Management and monitoring of a Wi-Fi network for passenger counting in public transport

Marta Martins Bernardo
marta.bernardo@tecnico.ulisboa.pt

Master Programme in Telecommunications and Informatics
Instituto Superior Técnico, Lisbon, Portugal

Abstract -* Over the past few years there has been an increasing need to ensure better conditions in the area of public transportation due to the increased demand for these services. This demand leads to greater competitiveness from the service providers, being the quality and efficiency of the service the differentiating factors. In this way, knowing the number of passengers in each bus, in real time, is a parameter that can help to improve this offer. The present thesis of Master in Telecommunications and Informatics Engineering intends to provide the companies that supply this service with a way to improve their quality. A bibliographic review was initially carried out, covering the subjects related to real time passenger counting, the methods used and the systems currently used. After an analysis, it was concluded that for this context, where is intended to count for passengers of a bus almost always in motion, the best solution would to be use the Wi-Fi signal. With the implementation of a Wi-Fi network inside a bus it was possible to understand the number of people who make use of the service, as well as the average time spent inside the bus. About 67.7% of passengers spend less than an hour enjoying the Wi-Fi network service. This also corresponds to the approximate average time spent inside the bus.

Keywords: Mobile devices, Wi-Fi, public passenger transports, Rodonorte.

Introduction

According to studies [1], almost half of the Portuguese population currently lives in urban areas. It is estimated that this trend will become more accentuated over the years. Smart mobility can be a plus in the process of making the city smarter. With this in mind, it is necessary to acquire operational data to monitor the entire transport network. One of the important data sets to improve service is the number of passengers in real time.

In July 2017, 71.6% of individuals had mobile phones with multiple extra features, eg accelerometer, Global Positioning System (GPS), wireless Fidelity (Wi-Fi), etc. This type of functionality has boosted the use and development of multiple applications and services based on location.

These devices are part of the day-to-day operations of passengers and as such, they can be used to get a passive collection of data relating to their mobility with the aim of improving the quality of public transport services. By using this type of device, you can increase the amount of data collected with no additional cost. In addition, no maintenance of the equipment or transport of any type of device is required for such data collection.

Wi-Fi technology offers advantages for use in a location system, since it allows you to acquire indoor and/or outdoor location data.

Through a Wi-Fi network it is possible to determine the number of devices connected to a particular AP, allowing a simple calculation of the number of passengers on board and therefore optimizing the planning and quality of public transportation.

The main goal of this dissertation is to provide data that allow companies that provide public passenger transport service to have a better knowledge of the use of each vehicle in real time. Regarding this, a Wi-Fi router is implemented inside a Rodonorte bus whose data will be analyzed later, in order to count for its passengers.

LITERATURE REVISION

In this section we analyze topics related to passenger counting, current methods and systems, the concept of cloud computing and location based on Wi-Fi networks.

I. Methods of counting passengers

The number of passengers using the public transport service can be counted by three different methods: (1) carrying out questionnaires to the passengers themselves;

(2) use of equipment installed inside the vehicle; (3) voluntary provision by the passenger himself. [2]

The first situation, corresponding to the performance of questionnaires, presents high execution costs. In this way, they do not allow data collection in real time, making it reduced to only a small sample of passengers.

The second method is the most used today, because it is easy to use. This method corresponds to the installation of equipment specially developed for monitoring and collecting data such as the card reader system present in the bus entrances.

The third method suggested is increasingly becoming the most viable and used option, due to the great evolution and popularization of smartphones worldwide. In this type of method, passengers provide access to their smartphones to enable the collection of operational data - crowdsourcing.

II. Wi-Fi Location

A Wi-Fi network is composed of several devices that communicate with each other, being possible to highlight Base Stations (BS), which in this case will be smartphones, and Access Points (APs). The AP is responsible for ensuring communication with the server (via cable) and with the other network stations (through radio frequency signals). [3]

Radiofrequency signals, when picked up by a receiver have three attributes that may be important for location effects: the angle of arrival of the signal, ie the angle of incidence at the receiver; the arrival time of the signal, which can be used to estimate the travel time between the sender and the receiver; the power with which the signal is received, ie the Received Signal Strength Indication (RSSI).

There are a couple of techniques that can be used when estimating location through the radiofrequency signals, each one with its advantages and disadvantages.

In triangulation the position of the device is estimated by forming a triangle between the APs and their respective angles of arrival. This technique is known as Angle of Arrival (AoA). On the other hand, in trilateration, if the time of arrival or the power to estimate the location is used, it is necessary to know three reference points (APs) and their range radius.

III. Current passenger count systems

Currently there are technologies implemented with the goal of counting people in certain contexts. In the intended context, that is, for real-time passenger counting of a particular bus, two types of systems can be distinguished [4]:

1. Infrared sensors;
2. Video sensor and image processing.

Infrared sensors have numerous advantages, including reduced cost and size, easy installation and reliability. These factors justify their use in several counting systems. However, it has a high sensitivity to noise in situations with a large number of people.

On the other hand, the counting system through image capture is not new in the present day, thus presenting several techniques developed to deal with the counting problem:

- **Motion Detection and Analysis-Based Techniques:** the first step in this technique is to detect movement corresponding to individuals. Then use these results to determine the trajectory of the movement, in order to count the people who crossed a predefined virtual area or line;
- **Techniques based on edge analysis:** corresponds to the extraction of edges for the detection. The objects of interest have a certain shape and organization, corresponding to a set of edges. In the specific case of counting people, a head corresponds to a circular shape;
- **Model-based techniques:** pretends to find regions in the rendered images that correspond to predefined templates. The major disadvantage of this model is that it requires a larger learning database;
- **Spatiotemporal techniques:** involves the selection of lines of interest in the acquired images and their organization in time. The second step involves the application of statistical models to determine the number of people crossing these lines, in order to determine their direction. These techniques have the advantage of being simple to implement, however, a person can be counted several times if it it remains in the same place.

IV. Cloud Computing

After completing the phase corresponding to the collection of sufficient data for the counting of passengers it is necessary to store it for later observation and analysis by the corresponding companies.

In recent years there has been a great development in the area of cloud computing and consequent increase in its use by all companies worldwide, as this service enables new and better forms of computing.

The definition that was accepted by its industry was composed by the Institute of Standards and Technology (NIST) in 2011 and defines cloud computing as:

"Cloud computing is a model for enabling ubiquitous and convenient network access to a shared set of configurable computing resources (eg networks, servers, storage, applications, and services) that can be quickly provisioned and released with minimal effort management or interaction with your service provider. This cloud model consists of five essential features, three service models, and four deployment models." [5]

According to NIST, the Cloud Computing model consists of three service models and four implementation models.

a) Service Models

- **Software as a Service (SaaS)** is a service model intended for use by the end user;
- **Platform as a Service (PaaS)** corresponds to the ability of the service provider to provide the user with the ability to develop applications through a platform or to acquire applications already developed.
- **Infrastructure as a Service (IaaS)**, where the user does not have any control over the physical infrastructure, but has control over the operating systems, storage, installed applications and in some cases, limited control of network resources through virtualization tools.

b) Implementation Models

- **Public Cloud**, where the cloud infrastructure is available for open usage, without exception, of the general public. A pay-per-use system is used, i.e. the user pays only when using a service.
- **Private Cloud** whose infrastructure is used, solely, by a single organization, for its exclusive use. This model can be managed by the organization itself, by outsourcing, or by a combination of both.
- **Cloud Hybrid**, which allows organizations to make the most of both models, combining their needs with the characteristics of each model.
- **Community Cloud** whose cloud infrastructure is shared by a community of organizations with common interests, such as security requirements, privacy policy, etc. This is available for the exclusive use of organizations of a particular community.

FUNCTIONAL ARCHITECTURE

In the present dissertation we intend to integrate different platforms and technologies putting them to work together. In this chapter an approach will be made to these same technologies, their operation and characteristics.

I. IBM Bluemix

IBM recently created a platform that enables the creation, implementation and management of applications quickly and easily through the organizations / companies designated IBM Bluemix. This platform can be classified in one of the types of service described above, more specifically in the PaaS service type.

IBM Bluemix provides enterprise-class entry-level services required by organizations so that their applications are always available for use by customers at any time. Since it is an open source platform, it ensures flexibility to integrate developments and services tailored to the needs of each user.

a) IBM Watson

IBM Watson, PaaS provided by IBM, allows an application to communicate with devices. Figure 1 presents an overview of this platform.

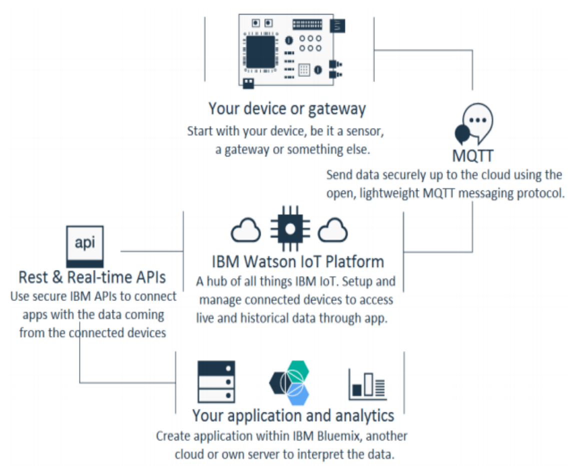

Fig. 1. Overview of IBM Watson.

b) Node-RED Platform

Node-RED is a programming tool used to connect hardware devices, Application Programming Interfaces (APIs) and online services, in new and interesting ways. This provides a browser-based editor, which makes the process easier and has a vast directory of nodes that can be used with a simple click. There are eight main categories of nodes: input, output, function, social, storage, analysis, advanced and Raspberry Pi.

Figure 2 shows the tool interface with the node directory represented on the left, the editor in the center, and the information and debug tests on the right.

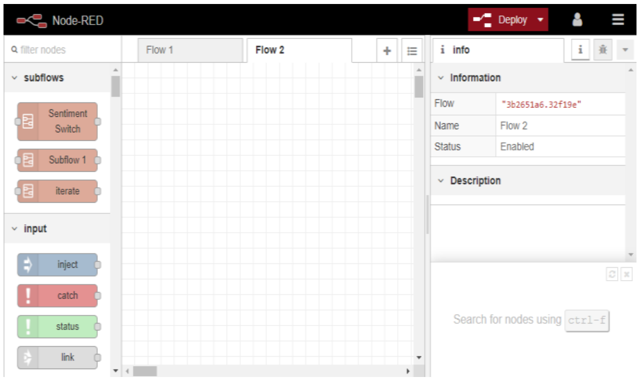

Fig. 2. Node-RED tool interface.

II. *Message Queuing Telemetry Transport (MQTT) Protocol*

To ensure the integration of multiple devices into a network, a standard communication protocol is required. In this case the MQTT protocol will be used, since it is the protocol that is preferentially used for communication between the IBM Bluemix and Node-RED platforms.

MQTT works over a TCP protocol and is based on a publisher/subscriber standard. It was designed with the objective of being open and easy to implement, allowing several clients to access the same server.

As shown in figure 3, the MQTT protocol consists of three components: publisher, subscriber and broker.

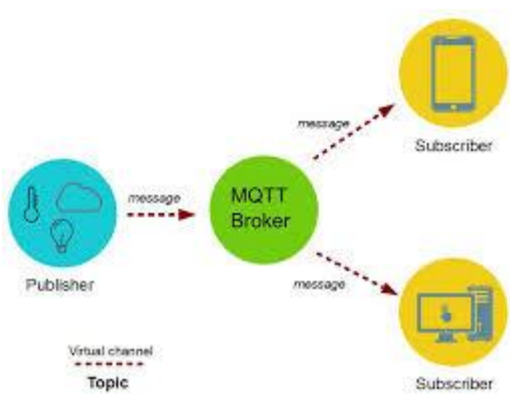

Fig. 3. MQTT protocol architecture.

The publisher is the device that connects to the server for the purpose of sending information, while the subscriber is the device that connects and selects the information that you want to receive. The server that makes this connection, between both devices, is the broker.

Since messages are organized by topics, when sending a message, the publisher should identify the topic of the message. Likewise, subscriber can subscribe to one or more topics that you are interested in receiving messages about. So, whenever a publisher conveys information, it will be sent to interested subscribers via the broker. [6] Figure 4 illustrates the process described.

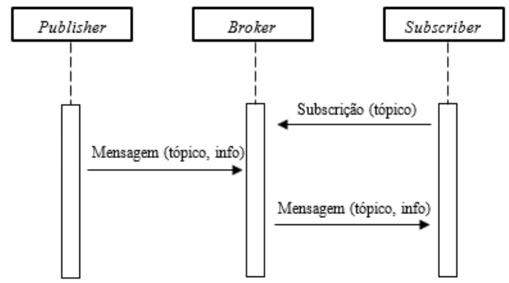

Fig. 4. Publisher/subscriber process useb by MQTT.

SOLUTION PROPOSAL

In order to fulfill the main goal of the dissertation, corresponding to the real-time counting of passengers in a moving bus, a router was set up inside a Rodonorte bus, responsible for collecting the required data. It is also necessary to understand whether the system is applicable to different environments of the public passenger transportation service.

I. *Solution proposal architecture*

The architecture of the solution has been developed to be flexible to several types of networks, allowing to use different types of technologies and topologies.

The architecture of the proposed solution is divided into three blocks (figure 5).

The Wi-Fi network is inside of the bus and ideally it is possible to implement it with any type of equipment intended by the customer. Responsible for providing access to the Internet, this equipment allows collecting operational data for passenger counting.

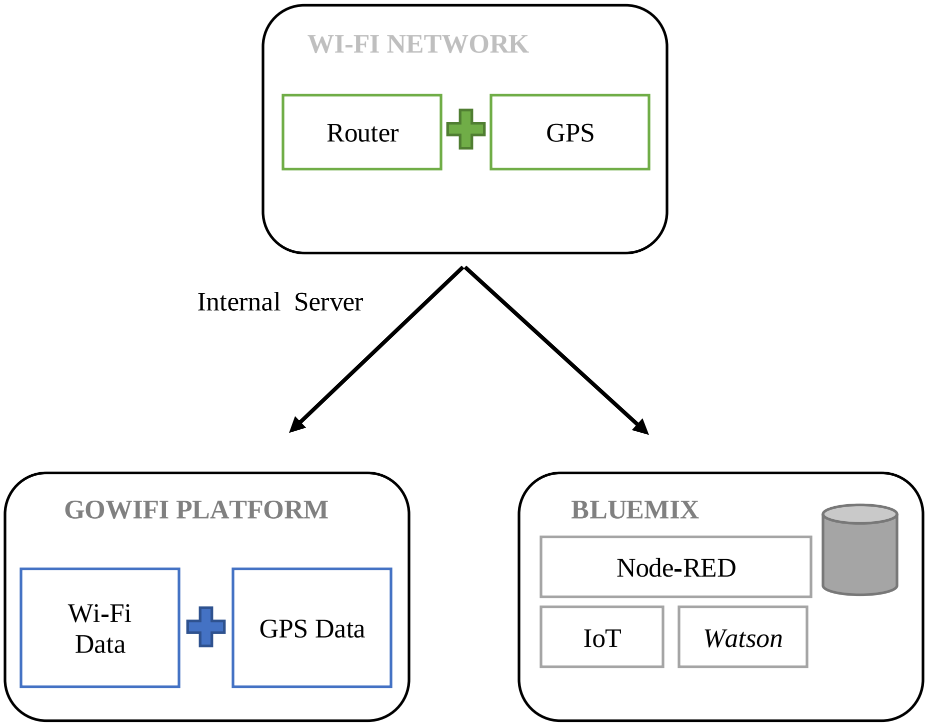

Fig. 5. Logical structure of the solution.

After the connections between devices and AP, the data in each one of these connections is sent to different platforms: through an internal server, to a cloud platform developed by the company and to the cloud platform IBM Bluemix that will be integrated with the Node-RED. Through the GPS data received, with a time interval of 20 seconds, it is possible to cross-reference location information with the number of passengers. The service provider is allowed to understand which stops have more flow of passengers in and out.

II. *Equipment chosen*

There was the possibility of choosing between two different equipments for the implementation of the system, being the initialization time and the probability of failure the main factors for the choice. The equipment provided corresponds to a Teltonika RUT955 router and a Mikrotik 1100 router.

Regarding the startup time, the Teltonika RUT955 is the one that best meets this requirement, with a startup time of one minute. In order to understand which equipment had the most faults, a load test was performed with the aid of software, Apache JMeter, which is regularly used for performance testing.

For the accomplishment of these tests the following parameters were defined:

- 50 users making simultaneous orders, since it is the average number of passengers that a bus can carry;
- HTTP requests to 10 different links, randomly, to ensure that multiple users can access the network;
- The duration of the test was 15 hours, to ensure all-day operation.

Table 3 shows that the best option would be the Teltonika RUT955 router with a total error of 9.47% versus the 22.66% of the Miktotik router.

Teltonika RUT955

In this equipment, the 3G / 4G network is captured through Long-Term Evolution (LTE) antennas while the distribution of Wi-Fi coverage is the responsibility of the Wi-Fi antennas. There is also a GPS antenna that is responsible for collecting the GPS coordinates of the bus.

III. *Methodology*

To achieve the desired goal and to ensure that public transportation service providers have access to the information needed to improve their services, a 4-step methodology, shown in Figure 6, was followed.

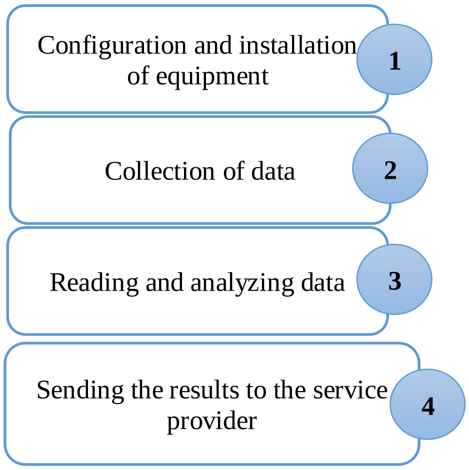

Fig. 6. Used methodology.

In a first phase (1), the equipment chosen was installed inside a bus of the transportation company Rodonorte. In order to have a functional equipment it was necessary to make some configurations in the equipment, particularly in the area of Virtual Private Network (VPN) services, more specifically the creation of a Point-to-Point Tunneling Protocol (PPTP) instance whose function is to create a secure connection between the device and the company's local server.

In order to be possible and facilitate the collection of data it is necessary that there is a certain page of authentication in the network, being necessary also to configure the Wireless Hotspot functionality, which in practice works in a similar way to an AP, however, it is more versatile when it comes to monitoring, management and authentication of users. The defined authentication page followed the standards set by the GoTVee company to ensure compatibility with its data management platform. It would be for this same platform that all the collected data would be redirected, concluding the second phase of the process (2).

Next (3), the data is extracted from the platform in Excel format, having a unique identifier for each device that has connected to the AP as well as the time of connection and the type of device. By changing the settings of the authentication page, it is also possible to register names, emails or mobile phone numbers if it is desired by the service provider.

After completing the analysis phase, the next step (4) is to send all results obtained with the data obtained so that the company in question, in this case Rodonorte, can use that same information to improve its services, in particular whether the resources used are sufficient or exaggerated according to the flow of passengers.

IV. *Applicability to different environments*

Depending on the type of environment in which the system is to be deployed, the system requirements may differ. However, the system must be able to match as well as possible regardless of the environment in which it is inserted.

a) Urban environment

Presents a greater number of stops and greater proximity between them, and consequently a greater flow of passengers in and out. In this case, it would be necessary to ensure that the time interval for the renewal of connection between the device and the AP is smaller, since there are stops approximately every five minutes.

b) Inter-Urban environment

Corresponds to an environment in which the risk of occurrence decreases considerably, typically presenting longer trajectories, with fewer stops and greater distance between them. In this environment, the renewal period of the connection used in the practical case, between fifteen and twenty minutes, is enough to guarantee the collection of reliable and real data.

Results and Discussion

Once the Excel file has been exported, the analysis and filtering of the collected data is carried out, in order to facilitate the interpretation of the data by the company that provides the service.

The data presented correspond to about a week of collection, specifically from April 24, 2019 until May 1, 2019.

The file contains an unique identifier for each device, that is, the MAC Address assigned when connecting the device to the AP. It is also possible to obtain information on the type of device used, in particular its manufacturer. Within the passenger count, the parameter that contains the date and time of the connection is highlighted, and the connection is renewed after about fifteen to twenty minutes. This renewal on the connection makes it possible to estimate the time spent by the passenger on the Internet service and, consequently, to estimate the amount of time spent on the bus by each passenger.

It began by analyzing the number of people on the bus, in order to perceive the influx and flow of passengers over the days. The data obtained is shown in table 1.

Date	Number of passengers
24/04/2019	14
25/04/2019	25
26/04/2019	20
27/04/2019	13
28/04/2019	23
29/04/2019	22
30/04/2019	13
01/05/2019	11

TOTAL:	141

Table 1. Estimated number of passengers per day.

Based on this data it was possible to perform a graphical representation, which is expressed in figure 7.

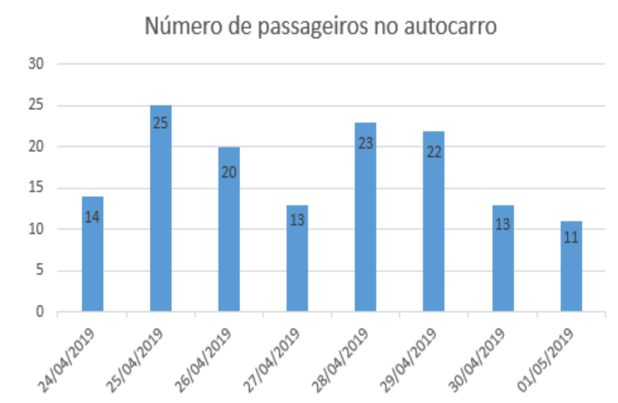

Fig. 7. Graphical representation of the number of passengers on the bus, per day.

As mentioned previously, there is a renewal of the connection in almost constant time intervals which allows to check approximately how many times each device makes a connection to the AP and, consequently, to estimate the average time of use of the service and possible time spent inside the bus. Table 2 shows an estimate of the average time that users spend connected to the network during their travels, which roughly corresponds to the time of travel. Next, figure 8 corresponds to the graphical representation of the same data.

Less than 1h	About 1h	About 2h	About 3h	More than 3h
67,7%	6,3%	9,4%	6,3%	10,3%

Table 2. Average access time to the Wi-Fi network by the passengers.

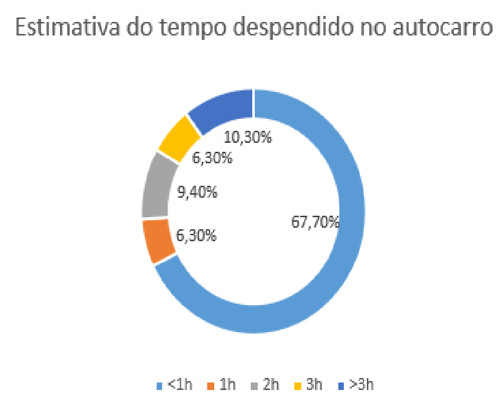

Fig. 8. Graphical representation of the average time spent on the bus.

In relation to these data, it was found that one of the devices maintained the connection to the AP for approximately nine hours which may indicate that it is the bus driver instead of a passenger.

It was also observed that, on two occasions, the same passenger traveled on two separate days, that is, both passengers were counted twice in this scenario.

The data collected may also be used by businesses for commercial purposes, particularly advertisers because they can access the type of device that has connected to the AP. In Table 3 and Figure 9 these same data and their graphical representation are shown, respectively.

Apple	Samsung	Asus	Xiaomi	Outros
27,6%	19,3%	2,8%	2,8%	47,8%

Table 3. Types of devices used by passengers.

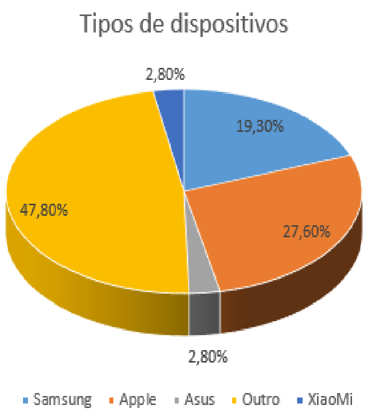

Fig. 9. Graphical representation of the types of devices used.

The collection of this type of data can allow any company to be aware of the most common devices in its passengers which can be used in advertising issues in the case of their respective brands.

CONCLUSIONS

The main goal of this dissertation was the real need of a company that provides services in the area of public passenger transportation, Rodonorte, customer of the portuguese company GoTVee.

The main need was to find a low-cost, easy-to-implement passenger counting solution. Existing systems through ticketing, infrared sensors, or image collection are often very costly and require a high degree of implementation and maintenance processes. These two systems still have the drawback of being affected by sunlight as buses circulate in outdoor spaces.

As an alternative to these systems Card4B and GoTVee are looking for an solution solution with a lower cost based on the use and detection of Wi-Fi signals as a method of counting passengers. The only downside to this solution is the assumption that all passengers carry a mobile device with them.

The literature review allows anyone who has an interest in developing a system in this area to understand the basic concepts for achieving it.

The main goal of this dissertation was reached, and Rodonorte was provided with a set of data and operational results that can be used to improve the quality of services. However, compatibility with the IBM Bluemix platform was not possible as it is incompatible with the type of equipment provided, specifically with the Teltonika RUT955 router.

In a next phase, the network authentication page may also allow the collection of names, e-mails and telephone contacts if the public transportation company wants. This additional data will enable better knowledge of your passengers. This data was not collected throughout the dissertation since Rodonorte wanted to only carry out a test implementation. However, following this dissertation, the company decided to maintain the service and increase the number of parameters collected, for the future.

As a way to make the equipment compatible with the intended platforms, as future work, the ideal would be the replacement by compatible equipment. However, due to the high cost of this action, changing the cloud computing platforms would be enough.

REFERENCES

[1] https://www.jornaldenegocios.pt/economia/detalhe/vivem_muitas_pessoas_nas_cidades_na_europa_40_da_populacao_e_em_portugal_quase_44; Acessed 28 november 2018

[2] MELO, Ângelo dos Santos. Método de sensoriamento de passageiros do transporte público com uso de smartphones e bluetooth – Universidade Federal de Santa Catarina, Florianópolis, 2016.

[3] MOURA, Andre Iasi. WBLS: um sistema de localizacao de dispositivos moveis em redes Wi-Fi. 2007. Dissertacao (Mestrado em Engenharia Eletrica) – Escola Politecnica da Universidade de São Paulo, 2007. Available at: <http://www.teses.usp.br/teses/disponiveis/3/3141/tde-07012008-171855/pt-br.php>. Acessed 2 may 2018.

[4] T Yahiaoui, L Khoudour, C Meurie; Real-time passenger counting in buses using dense stereovision; 2013; available at: https://hal.archives-ouvertes.fr/hal-00855979/document; Acessed 28 april 2019

[5] Mell P., Grance T; The NIST Definition of Cloud Computing; 2011; available at: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf - Acessed 14 march 2019

[6] LAMPKIN, V. et al. Building Smarter Planet Solutions with MQTT and IBM WebSphere MQ Telemetry. IBM Redbooks, p. 270, 2012. Available at: https://books.google.pt/books?id=F_HHAgAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false