5. Validation: is a second complicated phase to keep or leave. In this phase, the individual decides if he/she keeps it or not.

According to DOI, there are several key elements that influence the decision of an individual when making a choice to keep or leave an innovation. The elements are briefly listed as follows:

1. Upgrading Advantage: demonstrates how new version has been designed to be more efficient than previous version. This promotes a new experience where the user try to discover new feature and advantages.

2. Compatibility: this element assess the suitability of the new innovation on the user environment.

3. Ease or Difficulty: this element evaluate the degree of ease of the innovation, more it is complex , more the individual is likely to leave.

4. Trialability: the user test of the innovation, more it is easy more the individual is likely to keep.

5. Observability: demonstrates the way the innovation spread among social system. The more it is correctly developed, the more probable it will be kept by individuals.

The last critical issue that Rogers mentioned is the clustering of the users. Rogers, categorized the people into five clusters:

1. Designers: the designers of innovative ideas.

2. Early keepers: one of the largest group, a group of young individuals with high education level and revenue.

3. Early majority: who are required quite a while to embrace another innovation, persons who have over the normal economic wellbeing.

4. Late majority: the individual who wait for the others decision.

5. Laggards: the last cluster according to Rogers, the individuals among the cluster are related to cultural issues.

6.3 Telecommunication Technologies

Successful mobile service/application is conditioned by the advance in mobile technologies, the user sustain the usage of a service only if he/she feels comfortable with. Thus, to better know the process of users' acceptance, we present an overview on technological issues. Global System for Mobile correspondences (GSM) was introduced in 1982, a short period after this telecommunication revolution, new technologies emerged such as Wireless communications. Other technologies, namely, Enhanced Data rate for GSM Evolution (EDGE) and General Packet Radio Service (GPRS) came after to upgrade the GSM network. The main advantage of a such technologies is the low-cost to receive data and being IP-based networks. In addition, EDGE gives a possibility to use an IP platform which is separated from communication standards. The highly advance and quick materials development including the growth of the internet connection requests through cell phones has prompted the third generation 3G of mobile phones. Thanks to 3rd generation mobile technologies we are able to have wide-area wireless for phone communications, Internet access, video calls and mobile TV. The new generations are more advanced which is carried out in an advanced networks infrastructures to operators giving possible access to different services by users. In addition to this, the growth of users' demand to access the Internet network in their mobile devices has promoted the introduction of 4th generation of the network technology, namely, Long Term Evolution (LTE). Thus, the telecom operators having 3rd generation network technology are upgrading the system to LTE. Today, LTE is the latest version in the mobile network technology which was developed to increase the coverage and performance of the networks. The prior studies has argued that GPRS network will soon be replaced by LTE. To conclude, by the mean of LTE the IT will live a significant growth especially multimedia.

6.4 Service Design

This section explores the design of mobile services; the main goal is to list the key factors which promote the keeping or leaving a mobile service. The objective of this section is to explore all the services' design that were not mentioned in the aforementioned theories. Thus, we can evaluate the importance of service design compared to traditional theories. Finally, the conclusion of this section will reveal some insight to our critical view to the individual IT acceptance.

6.4.1 Mobile Application Innovations Typical Features

Smartphones having programs already installed which are necessary to help on the users in their daily life. This new technology opens more possibilities for optimal life and gathers several advantages. Basically, smartphones having possibility to install applications and gadgets allow the users to reach real time data, asses the social networks, communicate, download applications, online shop and bank transaction etc. These devices give possibility to run several tasks everywhere and in any time and the manner in which they were designed make the acceptance and the relation user-device unavoidable. Whereas, likewise other technologies, prior studies have argued some failures related to mobile applications acceptance (Lopez Nicolas and Bouwman, 2008).

In general, the Smartphone holders need to use the mobile services and applications everywhere and whenever they need it. i.e. the mobile phone is now used wherever we are and whenever we want to. Consequently, we can highlight that the mobility assess a critical part in acceptance or the adoption of such services. Mobility is important with two fundamental parameters time and location. The devices are carried out at anytime and anyplace. Although, the advantage and benefit we are able to gain from mobile services, the mobility attribute is not the only important device's characteristic. Mobile programs and services, because of its variety, ought to be ordered into a wide range of approaches to better understand users' acceptance toward it. Thusly, it is important to cluster these services and applications in profound and reasonable way. The clustering process of these services can meet with some problems due to the technology advances and several groups can share same services, whereas, classified them as follows:

- Communication services (mobile call services and SMS)
- Information services (mobile weather, map information and mobile news)
- Entertainment service (mobile game, mobile music an mobile TV)
- Web 2.0 (mobile health)
- Transaction services (mobile banking, mobile shopping)

Services such as telephony and SMS have been available to users for a quite long time, while some other services have recently been introduced and are available to use such as social network services (Facebook, Viber and WhatsApp) and Mobile monitoring RFID (Radio-Frequency Identification) information. While some services are designed specifically for individual, some other services are only designed for groups.

Several services were introduced to consumers for a long time such as voice calls and SMS, while just recently, open source programs including apps stores have been available and are now accessible we can list for example, Facebook, sensors, camera etc.

Among the existent applications we have a personal-based and group-based. In addition, while different administrations give data to the clients, for example, notices by means of SMS. Several mobile services are used to access online diary information, others uses notification system and some of them explores sensors embedded in the smartphones. A large group of applications does not require user manipulation while some others require users participation purposes and the basic example is the messenger or uploading a video in YouTube etc. Shao in (Shao, 2009) classified mobile services into two groups, (i) consumption, and (ii) contribution.

Generally speaking, the services can be classified in two dimensional manner, (i) application stores and innovative design such as mobile games, and (ii), already existent tools converted to mobile such as banking, email and using sensors data which were specifically developed for mobile location based services or mobile augmented reality.

6.4.2 Usage Situation

A mobile service should offer particular standards to its users so that they will easily adopt it i.e. the service should be able to enhance the users' performance while they are carrying out any task. Yet some researchers conducted on the field have shown that the situation in which the service is used has an effect on both the users' attitudes and behaviors. In case the situational features fit what the users need to carry out any specific service, then the approval of the latter will be enhanced.

Today, mobile phones have become so indispensable that people cannot move without them, so they carry them everywhere. As a result, the usage situation has widened. However, in some cases (areas) it is restricted and the use of mobile services is impossible due to the absence of technological infrastructures.

In summary, the acceptance and adoption of mobile services depend primarily on the usage situation. In most cases users expect to have a ready availability of mobile services when and where needed. That is why mobile service companies should take into consideration the situational usage of mobile phones while making their services.

6.4.3 Mobile Users' Adoption of Innovative Technology

Innovations in mobile technology have had a great impact on the users' service perception. Cellular systems have experienced exponential growth over the last decade and there are currently about two billion users worldwide (Goldsmith, 2005). The highest number goes for mobile devices and particularly smart-phones which have become an indispensible tool and a part of everyday life worldwide, exceeding largely the number of the fixed access lines. According to International Telecommunication Union (ITU, 2016), mobile-broadband networks (3G or above) reach 84% of the global population, and the total number of mobile-broadband subscriptions is expected to reach 3.6 billion by end 2016. Today, mobile users can have access to a wide range of services and information: look up information about a health condition, do online banking, get information about a job and submit a job application, do business, take a class or get educational content, etc. So if those services fit and meet the users' preferences and expectations, then users' adoption of mobile technology will be greater than before.

6.4.4 Cultural and Social differences Impacts

Many years ago, several studies were conducted to treat cultural diversities, this issue has likewise been talked about in mobile applications selection and how culture affects the user acceptance toward a technology (Gallivan and Srite, 2005). In addition, applications' feature were also influenced by user cultural attributes (Hiller, 2003). The most critical impact is on associations and behavior (Sarala, 2010). First, in various societies individuals have distinctive qualities on the same subject. Besides, traditions likewise prompt various kind of behavior. Finally, state of mind is likewise distinctive in various nations. As an example, because of the social contrasts between the Asian and the European clients, some mobile applications are seen to be more valuable than others. Gaming, for instance some are extremely well known in Japan, South Korea and China, however not in the European nations. Another illustration, is mobile TV, as individuals in Japan and Korea invest hours for driving in day by day bases, thus they are more likely to utilize more diversion apps. Then again, individuals in Finland for instance, they don't invest much energy as much as Asian for driving, in this manner the reception and acceptance of mobile TV has fizzled.

6.5 Mobile Application Platform And Operating Systems

First, a platform in prior researches has been defined as hardware configuration, a working framework, a program system or some other regular substance on which various related segments or applications run (Ballon and Van Heesvelde, 2011). An introduction of mobile application platforms and brief comparing between each mobile service platforms are discussed in this section. In addition, Mobile Operating Systems (OSs) will be explored profoundly and many of the well known OSs will be presented. Because we gain something from using mobile applications, we are always requesting a service. Their significance in telecommunication market turns out to be considerably more obvious on the off chance that we take into account service provisioning instrument. Generally, telecom operators' entry were the unic possible to application access; whereas after the development of App-Stores the conditions have been changed. By 2007, when the principal official application store has been put online, users have more the chance to acquire their needed applications by means of the application store. Apple has been the first to present such store-based service. Application store idea has been adopted by numerous different device providers, for example, Android with Google store , BlackBerry with BlackBerry App World, HTC with Market and Windows with Microsoft store. These open application-stores are known as the platforms which go about as the main in-between the phone and developers of applications. Some of these platforms offer boundless number of uses –like Apple and Android application store, while others have just set number of apps Microsoft store for windows phone. There are currently three major mobile service platform providers in the telecommunication market. These are known as the device-providers, developers and Operating systems platforms, in the following section these three platforms will be discussed in details. Users to acquire the applications are required to get to the application stores by means of their phones or supplier's site page and afterward select the application they need to download. While a portion of the applications are allowed to download, others are costly. That is to say, purchasers are required to pay to have access to download link. Another issue with these platforms is the service improvement procedure and how the application designers can take an important part in the advancement of such process. Among the platforms suppliers set certain constraints for developers' cooperation; the application developers or designers must respect the rules to have authorization to join the service improvement. Apple is for instance one of such platforms. In addition to this, and Google with Android namely Open platforms and application designers have more flexibility keeping in mind their important contribution in developing the platform.

6.5.1 Device providers based stores/platforms

Gadget producers have lately entered to mobile service market like Nokia, Apple (iPhone), HTC. They are putting top of the line application on the client handset, adding to that propelled communication platforms which are recognized as the Device-Centric Platform model. Commonly, gadget makers utilize their special exacting platform to achieve the end-clients. Applications for advanced mobile phones i.e (smart phones) are open by means of application stores which are given by gadget makers' sites or App-stores. These business sector players likewise characterize what particular platform is utilized to give administrations to the clients and they may likewise give devices and assets in the types of a Software Development Kit

(SDK) to service engineers to propose new services. Application designers, with a specific end goal to take an interest in service improvement process need to take after specific principles allotted by gadget makers. Some gadget creators (Apple and BlackBerry have an extremely firm approach to be tracked by engineers, while some other gadget producers have more advantageous controls (HTC and Nokia). Platforms afforded by Apple and BlackBerry are thought to be Closed kind of platform (closed source), while platforms supplied by Android are thought to be (open source). The contrast between the Closed and Open platform is essentially because of approaches and policies that gadget makers or platform suppliers relegated. The platform is called Open type if the chance that the investment of the application designers' and outsiders is allowed to everybody, however; if the interest ought to be trailed by strict tenets, then the platform is considered as Closed.

6.5.2 Application providers based platforms

Furthermore, administrator driven and gadget driven platforms, there is another sort of platform which is provided by exclusive service suppliers. Applications suppliers, for example, Windows, Google, Facebook and Skype are hoping to disorder mobile media transmission market the profile of service oriented model. These business sector players give simple-to-use, frequently valuable, new services and resolutions to consumers. These purported full IP-based organizations and restrictive platforms utilize their own particular platform to offer services to the end-consumers. services are regularly offered taking into account the best exertion service conveyance. This is for the most part because of the way that Internet service contributors don't possess the system infrastructures and they depend on the Internet Protocol (IP). These kinds of services are not commonly skilled enough to guarantee the most astounding Quality of Service (QoS), user security, and privacy. A standout amongst the most utilized and client acknowledged such administrations is Voice over Internet Protocol (VoIP), which is fabricated in view of SIMPLE, (Session Initiation Protocol for Instance Messaging and Presence Leveraging Extensions). SIMPLE is derived from SIP (Session Initiation Protocol). SIP is a procedure that is initially intended for stimulating media sessions, including voice over IP (VoIP).

6.5.3 Operating Systems

A mobile working framework is a part of programming that is mindful to keep running on cell phones, for example, smart phones, tablets, PDA or some other mobile gadgets. A mobile working framework is program which is incorporated on the cell phones to empower different services and applications to use and function. Broadly speaking, device producers have their own exclusive operating frameworks; however there are some device makers where they utilize mobile working systems planned by others on their devices (Samsung, LG). Presently, there are several entrenched working frameworks in mobile communication market that the greater part of device producers are utilizing them. The most prominent mobile working systems are for example Apple operating system is 'iOS', Google with "Android" and Microsoft with 'Windows Mobile'. In any case, there are a few contrasts between these mobile operating systems, while some of them can be kept running in a wide range of brands (devices), others are just intended to be utilized particularly just on one specific brand (device). As an example, Apple iOS and BlackBerry OS are composed just for Apple's and BlackBerry's items and they don't permit these OSs for setting up on non-Apple and non-BlackBerry devices. Not at all like, Apple and BlackBerry, Android from Google can be introduced on other hardware; this infers other device makers which they don't have their own particular restrictive mobile operating systems can exploit them. As indicated by, IDC, the overall mobile OS industry use is ruled by Android with 82.2% of shares; this is trailed by Apple iOS with 13.9% market shares. BlackBerry OS has 0.3% and Microsoft has 2.6% market shares. Other working frameworks have just little market share 0.4%. The following table illustrates the main mobile operating frameworks with their market share in August 2015 .

Period	Android	iOS	Windows Phone	BlackBerry OS	Others
2015Q2	82.8%	13.9%	2.6%	0.3%	0.4%
2014Q2	84.8%	11.6%	2.5%	0.5%	0.7%
2013Q2	79.8%	12.9%	3.4%	2.8%	1.2%
2012Q2	69.3%	16.6%	3.1%	4.9%	6.1%

FIGURE 6.4: The Mobile Service Operating Systems. source: http://www.idc.com/prodserv/smartphone-os-market-share.jsp

6.5.4 The service Data Security, Confidentiality And Quality Of Service

The security and protection of the users' data and information is of conclusive unease toward a large portion of the Internet-base organizations and application suppliers. Generally, services which are offered by Telecom operators are more secure and dependable and ensured service quality. This depends on the way that the operators have vigorous system foundations available to them. Then again, device producers and service suppliers are just able to offer service quality to what is named as 'Best-Effort'. Best effort conveyance depicts a network service in which the supplier does not give any insurances that information is conveyed with a specific need or that a consumer is given an ensured quality of service level.

One of the top priorities for service and application supplier is securing users' data and individual details. Application suppliers, for example, Skype, Yahoo, and Facebook have put intensely in creating controls to ensure their users' data and profile.

Yet, services offered through Internet-based organizations are based with respect to best exertion conveyance which implies it doesn't promise the most noteworthy service quality. Consequently, consumers here and there experience poor voice, video quality or absence of enough accessible system transmission capacity. This can prompt supposition that, Telecom operators can utilize the shortcoming of the Internet-based organizations as allowed to pick up an aggressive situating by offering high service quality, more secure and dependable services.

To summarize, the platform's war in mobile media transmission business sector is pushing mobile operators to reevaluate their business sector part in the event that they need to have a key part in the development of mobile telecom incomes. One conceivable option for them is to concentrate on bit-channel arrangement. This implies that they just give system infrastructures and Internet admission and don't contribute in mobile service provisioning. The other option is that they put resources into new services that can struggle with the IP-based organizations and device makers' platforms. From end-consumers point of view, this war between various players in the media transmission business sector may prompt immense decreases in costs, and may bring about perplexity on where to choose the services. This business sector fracture may bring about developing of different service provisioning models for complex communication services.

It is worth mentioning that operator-centric platform model offers all control and power at the administrator; then again, benefits supplier –centric and device-centric are picking up business sector upper hands by suggesting services that acquire less time to be dispatched in the business sector. Usually, mobile Telecom holds ought to manage numerous issues, for example, institutionalization before propelling their support of the mobile service market.

Chapter 7

AFC to understand travel behavior in the greater Paris

7.1 Overview

This chapter (6) explores my practical work done in LVMT about the flat fare impacts in travel behavior under the transit system in the greater Paris also called Ile-de-France, the ways in which that changes affect the socio-economy, and the additional impacts of transits system changes on users' behavior. Similar topic objectives have been discussed in a series of past reports, including LVMT report published few years ago (Exploring the relationships between work and travel behavior on weekdays - An analysis of the Paris region travel survey over 20 years). However, the nature of transit system changes over time, the structure of the socio-economy and the analysis methods continue to improve. Consequently, the aim of this study differ from those of earlier works, both in aims and results.

This chapter is organized into eight sections.

1. Summary - examines the objectives of understanding the travel behavior and introduces the context of transportation.
2. AFC and other methods - introduces the inter/extern relation between AFC and other methods used to measure socio-economic impacts.
3. Transit system in greater Paris - presents the case of study and provides all the essential information related to the data used.
4. Tap-in Tap-out data from Navigo validation system - describes the data collection process.
5. AFC to estimate trends of traffic state under constant supply - demonstrates the steps followed to understand travel behavior under TS of the greater Paris.
6. AFC to estimate demand response to supply change - discusses the results.
7. Conclusion - a conclusion of chapter (6) findings

7.2 Summary

Today the world is experiencing unsustainable growth in many sector of human life among them mobility and transportation. Due to an increasing population growth, focus should be placed on these two issues. Companies introduce innovative and often costly policies which unfortunately are liable to fail when we do not promote significant updates to the working systems.

7.2.1 The urban transport system

The life of a city depends on its transportation system and a healthy urban economy requires a smooth and efficient transportation. Transportation of both people and goods, is a service without which the city cannot live and exchange. The transportation function plays a key role in each city. It affects directly what constitutes the energy and the main reason of urban forms: the variety and density of exchange opportunities made available for citizens. If the ease and quality of these urban exchanges (work, shopping, leisure, ...) depends on the richness and equilibrium of the urban structure, it is also conditioned by the efficiency of the transportation system.

To analyze the urban TS , we should place it in its context: the city. By analogy with the functioning of the human body, the urban TS can be compared to blood and nervous system that irrigates different places of the city, bring them into relationship. This system is not only characterized by materials for network communication, but also by the mobility made by persons and goods. TS basic function is to exchange.(Bonnafous, 1983) call it "urban mobility system" and they assign the following definition: "It is constituted at once the flow of goods and people through the city and what we call the TS, which constitutes the support" (BONNAFOUS, 1983). The TS is a complex system. Its analysis depends in some basic fundamentals, to conceptualize it. (Manheim, 1979) developed two fundamentals of urban mobility system analysis:

- The main system of mobility inside an urban area should be considered as a unic and multimodal system. That is to say that all modes should be considered and all the elements of the TS must be taken into account (people and goods transported, vehicles involved, network and equipment), etc. Based on this detailed definition of the mobility system, the analyst needs to limit the analysis to those elements that are fully involved with the objectives. This operation leads to considering explicitly the hypothesis introduced by rejecting all the individual elements of a highly complex and interconnected system.
- Mobility system considerations cannot be independent from the social system considerations and the economic policy of the urban area. (Manheim, 1979) and (Bonnafous, 1983) likewise adopt a systemic approach of the mobility system . The urban transport system is highly interconnected to the socio-economic system. The link transport-urban constitutes the characteristics of the urban system. The TS affects the manner of how the socio-economic system evolves and changes, these changes in the socio-economic system will consequently lead to changes in the transportation system. This interrelation is fundamental for Manheim, and allows him to design the mobility system with three basic variables:
    - T, system of transport;

– A, system of activities;
– F, the structure of the flow in the TS, that is to say, the origins, destinations, routes, volumes of transported goods and people moving;

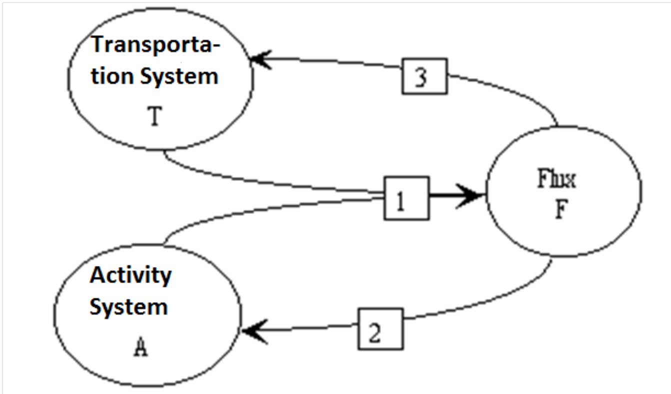


FIGURE 7.1: Manheim identifies three kinds of relationships between these three variables

Relation 1 in Figure 7.1 shows that the structure of flux in the transportation system is determined by both the physical transport system and the activities system. Relation 2 indicates that the structure of the current flux may show effects on the system of activities during the time. The relation 3 indicates that the current structure of flows can also lead to modification in the transportation system (in relation to the current or anticipated flux, entrepreneurs and authorities will develop new transport services or update the existing ones).

Manheim reported that the system of activities denoted by A consists of a subsystems set which are overlapping and interrelated in between them such as social structures, political institutions, housing markets, etc. while the mobility system is also considered as a subsystem of A. The evolution of the system of activity is determined by a large number of efforts and pressures. The internal dynamics of this system is very complex and our understanding of these dynamic remains incomplete. Transport plays a role in the evolution of the system of activities, whereas in specific situations, it is not the unic determinant of this evolution. Thus as Manheim highlighted, the development of vehicles and fast roads did not boost the sub-urbanization and dispersion of metropolitan areas. Indeed, this phenomenon is strongly related to income level, changes in housing markets and jobs, etc. The interrelationship between the mobility system and the system of activities appears fundamental to understand the functioning of such a system. Manheim concludes that transport should be seen as a technology, a system of physical elements managed by human organizations to move people and goods.

This systemic approach of the mobility system is taken over by (Bonnafous, 1983). These authors propose a conceptualization of the city, and developed "the hypothesis that a city can be interpreted as the interweaving of three subsystems, each with its own operating logic and processing, but which articulate with each other in complex and causal relations "(BONNAFOUS, 1983). The three subsystems that these authors consider are: a localization system (which refers to the problem of land use), a mobility system, a system of activities and social relations (which refer to the operating mode of the society). We recover the approach of Manheim, with more explicit content of activities system , which is divided into two subsystems, the first highly refers to the space constraints of the urban system and the second to the temporality of urban activity.

The analysis of the mobility system constitutes on studying the forming of flux, its structure, and its determinants. Systemic analysis of Manheim helps us to conceptualize the forming of movements in urban space. The movements made are resulted from the match between the transmission service (networks, access services) and traffic demand (which depends on the system of activities). This economic equilibrium is an explicit relationship between supply and demand.

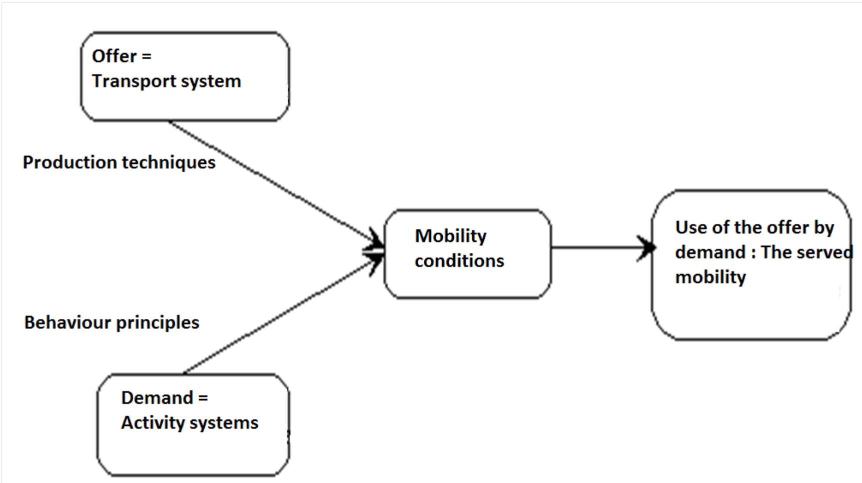


FIGURE 7.2: Representation of the transmission system: the confrontation demand and transport supply

Figure 7.2[1] shows that the urban mobility is the result of the confrontation between an offered service of transport , the transportation system materials and travel demand.

Whereas, the transport service can easily be identified, transport demand, on the other hand should be clarified. The transport demand refers to the choice behavior in the individuals' mobility. It is a representation of individual mobility behavior. Moreover, it is a derived demand in the way that the mobility is not needed for itself but it is a travel instrument driving to a certain locations at a certain time, and that purpose is itself derived from the need to achieve some patterns of activities. Thus, the transport demand is conditioned by the system of activities. The need to travel is resulted from the individuals exchanges needed in the city and the dispersion of business places throughout the city. In this regard, Vermot-Desroches pointed out that transport is only an effect of spatial interaction needs. Travel demand is therefore determined by the individuals' behavior with regard to their activities and thus refers to a group of a complex and spatially varied activities such as work, leisure activity, shopping and residences. Travel demand is a intermediary derived demand.

[1]Source: Ministry of Economy and Finance Equipment, Forecasting Division (1997).

7.2.2 The urban transit system

Transit or public transport is a group of different vehicles used to the simultaneous reception of people. The passenger has access to its services only after buying a ticket, pass etc. The payment depends on the functionality and the applicable regulations (public or private) .

Public transports are often organized by the public authorities (namely STIF in the greater Paris). This is what explains their frequent given name "public transport"; Since the public transport are often organized by the governmental authorities, it would be appropriate to call it "passenger public transport", as it is often the case, or even better "passenger transport organized by the public authorities". The different kinds of the land passenger public transport are listed bellow:

- Passenger Road Transport : bus (for urban services), bus (for interurban services), trolleybuses, traditional or shared taxi, large discount vehicle, small discount vehicle, carpooling
- Passenger Rail Transport: train, TER, RER, subway/metro , trolley, Rubber-tyred metro , tram-train and train tram.
- Other guided transport of persons: monorail, skytrain, funicular, rack railway, cable car

Compared to autonomous vehicles , the public transport are much more efficient than private vehicles, in terms of

- Energy consumption
- Easy mobility : a bus can carry more than 60 people using the same surface area as a couple of cars that are occupied by 2 persons as an average.
- Avoid parking service occupation

However, public transports presents also some drawbacks in terms of :

- Investment costs : public transport require major investments, they are depending on the techniques used and therefore it is justified in very densely populated areas.
- Flexibility : public transports do not offer the same service as autonomous vehicle transport mode: (i) they impose a respect of timetable, (ii) correspondences are often complicated, (iii) they do not provide door-to-door service, (iiii) Going home at lunchtime is rarely possible.

Each of these advantages and inconveniences affects somehow the travel behavior under the TS .

7.2.3 Smart card fare collection systems

Smart card fare collection systems are now implemented all over the world. The concept is well advanced in Europe, especially in France (Marie-Pier Pelletier and Morency, 2009). The smart card fare systems were introduced to facilitate the revenue collection. However, they are also designed to store large amount of data, (Bagchi, 2005). These data have potential applications. The information collected can be used in different ways by transit planners and researchers. Various uses are possible at three levels of management, (Marie-Pier Pelletier and Morency, 2009), namely the strategical (e.g. long-term planning), the tactical (e.g. service adjustments and network development), and the operational (e.g. ridership statistics and performance indicators). Recent studies have demonstrated how much interest exists in using smart card data for transit planning.

Concerning the impact of fare change, from the user standpoint, a large number of studies have analyzed passengers' responses to fare change using stated-preference (SP) surveys, revealed-preference (RP) surveys, or a combination of both (Sharaby, 2012). Passengers' responses to fare change are generally referred to as price elasticity of the demand, represented as a measure that gives the percentage change in quantity demanded in response to a one percent change in price. Price elasticity, on the other hand, is a factor in predicting transit ridership and the change in the global revenue due to fare change (Pelletier, 2010). A literature review indicates that the price elasticity of travel demand varies greatly in terms of variables such as time span, transit mode, original fare level, income level, journey distance, data paradigm, type and direction of price change, demographic and geographic conditions (Farber S, 2014).

In this thesis we interrogate the potential of AFC to capture the mobility behavior of individual passengers, especially so in order to assess the trend of the demand under constant supply and, distinctively, its response to a significant change in transport supply. We study a case of transit fare scheme in greater Paris, in which a number of subscriptions between concentric zones were replaced by a unique, flat fare subscription on September the 1st, 2015. Our approach combines a computer treatment with a geographical and socioeconomic analysis.

7.3 AFC and other methods

One important premise of this work to pursue the proposed goals is that AFC data conveys relevant information about the mobility of people. Among other pieces of such information we have time and position in the gate of validation. Combining this information with the timetables of trains, we can construct a leg data table. A leg is generated systematically using two successive validations made by a user using tap-in and tap-out information at the access station and egress station respectively. Thus we can measure the quantitative and qualitative evolution of user's visits to the stations.

This is a technical-based investigation with (i) a big influence of technology challenges which highly relates to eventual maintenance of the system and (ii) the inability to satisfy the demand in some stations with high number of visitors, that is to say the information is influenced by the delay caused. Nevertheless, using AFC data represents a novel and efficient way if compared to the regular SP/RP surveys [2]. This is specially true if we consider the finance purpose as highlighted elsewhere (WBCSD, 2016), in which authors argue that online process compared to face-to-face surveys

[2]Several questions submitted face-to-face or using online applications.

had the additional advantages of being cheap. This is a similar case compared to AFC. However, in online surveys the conveying mobility information is still not guaranteed due to the penetration rate of such an approach; it is not always guaranteed it will reach all the population. Additionally, time constraints and privacy issues are also very important concerns, needing appropriate treatment.

The data collection methods used in the submission of a survey includes several questions which can lead to the desired information through the analysis of people feedback . In their research (), authors tried to understand how the activities of people are conducted before, during and after their journeys so as to have a perspective about it. Such a perspective involves analyzing the activity of the subject on several occasions over time which consists in submitting the same survey each time, which would demand additional funding.

As for AFC, the information given from tap-in and tap-out can be used for the same purpose which is to understand the mobility of people and will be substantiated by the subsequent details.

7.4 Tools and Materials

The practical work in hand has been conducted during a period of four months and the results have been presented to the STIF (See Giraldo, 2016). To explore the data, many programming languages and software environment for statistical computing and graphics have been used. Each of these has been particularly efficient for a specific task. For instance, R language was very useful for a Big data analysis (+/- 6 millions of entries-journeys). The tool provided a capacity of running several algorithms on the data, while this also depends on the computing power of the machine. Three major R packages have been used:

- data.table: this package contains several useful functions among of it, we have "fread" which avoids the system breaks when the data are bigger compared to "read.table" function;
- ggplot2 : this package contains several type of graphical representation alternatives. Adding to this the good design compared to the R plot package;
- sqldf : by using this package, I was able to lunch SQL queries in R. The package was efficient to run algorithms that are quickly dealt in sql then in R;

In addition to R langage, Excel and SPSS have also been used however, they are only suitable for small data analysis. The algorithms used and applied to the data has each took a long time to provide results (max. 11hours). However, by the means of optimization process the hole data could be analyzed.

7.5 Transit system in greater Paris

Before we can further into detailing the proposed approach, it is important to understand briefly how the transportation system in greater Paris is structured, and how its Navigo system works.

7.5.1 Navigo system: AFC data of urban rail in the greater Paris

The two main systems of urban rail in the greater Paris metropolitan area according to (Aguiléra, 2014) are: (i) the semi-closed Metro system which includes 14 lines. This system is equipped with tap-in gates at access only; (ii) the closed RER (*Réseau Express Regional*, which is the Regional Express Network) equipped with both tap-in and tap-out gates for external transfers with 5 RER lines. The AFC system of this area implemented by STIF , namely "SIDV", allows the usage of the smart card through the network and stores anonymous passenger information including the smartcard number which is anonymized (with anonymous number that is maintained during 3-month periods), the date, the validation instants at tap-in / tap-out gates, the gate IDs and name of access/egress stations. During the peak period on workdays, more than 90% of the trips by public transport are home-work or home-study trips using network subscription hence the smartcard. Although, the information of validation data is collected systematically, it remains incomplete due to the absence of validations in the exit, except in some cases such as the RER (for instance the RER A) and trains. Thus we focus on the RER A during this study.

7.5.2 Line RER A and related data set

The line A of the RER network in the greater Paris, often simply called RER A line (depicted in Figure 7.3 is one of the world's busiest lines, and the busiest line in Europe with around one million passengers each day. It contains 46 stations in total 109 km and is structured around a central trunk into which five branches are grafted: two eastward branches to northeast terminal Marne-la-Vallée and southeast terminal Boissy, and three westward branches to northwest terminal Cergy, central-west terminal Poissy and southwest terminal Saint-Germain, respectively. The central trunk between Vincennes and La Défense stations passes through the largest underground train hub in France, Châtelet-Les Halles in the center of the city, and serves the major business district La Défense. It serves 2 million jobs, representing 41% of regional tissue.

7.5.3 Navigo Pass: Re-zoning of greater Paris and the new fixed-price scale scheme

The greater Paris covers an area of 12.012 km2 and accommodates a population of 11.5 million (EGT, 2012) as well as 5.6 million jobs representing 29% from the GDP of France. It is characterized by a dense transit system and a great diversity of urban transit modes and services, including: 16 metro lines of over 200 km; 1.500 bus lines of over 25,000 km. The network is composed of 14 train lines of over 1.500 km; 8 tramway lines. In terms of daily mobility, it counts over 41 million daily journeys, 39% of these journeys are made by foot, 38% by car and 20% by public transport. As we

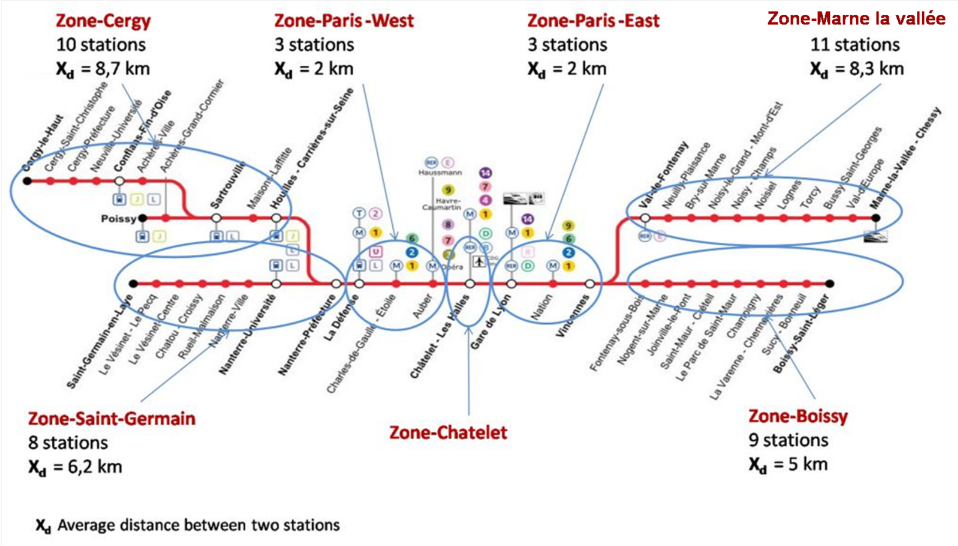


FIGURE 7.3: Line A of RER in the greater Paris, including correspondences (RER, Transilien, metro and tram).

move away from the city center, the private car is over-represented compare to the other modes with 2/3 of daily journeys in outer suburbs. Public transport counts 8.3 million journeys, 51% of them are made between home and work or study place (EGT, 2012).

Since 1975, transit pricing of Paris region was based on a system of concentric zones. The region was divided into homocentric areas. The journey price was depending on the number of zones that the passenger crossed. In September 2015, STIF introduced a flat fare system with the creation of a single price for the Navigo Pass.The zones 1-2, 1-3, 1-4, 2-4, 2-5 and 3-5, became Pass "all zones" (1-5) , that is to say the ability to travel throughout all the region. The monthly price was set to 70 euros while the maximum was 113 euros (zone 1-5) and the minimum was 67 euros (zone 1-2) before the changes. Navigo system offers two types of membership, both weekly and monthly:

- Monthly subscription is valid for one month (from the first to the last day of the month)
- Weekly subscription is valid for one week (Monday to Sunday)
- The total number of weekly and monthly subscriptions is around 1.5 million

7.6 Tap-in Tap-out data from Navigo validation system

Navigo Pass allows the access to a rail stations' gates by passing a card, which contains user's identity data, near an electronic reader tap-in/tap-out. These passes are used to access vehicles of the RATP, the SNCF (within the Transilien network), and companies under the aegis of the Syndicate of transports in le-de-France (STIF) (Wikipedia).

In the analysis below, AFC data spanning the working days of one week (i.e. weekends excluded), from three different years, are considered: Monday-Friday from October 2013, 2014 and 2015. This data gathers validations: information recorded by AFC system aforementioned above and legs which are systematically generated from the validations data using the tap-in and tap-out details. The validation data is exploited using a specific dynamic O-D (Origin-Destination) matrix inference scheme devised by (Xie, 2015). This scheme extends previous work by (Zhao, 2013) to infer rail transit O-D matrix and to generate leg choice from Oyster smartcard data in London, based on data processing and analysis methods supported by technologies of database management systems (DBMS) and geographic information systems (GIS).Finally, the legs data table now contains the date, the validation instants at tap-in & tap-out gates, the gate IDs, the access & egress station name , duration & distance of the O-D and the carrier information : bus or rail system (Metro, RER) including the lines i.e RER A, B etc.

As for validation and leg, journey data table of all the network was included in our raw data. The method of SIDV for journey construction is to extract all O-D pairs related to each individual from leg data table. One journey contains all the information about the first and last validation of each individual after the conclusion of several legs. One individual can have multiple legs and journeys.

For the current study, a specific scheme has been pursued for the data training which involves five steps as following: (1) extracting the data related to RER A from the dataset of the entire network Table 1; (2) creation of the average day-matrix for each year data to generate average day-matrix of 2015-hat; (3) focus on the rush hour of the morning (RHM) from 4:30 to 11 am according to a survey data made in France called Enquete Globale de Transport (EGT), a questionnaire submitted every 10 years in France in which citizens such as employees and students are asked about their daily transport activities including the time of their going out for work (see Figure 2); (4) generating an O-D matrix which contains the number of legs for each O-D using the leg data table for both 2015 and 2015-hat; and finally, (5) discuss the traveled distance.

As aforementioned, the AFC data assembles anonymous passenger information. Dealing with anonymous users cannot allow the process of Panel methodology for specially, a qualitative analysis of the public transport users. The methodology of the panel can be applied with different techniques for mobility analysis such as:

- The basic technique used consists on doing a personal and well-structured interviews: face-to-face or by phone. The advantage of this methodology is to avoid the drawbacks related to each technique and provide a diachronic approach of the traveler to understand how the journey is conducted.
- More advanced technique aims to use AFC data, in case that the anonymizing process has not been applied to the passenger cards, which may assist in obtaining both qualitative and quantitative data. In our case of study, the profession was missing in the analysis for more accuracy in term of the third step in the scheme aforementioned above.By applying step (3) we try to focus on the employees and students. This

TABLE 7.1: NUMBER AFTER THE EXTRACTIONS

Year	Table	Monday	Tuesday	Wednesday	Thursday	Friday
2013	Validation	1273964	1317334	1238193	1353485	1304540
	Leg	940730	979194	924376	1001201	954489
	Journey	940534	978993	924199	1000986	954322
2014	Validation	1261058	1303917	1251918	1270060	1273338
	Leg	929489	962662	923193	927330	924776
	Journey	929449	962656	923186	927326	924774
2015	Validation	1353514	1391997	1374069	1413155	1205928
	Leg	1013397	1057082	1031331	1061110	925967
	Journey	1013297	1057082	1031331	1061109	925966

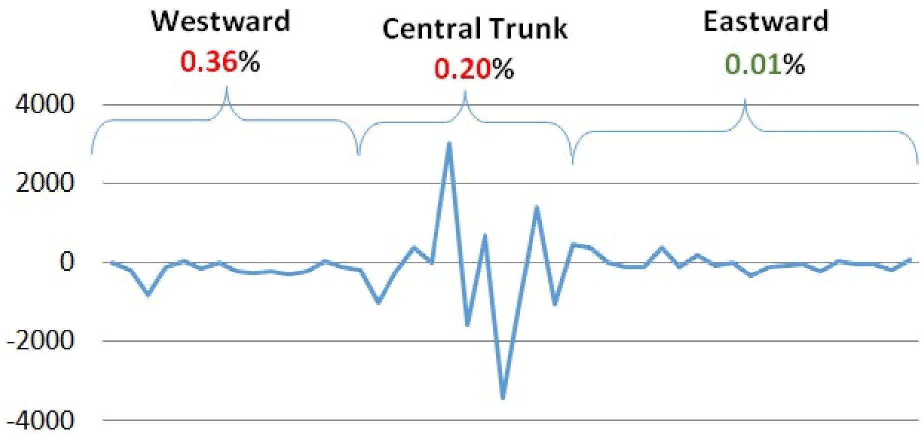


FIGURE 7.4: Evolution of the total number of legs between 2013 and 2014 For each stations

step remains imprecise because RHM sample does not totally exclude other types of users (non-student and non-employee).

7.7 AFC to estimate trends of traffic state under constant supply

The extraction of data related to RER A from the dataset of the entire network was based on the validation information i.e. the data that belong to any validation in RER A, this was done using carrier information and station name as a conditional variables. Then, from each validation we take two values to produce RER A legs extraction: User ID and leg ID which is a foreign key that relates each leg with its two validations, thus these two validations have the same user ID and leg ID .The total number of validations is high compared to the number of legs. Each journey is also matched with all the legs that are related to it using journey ID that exists in both leg and journey of the same user ID. As results, we have RER A data tables of validations, legs and journeys.

The generation of the average day-data was done using several matrices. First, a matrix O-D for each day of the week starting from Wednesday and ending to Friday excluding Monday and Tuesday. The reason of this exclusion is due to habitually problems in the urban rail transit of greater Paris. The problem was detected after analyzing all the days-matrix, a high decrease was observed in the number of legs with origin/destination from several stations of the central trunk of RER A in 2015. The average day-matrix is the mean between all the days-matrix. After computing the average day-matrix for both year 2013 and 2014, the evolution shows a 0.55% decrease in the total number of legs between the two years respectively. From 2013 to 2014, the evolution 7.4 shows a high decrease of the legs number related to station Charles de Gaulle Etoile,Châtelet-Les Halles and Gare de lyon with 0.20% of decrease in all central zone. In the westward, we observe a decrease of 0.36% while we note 0.01% increase in the eastward.

According to the evolution between 2013 and 2014, we generated a new data denoted by 2015-hat. For instance, if the number of legs between station A and station B has increased from 2013 to 2014 with x%, this means that the number of legs between A and B will increase with x% between 2014 and 2015-hat.To generate the leg data table of an average day 2015-hat, we subtract the average day-matrix 2014 denoted by X from average day-matrix 2013 denoted by Y to compute X-Y denoted by Z. The matrix Z is used as input by an algorithm-based method, this method consists on testing the value of each O-D in the matrix Z and adding a specific legs from Thursday's leg data table if the value is positive ; or deleting a specific legs from the same data table if the value is negative i.e. let's suppose that the number of legs of Chatelet-Cergy as an O-D is equal to p^+ in the average day-matrix and the number of legs of the same O-D is equal to p^- in Thursday-matrix. Thursday 2014 was chosen as reference day-data to generate the 2015-hat in the following function : sequence(p^+-p^-, length(legs as Chatelet-Cergy))

```
Result: Average day-data of 2015-hat is the new T
T = Thursday-data 2014;
1: function SEQUENCE(x,y)
   i = ⌊y/x⌋;
     if x > 0 then
       while i ≤ length(T) do
           p = select i^th leg from T;
           i = (i + ⌊y/x⌋) + 1;
           Insert p into T;
       end
     else
       while i ≤ length(T) do
           p = select i^th leg from T;
           i = (i + ⌊y/x⌋) + 1;
           Delete p from T;
       end
     end
2: end function
```

Algorithm 1: BODY OF THE FUNCTION SEQUENCE

TABLE 7.2: THE NUMBER OF LEGS IN THE AVERAGE DAY

	2013	2014	2015-hat	2015
Legs	957013	928429	919762	998788

7.8 AFC to estimate demand response to supply change

The average day-data contains the legs from the whole day Table 7.2. We determined a strategy to focus only in the rush hour of the morning Table 7.3. This decision aims to select only the legs related to the users that use the public transport to go for work or study. We assume that the majority of those selected users are actually citizens of the greater Paris.

TABLE 7.3: Number of legs in RHM

	2015-hat	2015
Average day	919762	998788
RHM	376890	403428

The data included some incoherent legs and thus some data cleaning is required. An incoherent leg contains an impossible O-D using only one mode of transport which means that the user has missed a validation somewhere in the exit of the first mode.

Let us suppose the following two successive validations of a specific user: the tap-in was made at (RER B) Antony and the tap-out at (RER A) Auber. However, going from Antony to Auber we normally need to make a transit in Chatelet Les halles.

The correct form of this leg is retraced as follows:

- Leg 1: (RER B) Antony to (RER B) Chatelet Les halles
- Leg 2: (RER A) Chatelet Les halles to (RER A) Auber

The reconstitution process has been dedicated to all lines RER B, RER D and RER E which have Chatelet Les halles, Gare de Lyon and Val de Fontenay as common stations with RER A respectively. The data contains some incoherent legs, including Metro and other modes which were not taken into account in the process of the reconstitution. The goal behind this correction/reconstitution is to give much efficacy to our results by recovering lost legs and it is also considered as a big step of data cleaning. Table 7.4 represents the number of reconstituted legs in each average day.

TABLE 7.4: THE NUMBER OF RECONSTITUTED LEGS

	RER B	RER D	RER E
2015-hat	116536 (12,7%)	97048 (10,6%)	39505 (4,3%)
2015	129703 (13%)	91384 (9,1%)	50272 (5%)

According to 2014 (the last year of the concentric fare transit system) and 2015 data, we note that the number of journeys between the center of Paris and suburbs are the one that increased the most by 2.8%.

Figure 7.5 represents a matrix of O-D between the RER A defined zones, with the indication of the decrease and the increase of the flux between each zone in the 2015 and 2015-hat , the increase goes to red and the decrease goes to blue.The values corresponds to the total number of legs in 2015 divided by 1000. The first effect noted is the increase of O-D in the eastward region, particularly northeast branch Marne-la-Vallée.

	Zone-Cergy	Zone-Saint Germain	Zone-Paris-West	Zone-Chatelet	Zone-Paris-East	Zone-Marne la vallée	Zone-Boissy	Other destination	Total
Zone-Cergy	4,438	1,652	16,521	1,74	2,547	1,179	0,097	17,05	45,224
Zone-Saint Germain	0,109	2,603	10,785	1,377	1,887	0,384	0,117	4,042	21,304
Zone-Paris-West	1,325	4,664	20,069	5,288	5,619	4,276	0,986	48,842	91,069
Zone-Chatelet	0,8	1,167	6,011	1,203	3,223	2,376	1,297	17,925	34,002
Zone-Paris-East	0,74	2,388	13,331	5,338	8,903	8,353	2,85	60,901	102,804
Zone-Marne la vallée	0,212	0,545	10,886	3,819	14,992	10,369	0,549	8,864	50,236
Zone-Boissy	0,101	0,558	8,618	2,502	9,256	0,753	3,872	5,123	30,783
Other origin	4,477	3,439	18,078	5,77	8,237	6,217	1,574	0	47,792
Total	12,202	17,016	104,299	27,037	54,664	33,907	11,342	162,747	423,214

FIGURE 7.5: Matrix Origin-Destination of the RER A (2015)

The graphs in Figure 7.6 represent the density of distances .The long distances (more than 20km) are the distances that have increased the most in the northern branches of the RER A Marne-la-Vallée and Cergy. While in the other side, the short distances (less than 10km) have decreased. In addition to this , the mean duration spent in the RER A has increased by two minutes in the suburbs and remained constant in Paris central area (Paris-west , Chatelet and Paris East).

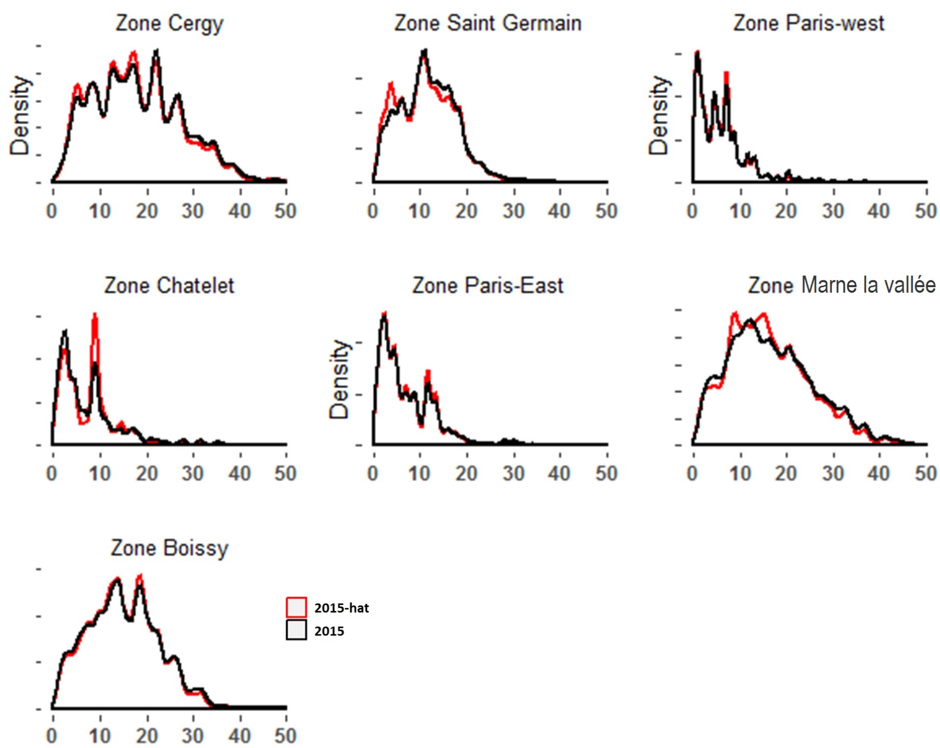


FIGURE 7.6: The effects on the distances involved on the RER A

The weighted distance (distance between stations multiplied by the number of O-D of each zone) is an indicator through which we note that the users of Marne-la-Vallée branch make the longest traveled distance.

After the flat fare application, the results show that the number of legs with long distances in the RER A increased. The same trend is observed in the evolution of the duration, specially for users who spend more than 45 minutes in RER A. According to the O-D matrix of the RER A stations, we can see that the number of legs increased in the east side of Paris central area as well as in the eastward and northwest zone (Cergy). A decrease in the number of legs is observed on the west side of Paris central area and some stations in the southwest. This study was made few months after the fare reform; the users behavior vis-a-vis public transport takes time to change.

The frequency of people entering RER A stations (origin) shows an increase of 7.54% in Paris central area in 2015, which is the opposite of the eastward and the westward which has a decrease of 2.88% and 5.64% respectively. The frequency of leaving stations (destination) shows an increase of 11.49% and 0.48% in Paris central area while in the westward the frequency decreased by 1.00%.

Conclusion

Many technologies were developed to solve a particular problems, however according to several studies, these technologies can solve several others problems that were not designed for. The thesis topic is a good example that explains how do one technology integrates several areas at the same time. Understanding mobility behavior using traditional observation-based technologies is good enough to understand mobility and it was the main method used for several decades however their limitations are really significant specially with respect to the data coherence level that it may provide.The results of such techniques were complemented with the data collected using traditional surveys. These surveys being conducted face-to-face or using others means of communication, they have shown a really efficiency for a global understanding of mobility, these data are not highly accurate due to human brain abilities since the interviewed persons may forget important details about their mobility.

The transportation researchers have waited until the technology introduces Automated Fare Collection systems to the market. When this technology will be used by all the transit users as it will be the case in Paris, there is no doubt that it will promote a real revolution in the current transit systems design. Due to the accuracy of data provided by these systems, we are able to find solutions in short time compared to other techniques which take a long time to be implemented and concluded. Even if they record real time mobility data, the analysis in their side are not conducted directly. The main problem of these systems is their scope limitations. They only provide data related to transit practices and not the hole urban mobility information.

Moreover, the data analysis tools has also been subject of adaptation.The need of analyzing Big data obliged the researchers to find quick solution to both read and analyze this kind data. For example, the well known tool for statistical data analysis namely SPSS have been surpassed by the advance in the data quality and quantity. At the same time, these advances in data collection also promoted the development of new and the the updating of others algorithms and techniques in the data analysis area of study such as data mining and agents technologies adding to these other tasks e.g. learning machine.

In addition, smartphone in its side, is considered as the new trend of the topic. It combines traditional techniques with the innovative ones, using smartphone application, we can track the mobility by the mean of the sensors available in it such as mention, location and ambient sensors, which can provide real-time data and than display the analysis results shared between the user and the application owner.The current problematic of this technique is the user acceptance toward the application. Many factors are taking into account for this purpose where the most important is the privacy concerns. However, the acceptance theories have shown that the user is comfortable with a service when the later provide in one hand a high security level for data protection and in the other hand a useful service. Just recently, several studies have added a new factor to a lot of others which is the amusing design of the service. This explains the use of game aesthetics in several new applications. According to their fast development, We believe that smartphones' applications will be the leaders of mobility behavior data collection systems.

Corresponding author —
Yassine Baghoussi linkedIn
Email : Baghoussi.yassine@gmail.com
Tel : +33(0)768771509

Bibliography

A. Chakirov, A. Erath (2011). "Use of public transport smart card fare payment data for travel behaviour analysis in Singapore". In: *16th International Conference of Hong Kong Society for Transportation Studies, December, 2011*.

A. J. Eronen V. T. Peltonen, J. T. Tuomi A. P. Klapuri S. Fagerlund T. Sorsa G. Lorho and J. Huopaniemi (2006). "Audio-based context recognition". In: *IEEE Transactions on Au- dio, Speech, and Language Processing*, 321–329.

A. Mesaros T. Heittola, A. Eronen and T. Virtanen (2010). "Acoustic event detection in real-life recordings". In: *in 18th European Signal Processing Conference*, 1267–1271.

Aguiléra V, Allio S Benezech V Combes F Milion C. (2014). "Using cell phone data to measure quality of service and passenger flows of Paris transit system. Transp. Res. Part C Emerg. Technol." In: *CIRRELT-2009-46*, 198—211.

Ajzen, Icek (1991). "The Theory of Planned Behavior". In: *Organizational behavior and human decision processes*, pp. 179–211.

Akram Bayat Marc Pomplun, Duc A. Tran (2014). "A Study on Human Activity Recognition Using Accelerometer Data from Smartphones". In: *The 11th International Conference on Mobile Systems and Pervasive Computing (MobiSPC-2014)*, p. 605.

Amblard F., Phan D. (2006). "Modelisation et simulation multi agents applications pour les Sciences de lHomme et de la Societe". In: p. 751.

Amouroux, E. (2009). "GAMA an environment for implementing and running spatially explicit multi agent simulations." In: *In Agent computing and multi agent systems Springer*, 359–371.

— (2011). "KIMONO une methode de modelisation descriptive centree agent pour lexplication des systemes complexes une application en epidemiologie." In:

Anand, S.S. and A.G. büchner (1998). "Decision Support Using Data Mining". In: *Financial Times Pitman Publishers, London, UK.*

Antonio J. Ruiz-Ruiz Oscar Canovas, Ruben A. Rubio and Pedro E. Lopez de Teruel (2012). "Using SIFT and WiFi signals to provide location-based services for smartphones". In: *Department of Computer Engineering, University of Murcia, Spain*.

Armendariz, M. (2010). "Carpooling: A multi agent simulation in netlogo." In:

Atizaz Ali Jooyoung Kim, Seungjae Lee (2016). "Travel behavior analysis using smart card data". In: *KSCE Journal of Civil Engineering*, 1532–1539.

Axelrod, R.M. (1997). "The complexity of cooperation: Agent based models of competition and collaboration". In:

Bagchi M., White P.R. (2005). "The potential of public transport smart card data." In: *Transport Policy 12 pp.*

Balmer, M. (2009). "MATSimT Architecture and simulation times". In: *Multi agent systems for traffic and transportation engineering*, 57–78.

Barry (2002). "Origin and Destination Estimation in New York City with Automated Fare System Data." In:

Bazzan, A. L. C. (2005). "A distributed approach for coordination of traffic signal agents". In: *Auton. Agents Multi-Agent Syst*, 131–164.

Bekhor, S. (2010). "Integration of activity based with agent based models: an example from the tel aviv model and MATSim". In:

Bel, C. and W. Van Stokkum (1999). "A model for distributed multi-agent traffic control". In: *Multiple Approaches Intell. Syst.* 480–489.

Bernstein, D. and A. Kornhauser (1996). "An Introduction to Map Matching for Personal Navigation Assistants". In: *Technical report, New Jersey TIDE Center*.

Bertalanffy, L.V. (1986). "General System Theory: Foundations Development Applications Revised edition." In:

Bertalanffy L. von, Paulre B. (1993). "Theorie generale des systeme". In:

Bierlaire M., Frejinger E. (2008). "Route choice modeling with network-free data". In: *Transportation Research Part C: Emerging Technologies 16*, 187–198.

Bierlaire, M. (2003). "BIOGEME: a free package for the estimation of discrete choice models". In: *3rd Swiss Transportation Research Conference*.

Blythe, P. (1998). "Integrated Ticketing-Smart Cards in Transport." In: *IEE Colloquium - Using IT S in Public Transport and Emergency Services, Paper no. 4, 20 p.*

— (2004). "Improving public transport ticketing through smart cards." In: *Proceedings of the Institute of Civil Engineers, Municipal Engineer 157*.

Bohte W., Maat K. (2009). "Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands". In: *Transportation Research Part C: Emerging Technologies 17*, 285–297.

Bonabeau, E. (2002). "Agent-based modeling: methods and techniques for simulating human systems". In: *Proc. Natl. Acad. Sci. U.S.A. 99 (Suppl. 3)*, 7280–7287.

Bonnafous A., Puel H. (1983). "Physionomies de la ville". In: *Les éditions ouvrières, Série « Initiation Economique »*.

BONNAFOUS A., PUEL H. (1983). "Physionomies de la ville". In: *Les éditions ouvrières, Série « Initiation Economique »*, p. 44.

Bousquet, F. (1998). "Cormas: Common pool resources and multi agent systems". In:

Bouwman H., López Nicolás C. Molina Castillo F. J. Van Hat tum P. (2012). "Consumer lifestyles: alternative adoption patterns for advanced mobile services ." In: *International Journal of Mobile Communications*, pp. 169–189.

Boyle, Daniel k. (1998). "Passenger Counting Technologies And Procedures". In: *Transportation research board executive committee 1998*.

C. E. White, D. Bernstein and A. L. Kornhauser (2000). "Some map matching algorithms for personal navigation assistants". In: *Transportation Research Part C: Emerging Technologies*.

C. Morency M. Trpanier, B. Agard (2007). "Measuring transit use variability with smart-card data". In: *Transp. Policy 14 (3)*, 193–203.

Carlsson C., Carlsson J . Hyvonen K . Puhakainen J . Walden P. (2006). "Adoption of Mobile Devices/Services – Searching for Answers with the UTAUT". In:

Carroll, A. and G. Heiser (2010). "An analysis of power consumption in a smartphone". In: *in Proceedings of the 2010 USENIX annual technical conference*.

CERTU, centre d'Études sur les réseaux les transports l'urbanisme et les constructions publiques (2004a). "Les enquêtes déplacements : méthode pour les grandes villes, enquête pour les villes moyennes et enquête grands territoires". In:

— (2004b). "Étude méthodologique sur la connaissance des déplacements des périurbains". In:

— (2008). "Recueils de données «standard Certu» Les enquêtes ménages déplacements (EMD), ce qui change". In: *Journées d'échanges sur la mobilité urbaine 24 et 25 juin 2008*.

— (2010). "Les enquêtes déplacements par téléphone : mode d'emploie". In: *Mobilités et transports*.

CGDD, COMMISSARIAT GÉNÉRAL AU DÉVELOPPEMENT DURABLE (2010). "La mobilité des Français Panorama issu de l'enquête nationale transports et déplacements 2008". In: *La Revue du CGDD*.

Chapleau R., K.K. Chu. (2007). "Modeling transit travel patterns from location – stamped smart card data using a disaggregate approach," in: *11th World Conference on Transportation Research CDROM, Berkeley, California.*

Chen, H. B. and M. Bell (2002). "Instrumented city database analysts using multi-agents". In: *Transp. Res. Part C: Emerging Technol.* 419–432.

Chen, Jingmin (2013). "Modeling Route Choice Behavior Using Smartphone Data". In: *THÈSE 5649 (2013)*.

Chiang, Hsin-Hao (2013). "Magnetic Field Feature Analysis of Smartphone Application Activities Using Android MI Sensors". In: *Master of Science in Electrical Engineering*.

Chung E.-H., Shalaby A. (2005). "A trip reconstruction tool for GPS-based personal travel surveys". In: *Transportation Planning and Technology 28*, 381–401.

COST, European Cooperation in Science & Technology (2010). "COST 358 Pedestrians' Quality Needs". In: *PQN Final Report*.

Cubillos C., Crawford B. Rodríguez N. (2007). "MADARP: A Distributed Agent-Based System for OnLine DARP." In:

D. Guo Z. Li, J. Song and Y. Zhang (2003). "A study on the framework of urban traffic control system". In: *Proc. 6th World Congr. Intell. Transp. Syst. (Cat. No. 03TH8708)*, 842–846.

DAES, Département des affaires économiques et sociale (2010). "Guide pratique pour la conception d'enquêtes sur les ménages". In:

Daude, E. (2006). "Du complexe a la complexite". In: *Demain la geographie*, 261–267.

Dauphine, A. (2003). "Les theories de la complexite chez les geographes". In:

Davis F. D., Bagozzi R. P. and P. R. Warshaw (1989). "Consumer adoption of mobile TV: Examining psychological flow and media content." In: *Management science*, pp. 982–1003.

Davis, F . (1989). "Perceived u sefulness, perceived ease of use and user acceptance of information technology". In: *MIS Quarterly*, 319–340.

De Rosnay, J. (1997). "Le macroscope Psychol Bull". In:

Ding F., Zhang J. Wang J. (2010). "Accelerometer based transportation mode recognition on mobile phones". In: *2010 Asia-Pacific Conference on Wearable Computing Systems Ieee*, 47–50.

Donnadieu, G. et al. (2003). "L Approche systemique de quoi sagit il. Union Europeenne de Systemique." In:

Dupuy, J.P. (1994). "Aux origines des sciences cognitives." In:

E. Miluzzo N. D. Lane, K. Fodor R. Peterson H. Lu M. Mu-solesi S. B. Eisenman X. Zheng and A. Campbell (2008). "Sensing meets mobile social networks: the design, implementation and evaluation of the cenceMe application". In: *in Proceedings of the 6th ACM conference on Embedded network sensor systems*, 337–350.

Ebendt R., Sohr A. Tcheumadjeu L. C. T. Wagner P. (2010). "Utilizing historical and current travel times based on floating car data for management of an express truck fleet". In: *5th International Scientific Conference: Theoretical and Practical Issues in Transport*.

Edmonds B., Moss S. (2005). "From KISS to KIDS an anti simplistic modelling approach". In: *Springer*.

EGT (2012). "La mobilité en ile-de-France". In:

FangEmail, Xing and Justin Zhan (2015). "Sentiment analysis using product review data". In: *Journal of Big Data*.

Farber S, al. (2014). "Assesing social equity in distance based transit fares using a model of travel behavior". In: *Transport Research*, pp. 291–303.

Ferber J., Perrot J.F. (1995). "Les systemes multi agents vers une intelligence collective InterEditions Paris". In:

Fishbein, M. and I. Ajzen (1975). "Beliefs, Attitude, Intention and Behaviour: An Introduction to Theory and Research, Addison - Wesly, Reading, MA." In:

Fougeres, A.J. (2000). "Un systeme tutoriel intelligent pour la formation des regulateurs de trafic dune compagnie de transport en commun. In Colloque sur la multimodalite IMAG Grenoble." In:

Fourie, P.J. (2010). "Agent-Based transport simulation versus equilibrium assignment for private vehicle traffic in Gauteng." In:

Gallivan, M. and M. Srite (2005). "Information technology and culture: Identifying fragmentary and holistic pers pective of culture." In: *Information and Organization*, pp. 295–338.

Gao W., Balmer M. Miller E.J. (2010). "Comparison of MATSim and EMME on greater Toronto and Hamilton area network, Canada. Transportation Research Record: Journal of the Transportation Research Board". In: *Journal of the Transportation Research Board*, 118–128.

Gardner, M. (1970). "Mathematical games The fantastic combinations of John Conway s new game life". In: *Scientific American*, 120–123.

Gechter F., Simonin O. (2005). "Conception de SMA Réactifs pour la Resolution de Problemes: ne approche basee sur l environnement." In:

Gershenson, C. (2005). "Self-Organizing Traffic Lights." In:

Giraldo, Lia Patricia (2016). "Le passage au tarif unique en Ile de France (Focus sur le RER A) - LVMT". In: *Annual Conference of Chair STIF in LVMT - Ecole des ponts et chaussées (Paris)*.

Goldsmith, A. (2005). "Wireless Communications." In: *Cambridge University Press*.

Greenfeld, J. S. (2002). “Matching GPS observations to locations on a digital map”. In: *Proceedings of the 81th Annual Meeting of the Transportation Research Board*.

H. S. Zhang Y. Zhang, Z. H. Li and D. C. Hu (2004). “Spatial-temporal traffic data analysis based on global data management using MAS”. In: *IEEE Trans. Intell. Transp. Syst.* 267–275.

Ha I., Yoon Y. and M. Choi (2007). “Determinants of adoption of mobile games under mobile broadband wireless access environment”. In: *Information and Management*, 276–286.

Halteren, A.T. van and P. Pawar (2006). “Mobile Service Platform: A Middleware for Nomadic Mobile Service Provisioning.” In: *Information and Management*.

Hamilton, Paul and Suresh Sankaranarayanan (2013). “Intelligent Agent Based RFID System for on Demand Bus Scheduling and Ticketing”. In: *International Journal of Future Computer and Communication, Vol. 2, No. 5, October 2013*.

Hickman, M. (2002). “Robust passenger itinerary planning using transit AVL data.” In: *In Proceedings of the IEEE 5th Conference on Intelligent Transportation Systems*, 840–845.

Hiebeler, D. (1994). “The swarm simulation system and individual based modeling.” In:

Hiller, M. (2003). “The role of cultural context in multilingual Website us ability Electronic Commerce Research and Applications”. In: pp. 2–14.

Hoffman M., Wilson S. White P. (2009). “Automated Identification of Linked Trips at Trip Level Using Electronic Fare Collection Data.” In: *88 th Annual Meeting of the Transportation Research Board CDROM*, p. 18.

Hong S. J., Tam K. Y. and J. Kim (2006). “Mobile data service fuels the desire for uniqueness.” In: *Communications of the ACM*, pp. 89–94.

Hong S. J., Thong J. Y. L. and K. Y. Tam (2006). “Understanding continued information technology usage behavior: A comparison of three models in the co text of mobile internet.” In: *Decision Support Systems*, pp. 1819–1834.

Hugh Louch Brad Davis, Kim Voros Kristen O’Toole and Alta Planning + Design Sam Piper (2016). “Innovation in Bicycle and Pedestrian Counts”. In: *A Review of Emerging Technology*.

I.Gokasar K.Simsek, K.Ozbay (2015). “Using big data of automated fare collection system for analysis and improvement of BRT - bus rapid transit line in istanbul”. In: *TRB 2015 Annual Meeting*.

J. Chen A. H. Kam, J. Zhang N. Liu and L. Shue (2005). “Bathroom activity monitoring based on sound”. In: *in In Third International Conference on Pervasive Computing*, 47–61.

J. M. Zhao, S. Bukkapatnam and M. M. Dessouky (2003). “Distributed architecture for real-time coordination of bus holding in transit networks”. In: *IEEE Trans. Intell. Transp. Syst*, 43–51.

Jung Y., Perez Mira B. and S. Wiley P atton (2009). “Consumer adoption of mobile TV: Examining psychological flow and media content.” In: *Computers in human behavior*, pp. 123–129.

K. Muralidharan A. J. Khan, A. Misra R. K. Balan and S. Agarwal (2014). “Barometric phone sensors: More hype than hope!” In: *in HotMobile*.

Karim Ismail Tarek Sayed, Nicolas Saunier (2009). “Automated collection of pedestrian data using computer vision techniques”. In:

Kim, J. (1996). "Node Based Map-Matching Algorithm for car navigation system". In: *Proceedings of the International Symposium on Automotive Technology, Automation*.

Kulviwat S., Bruner G.C. Kumar A. Nasco S.A. and T. Clark (2007). "Toward a unified theory of consumer acce ptance technology". In: *Psychology and Marketing*, 1059–1084.

Kwapisz J., Weiss G. Moore S. (2010). "Activity recognition using cell phone accelerometers". In: *Proceedings of the Fourth International Workshop on Knowledge Discovery from Sensor Data*, 10–18.

L. Sun D.H. Lee, A. Erath X. Huang (2012). "Proceedings of the ACM SIGKDD International Workshop on Urban Computing". In: *Transp. Policy 14 (3)*, 142–148.

L. Zhao W. Y. Ochieng, M. A. Quddus and R. B. Noland (2003). "An Extended Kalman Filter Algorithm for Integrating GPS and Low Cost Dead Reckoning System Data for Vehicle Performance and Emissions Monitoring". In: *The Journal of Navigation*, 257–275.

Langlois, P. (2010). "Simulation des systemes complexes en geographie Fondements theoriques et applications Hermes Science Publications." In:

Le Moigne, J.L. (1999). "La modelisation des systemes complexes". In:

Lefloch, Damien (2007). "Real-Time People Counting system using Video Camera". In: *Master thesis*.

Lesas, Anne-Marie (2015). "Une perspective de la mobiquité au service de la gestion avant / pendant / après des séismes". In: *Prototypage autour du projet SISMAPP*.

Li, R. and Q. Shi (2003). "Study on integration of urban traffic control and route guidance based on multi-agent technology". In: *Proc. IEEE Int. Conf. Intell. Transp. Syst. (Cat. No. 03TH8708)*, 1740–1744.

Lopez Nicolas C., Molina Castillo F. J. and H. Bouwman (2008). "An assessment of advanced mobile services acceptance: Contributions from TAM and diffusion theory models." In: *Information and Management*, pp. 359–364.

Lu, H. (2007). "Network Smart Card review and Analysis." In: *Computer Networks 51*.

M. A. Quddus W. Y. Ochieng, L. Zhao and R. B. Noland (2003). "A general map matching algorithm for transport telematics applications". In: *GPS Solutions*, 157–167.

M. A. Quddus, R. B. Noland and W. Y. Ochieng (2006). "A high accuracy fuzzy logic based map matching algorithm for road transport". In: *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 103–115.

M. Azizyan, I. Constandache and R. Roy Choudhury (2009). "Surround-Sense: mobile phone localization via ambience fingerprinting". In: *in MobiCom*, 261–272.

M. C. Choy, D. Srinivasan and R. L. Cheu (2003). "Cooperative, hybrid agent architecture for real-time traffic signal control". In: *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, 597–607.

M. Rossi, O. Amft and G. Troster (2012). "Collaborative personal speaker identication: A generalized approach pervasive and mobile computing". In: *Elsevier Pervasive and Mobile Computing (PMC)*, 1574–1192.

M. Stager, P. Lukowicz and G. Troster (2004). "Implementation and evaluation of a low-power sound-based user activity recognition system".

In: *in Eighth International Symposium on Wearable Computers (ISWC '04)*, 138–141.
M. V. Belmonte J. L. Perez-de-la-Cruz, F. Triguero and A. Fernandez (2005). "Agent coordination for bus fleet management". In: *Proc. 20th Annu. ACM Symp. Appl. Comput.* 462–466.
Ma X., Wang Y. Chen F.-Liu-J. (2013a). "Mining Smart Card Data for Transit Riders' Travel Patterns." In: *the 92 nd meeting of the Transportation Research Board*, pp 13–3460.
— (2013b). "Transit Smart Card Data Mining for Passenger Origin Information Extraction". In: *the 92 nd meeting of the Transportation Research Board*, pp 13–2156.
Maciejewski M., Nagel K. (2012). "Towards Multi Agent Simulation of the Dynamic Vehicle Routing Problem in MATSim." In:
Mahatanankoon P., Wen H. J. and B. Lim (2005). "Consumer - based m - commerce: exploring consumer perception of mobile applications." In: *Computer Standards and Interfaces*, pp. 347–357.
Mallat N., Rossi M. Tuunainen V. K. and A. Oorni (2009). "The impact of use context on mobile services acceptance: The case of mobile ticketing." In: *Information and Management*, pp. 190–195.
Manheim, M.L. (1979). "Fundamentals of Transportation Systems Analysis". In: *Cambridge, MA: MIT Press*, p. 44.
Marcadon, J. (1997). "Les transports: geographie de la circulation dans le monde d'aujourdhui," in:
Marie-Pier Pelletier, Martin Trépanier and Catherine Morency (2009). "Smart Card Data in Public Transit Planning: A Review". In: *CIRRELT-2009-46*, pp. 5–19.
Marilleau, N. (2006). "Methodologie formalismes et outils de modelisation simulation pour letude des systemes complexes application a la mobilite geographique". In:
Markus, M. L. (1987). "Toward a "critical mass" theory of interactive media." In: *Communication Research*, pp. 491–511.
Meng L Liu S, Striegel A (2014). "Analyzing the impact of proximity, location, and personality on smartphone usage." In: *IEEE INFOCOM workshop on dynamic social networks*.
Min JK, Cho SB (2011). "Mobile human network management and recommendation by probabilistic social mining." In: *IEEE Trans Syst Man Cybern—Part B: Cybern*, 761–771.
Min JK Wiese J, Hong JI Zimmerman J (2013). "Mining smartphone data to classify life-facets of social relationships." In: *CSCW '13 Proceedings of the 2013 conference on Computer supported cooperative work*, 285–294.
Mirco Rossi Sebastian Feese, Oliver Amft Nils Braune-Sandro Martis Gerhard Troster (2013). "AmbientSense: A Real-Time Ambient Sound Recognition System for Smartphones". In: *International Workshop on the Impact of Human Mobility in Pervasive Systems and Applications 2013*.
Montoliu, R. and D. Gatica-Perez (2010). "Discovering human places of interest from multimodal mobile phone data." In: *In Proc. Int. Conf. on Mobile and Ubiquitous Multimedia*.
Morency C., Trepanier M. Agard B. (2006). "Analysing the variability of transit users behaviour with smart card data." In: *The 9th International IEEE Conference on Intelligent Transportation Systems – ITSC*.
N. Lecomte R. Patesson, P. Steinberg. In:

Nabaa M., Zeddini B. Tranouez P. (2007). “Approche decentralisee pour resoudre le probleme du transport a la demande.” In:

Nasri Bin Othman Erika Fille Legara, Vicknesh Selvam-Christopher Monterola (2015). “A data-driven agent-based model of congestion and scaling dynamics of rapid transit systems”. In: *Journal of Computational Science 10*, 338–350.

network, Aruba (2012). “Wi-Fi for In-Store Mobility: Engage Customers and Protect the Store”. In: *Aruba White Paper*.

Nham B., Siangliulue K. Yeung S. (2008). “Predicting mode of transport from iPhone accelerometer data”. In: *Technical report , Stanford University.*

Noopuran, Sivaguru (2014). “Exclusive: Proximity Sensing in Mobile Phones”. In: *Product Marketing Engineer and Priyadeep Kaur, Application Engineer, Cypress Semiconductor*.

North M.J., Collier N.T. Vos J.R. (2006). “Experiences creating three implementations of the repast agent modeling toolkit.” In:

N.Ravi and L. Iftode (2007). “Fiatlux: Fingerprinting rooms using light intensity”. In: *in na*.

Nysveen H., Pedersen P. and H. Thorbjorn sen (2005). “Intentions to use mobile services: antecedents and cross - service comparisons”. In: *Journal of the Academy of Marketing Science*, 330–346.

Ochieng W. Y., Quddus M. Noland R. B. (2003). “Map-matching in complex urban road networks”. In: *Brazilian Journal of Cartography (Revista Brasileira de Cartografia) 55*, 1–18.

Parker MJ Manan A, Urbanski S (2012). “Prospective evaluation of direct approach with a tablet device as a strategy to enhance survey study participant response rate.” In: *BMC Res Notes 5*, p. 605.

Paul Hamilton, Suresh Sankaranarayanan (2013). “Intelligent Agent Based RFID System for on Demand Bus Scheduling and Ticketing”. In: *International Journal of Future Computer and Communication*, 480–489.

Pelletier M., et al. (2010). “smart card data use in public transit: A literature review”. In: *Transportation Research part C 19*, pp. 557–568.

Pelletier M., Trepanier M. Morency C. (2009). “Smart Card Data in Public Transit Planning: A Review.” In: *TRB 2015 Annual Meeting*.

Poo Kuan Hoong, Ian K.T. Tan and Chai Kai Weng (2015). “A comparison of people counting techniques via video Scene analysis”. In: *ARPN Journal of Engineering and Applied Sciences*.

Quddus M. a., Noland R. B. Ochieng W. Y. (2005). “Validation of map matching algorithms using high precision positioning with GPS”. In: *Journal of Navigation 58*, 257–271.

Quddus M. A., Ochieng W. Y. Noland R. B. (2007). “Current map-matching algorithms for transport applications: State-of-the art and future research directions”. In: *Transportation Research Part C: Emerging Technologies 15*, 312–328.

Rachuri, Kiran K. (2012). “Smartphones based Social Sensing: Adaptive Sampling, Sensing and Computation Offloading”. In: *Cambridge University*.

Ravi N., Dandekar N. Mysore P. Littman M. L. (2005). “Activity recognition from accelerometer data”. In: *the Seventeenth Conference on Innovative Applications of Artificial Intelligence(IAAI)*, 1541–1546.

Reddy S., Burke J. Estrin D. Hansen M. Srivastava M. (2009). “Determining transportation mode on mobile phones, Wearable Computers”. In:

ISWC 2008. 12th IEEE International Symposium on Wearable Computers, 25–28.

Richmond, B.M. (1985). "Software for Bringing System Dynamics to the Other 98 percent". In: *the 1985 International Conference of the Systems Dynamics Society.*

Robert P., Rey A. (2001). "Le Grand Robert de la langue francaise Dictionnaires Le Robert". In:

Rogers, E. M. (2003). "Diffusion of Innovations (5 th edition)". In: *New York, NY: Free Press*.

Roozemond, D. A. (1999). "Using autonomous intelligent agents for urban traffic control systems". In: *Proc. 6th World Congr. Intell. Transp. Syst.*

— (2001). "Using intelligent agents for pro-active, real-time urban intersection control". In: *Eur. J. Oper. Res. , vol. 131, no. 2*, 293–301.

S. Mazilu U. Blanke, A. Calatroni and G. Troster (2013). "Low-power ambient sensing in smartphones for continuous semantic localization". In: *in Ambient Intelligence . Springer International Publishing*, 166–181.

S. Ossowski, J. Cuena and A. Garcia-Serrano (1998). "A case of multiagent decision support: Using autonomous agents for urban traffic control". In: *IBERAMIA—Progress in Artificial Intelligence*, 100–111.

Sarala, R. (2010). "The impact of cultural differences and acculturation factors on post-acquisition conflict". In: *Scandinavian Journal of Management*, pp. 38–56.

Schmitt, C. (2013). "Le modèle SimpopNet pour la simulation de la coévolution entre reseaux de transport et systeme de villes. In Modélisation et simulation spatiale." In:

Schneider Robert; Patton, Robert; Toole Jennifer (Toole Design Group) Raborn Craig (PBIC/UNC-CH) (2005). "Pedestrian and Bicycle Data Collection". In:

Schuessler N., Axhausen K. (2009). "Processing raw data from global positioning systems without additional information, Transportation Research Record". In: *Journal of the Transportation Research Board 2105*, 28–36.

Shao, Guosong (2009). "Understanding the appeal of user-generated media: a uses and gratification perspective". In: *Internet Research*.

Sharaby Y., et al. (2012). "The impact of fare integration on travel behaviour and transit ridership". In: *Transport Policy 21*, pp. 63–70.

Singh VK Freeman L, Lepri B Pentland A (2013). "Predicting spending behavior using socio-mobile features". In: *BioMedCom 2013*.

Stenneth L., Wolfson O. Yu P. S. Xu B. (2011). "Transportation mode detection using mobile phones and GIS information". In: *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 54–63.

Stern B.B., Royne M.B. Sttaford T.S. and C.C. Bienstock (2008). "Consumer acceptance of online aucti ons: an extension and revision of the TAM". In: *Psychology and Marketing*, 619–636.

Stopher P., Clifford E. Zhang J. Fitzgerald C. (2008). "Deducing mode and purpose from GPS data." In:

T. Franke P. Lukowicz, K. Kunze and D. Bannach (2009). "Can a mobile phone in a pocket reliably recognize ambient sounds?" In: *in International Symposium on Wearable Computers (ISWC 09)*, 161–162.

T. Kusakabe T. Iryo, Y. Asakura (2011). "Estimation method for railway passeng0rs train choice behavior with smart card transaction data". In: *Transportation 37*, 731–749.

T. Watanabe D. Kamisaka, S. Muramatsu (2012). "At which station am I?: Identifying subway stations using only a pressure sensor". In: *in ISWC*, 110–111.

Tao, C.C. (2007). "Dynamic taxi sharing service using intelligent transportation system technologies." In: *In Wireless Communications International Conference*, 3209–3212.

Treil, Sylvain (2005). "Comprendre l'evolution de la mobilite entre deux enquetes menages deplacements". In: *Laboratoire d'Economie des Transports – Ecole Nationale des Travaux Publics d'Etat*.

Trepanier M. Chapleau R., Tranchant N. (2007). "Individual trip destination estimation in transit smart card automated fare collection system." In: *Journal of Intelligent Transportation Systems. Technology, Planning, and Operations, Taylor and Francis*, 1–15.

Treuil J.P., Drogoul A. Zucker J.D. (2008). "Modelisation et simulation a base dagents exemples commentes outils informatiques et questions theoriques Dunod". In:

Troitzsch, K.G. (1997). "Social science simulation origins, prospects, purposes. Lecture Notes in Economics and Mathematical Systems". In:

Utsunomiya M., Attanucci J. Wilson N. (2001). "Transportation Technologies: Implications for Planning." In: *University of California Transportation Center paper 536*, p. 27.

— (2006). "Potential Uses of Transit Smart Card Registration and Transaction Data to Improve Transit Planning, Transportation Research Record". In: *Journal of the Transportation Research Board, No. 1971, Transportation Research Board of the National*, 119–126.

Venkatesh V., Morris M. G. Davis G. B. and F. D. Davis (2003). "User acceptance of information technology: Toward a unified view. MIS quarterly". In: *Management science*, pp. 425–478.

Vidal, J.M. (2006). "Fundamentals of multiagent systems: using netLogo models." In:

Von Neumann J., Burks A.W. (1966). "Theory of self reproducing automata". In:

Vuchic, V. (2005). "Urban Transit; Operations, Planning and Economics." In: *Hoboken, New Jersey: John Wiley and Sons.*

WBCSD (2016). "Methodology and indicator calculation method for sustainable urban mobility". In: *CIRRELT-2009-46*.

Westphal, C. and T. Blaxton (1998). "Data Mining Solutions, John Wiley, New York." In:

Wilensky U., Stroup W. (1999). "NetLogo HubNet Gridlock model." In:

Winter, S. (2008). "Intelligent Self Organizing Transport." In: *Artificial Intelligence and Spatial Mobility*, pp. 25–28.

Wu Y., Guan L. Winter S. (2008). "Peer to peer shared ride systems." In: *GeoSensor Networks*, pp. 252–270.

X. Shi J. Xu, Y. Xu and J. Song (2005). "A simulation study on agent-network based route guidance system". In: *Proc. IEEE ITSC*, 248–253.

Xie X, Leurent F Aguiléra V. (2015). "Commuter's individual travel time estimation: using AFC data of Paris transit system. COST Action TU1004

Model. Public Transp. Passeng. Flows Era Intell. Tr ansp. Syst. Paris Fr. May 11th-12th 2015." In: *CIRRELT-2009-46*, 198—211.

Y. Lou C. Zhang, X. Xie Y. Zheng W. Wang and Y. Huang (2009). "Map-Matching for Low- Sampling-Rate GPS Trajectories". In: *GIS '09 Proceedings of the 17th ACM SIGSPATIAL Inter- national Conference on Advances in Geographic Information Systems*, 352–361.

Yuan J., Zheng Y. Zhang C. Xie W. Xie X. Sun G.-Huang-Y. (2010). "T-Drive: Driving directions based on taxi trajectories". In: *ACM SIGSPATIAL GIS 2010*.

Zargayouna M., Zeddini B. (2012). "Fleet Organization Models for Online Vehicle Routing Problems." In:

Zhang, Sen (2012). "Smartphone Based Activity Recognition System". In: *THESIS*.

Zhao J, Frumin M Wilson N Zhao Z. (2013). "Unified estimator for excess journey time under heterogeneous passenger incidence behavior using smartcard data". In: *Transp. Res. Part C Emerg. Techno*, pp. 70–88.

Zhao, X. (2011). "Algorithmes pour les problemes de tournees a la demande." In:

Zheng Y., Liu L. Wang L. Xie X. (2008). "Learning transportation mode from raw gps data for geographic applications on the web". In: *Proceeding of the 17th international conference on World Wide Web , Beijing, China.*