When Technology Supports Urban Mobility: Improvements for Mobile Applications Based on a UX Evaluation

Rodrigo L.A. Almeida(✉), Lana B. Mesquita, Rainara M. Carvalho, and Rossana M.C. Andrade

Group of Computer Networks, Software Engineering and Systems (GREat)
Department of Computer Science, Federal University of Ceará, Fortaleza, Brazil
{rodrigoalmeida,lanamesquita,rainaracarvalho,rossana}@great.ufc.br

Abstract. The development of applications that helps urban mobility has been pushed by the increase of processing capacity and miniaturization of the mobile device as well as the improvement of speed and availability of Internet. Such applications can support users in several activities such as tracing routes and searching for addresses. In this scenario, this work aims to: *(i)* understand how users use urban mobility applications; *(ii)* evaluate the quality of interaction and interface into real applications called "Waze" and "Meu Ônibus"; and *(iii)* improve the applications based on the results of the evaluations. Moreover, this study also suggests a set of recommendations for urban mobility applications in developing countries like Brazil.

Keywords: Urban mobility · Mobile applications · Usability · User experience · Evaluation methods

1 Introduction

Urban Mobility represents a major problem in medium and large cities of Brazil [14,18]. The main reason is the amount of people living in urban areas. For instance, according to the latest census of the Brazilian Institute of Geography and Statistics (IBGE)[1], the number of people living in urban spaces continues to grow and represents 84% of the population [6].

Moreover, the infrastructure capacity and quality of the city do not correspond the growth of the population, what consequently creates many problems, for example, traffic jam, roads in bad conditions, low number of public transports, excess capacity of people in public transports, among others.

The evolution of the information technologies, especially mobile technology, enables the development of solutions to minimize the problems mentioned

R.L.A. Almeida—Scholarship (MDCC/DC/UFC) sponsored by FCPC.
L.B. Mesquita—Scholarship (MDCC/DC/UFC) sponsored by FUNCAP.
R.M.C. Andrade—Researcher scholarship - DT Level 2, sponsored by CNPq.

[1] http://goo.gl/ckJf7k.

M. Kurosu (Ed.): HCI 2017, Part II, LNCS 10272, pp. 111–130, 2017.
DOI: 10.1007/978-3-319-58077-7_10

before [3]. For example, we can cite the use of sensors to collect user information [5,7,8,16,27], the improvements in processing power of mobile devices, and the increased speed and availability of Internet [17].

Examples of these solutions are mobile applications for several means of transport (*e.g.,* car, bus, taxi, and bicycle). These solutions are changing the way people move. Currently, there are applications to trace routes, check traffic flow, monitor transports in real-time, and so on. Thus, the user experience of such technologies is an important factor to ensure their use and adoption.

There is previous research about user experience with urban mobility applications [4,9,28]. Nevertheless, more investigation is necessary to understand and to improve the user experience with such technologies in a large city in a developing country like Brazil. In this context, the XV Brazilian Symposium on Human Factors in Computational Systems launched an evaluation competition with the following theme "User Experience in Urban Mobility Applications"[2]. In this competition, researchers should use HCI evaluation methods to discover UX problems. In this scenario, our team, who are the authors of this paper, participated in this contest with two goals: *(i)* understand how people use urban mobility apps in Brazil; and *(ii)* evaluate the most used urban mobility applications.

To understand how people use mobile applications and what are their opinion about them, we used a *Survey*, in which we identified two most used applications: Waze[3], the Brazilian version, and "Meu Ônibus"[4], that means My Bus in English.

To evaluate the identified applications, we used a Heuristic Evaluation [21] and a Usability Testing [20]. The first one gives the necessary input to decide how to apply the next one. For example, we chose the problematics activities to explore in the usability evaluation by users.

Therefore, this paper presents not only an overview of the results from the competition [1], but also an evolution of this work by *(i)* performing a *Sentiment Analysis* in the Google Play reviews and *(ii)* proposing improvements for the identified problems. The purpose of the Sentiment Analysis is to identify the level of user satisfaction through the opinionated content and categorize it as positive, negative or neutral, as well as many other studies [10,23,24,29]. To suggest the improvements of the interfaces, first, we performed a brainstorming based on the previous results [1] with HCI experts to plan the improvements. After that, we applied a *Survey* with users to evaluated the suggestions of the interface. As a consequence of this research, this study also suggests a set of recommendations for urban mobility applications in developing countries like Brazil.

This paper is organized as follows: Sect. 2 presents the adopted methodology, which is composed of three phases and each one has a description of the used methods and instruments. Section 3 shows the primary results of all used

[2] http://ihc2016.mybluemix.net/br/cfp.jsp.

[3] Waze - https://play.google.com/store/apps/details?id=com.waze\&hl=pt.

[4] Meu Ônibus - https://play.google.com/store/apps/details?id=br.com.m2m.meonibus&hl=pt.

methods. Section 4 presents a discussion of the results, along with a set of recommendations for urban mobility applications in Brazil. Finally, Sect. 5 presents the conclusions of this work.

2 Methodology

The adopted methodology of this work is composed of three phases (See Fig. 1). The first two phases are from the evaluation competition, except the sentiment analysis, which we performed for this paper. The third phase is an evolution of the work from the contest, which is the main contribution of this paper. We present each one in the next subsections.

Fig. 1. Adopted methodology

2.1 First Phase - To Understand

This first phase aims to *understand* how users are using urban mobility applications. We defined the following three questions to achieve this goal:

1. What are the most popular applications used to support user locomotion?
2. How users use these applications?
3. What are its main strengths and weaknesses in the users' opinion?

The technique used to collect the data was a *Survey* since this is a good way to get answers to specific questions from a large number of people [25]. We use the *Google Forms*[5] platform to craft the online survey. It has four sections of fifteen questions. The first section concerns user consent to participate in the survey and confirm if he/she uses urban mobility applications. The second section concerns to user demographics. The third section deals with questions about the use of transport in daily activities. The fourth and final section is related to the use of technology for urban mobility.

Through the analysis of survey responses, it was possible to trace two profiles of users and to identify two most used applications ("Waze" and "Meu Ônibus"). Additionally, we could identify the most used features, the suggestions for improvement, several kinds of urban mobility, and technology issues. All this information was useful for the second phase of this work. The answer for the first-phase question is present in the results section.

2.2 Second Phase - To Evaluate

This phase aims to *evaluate* the user experience with two applications identified in the previous phase. Therefore, we collect data about efficiency, effectiveness, and satisfaction, as well as problems and interface improvement opportunities. Like the previous phase, we define three questions to support the evaluation, as follows:

1. Which interaction problems did users have to face?
2. How much effectiveness and efficiency users can achieve their goals?
3. What are the satisfaction the user about of these applications?

Then, we use three methods: Heuristic Evaluation [21], User Evaluation [20] and the *Sentiment Analysis*, applied on the Google Play reviews.

Heuristic Evaluation. The set of heuristics called SMASH (Smartphone Usability Heuristics) was used by three experts [11]. We chose this set because it has specific heuristics for mobile devices and applications, as follows: *1*- Visibility of system status; *2*- Match between system and the real world; *3*- User control and freedom; *4*- Consistency and standards; *5*- Error prevention; *6*- Minimize the user's memory load; *7*- Customization and shortcuts; *8*- Efficiency of use and performance; *9*- Aesthetic and minimalist design; *10*- Help users recognize, diagnose, and recover from errors; *11*- Help and documentation; and *12*- physical interaction and ergonomics. Three experts, who have at least two years of experience in researching human-computer interaction, performed the evaluation.

The activities for evaluators performing in each application were based on the most used features by users who answered the survey of phase 1. The following activities have been defined for "Meu Ônibus": *(1)* to view transport schedules; *(2)* to verify which lines pass through a particular place; *(3)* to find bus stops; and *(4)* to find an itinerary/route of a bus. For "Waze": *(1)* to verify traffic real-time; *(2)* to search by address; *(3)* to verify travel time between two locations; *(4)* to create a route; and *(5)* to follow the route real-time.

[5] The following link provides a copy of our survey: https://github.com/GREatPesquisa/2017-Research-Urban-Mobility-Applications.

User Evaluation. Ten users (seven men and three women) were recruited to use "Meu Ônibus" and "Waze". We selected users by taking into account the two profiles identified in the survey results (see Fig. 2). Five members represent the profile A, so they used "Meu Ônibus". The other five represents the profile B and then used "Waze". The results section contains the profiles explanation.

The activities performed by the users were the same as those carried out by the experts in the heuristic evaluation. The device used for both "Meu Ônibus" and "Waze" was the LG Nexus 5 model with 2GB of RAM and 2.3 GHz Quad Core processor. The versions used for "Meu Ônibus" and "Waze" applications were: 1.0 and 4.3.0.2, respectively. We delivered this device to the user with the application already configured for use. Finally, we performed two pilot tests, one for "Meu Ônibus" and another for "Waze", with different users from those recruited for the evaluations.

The data collection consisted of observing the users performing the activities. In "Meu Ônibus", all activities were performed in the laboratory. In "Waze", only one of the activities was carried out in the field due to its nature: *(v)* track the path in real-time. For that, the places of origin and destination were predetermined by the evaluators and previously said to the users. Thus, the execution of this activity required the user's car. Therefore, the "Waze" assessment was performed only by users who had a driver's license, own car and availability to use it during the evaluation.

During the observation sessions, we planned to obtain the following data: completeness of the activity, number of interactions and number of problems. The following instruments were used to collect them: think-aloud technique [22], audio recordings and videos of the screens, both performed by the Az Screen Recorder application[6], and also notes taken by the observers.

Finally, the interpretation and consolidation of results were performed. For the interpretation of the data, we analyzed the user interaction videos, *think-aloud* audio transcripts, and notes from the observer. Three HCI evaluators performed the analysis of the data which resulted in a list of problems and improvements opportunities.

Sentiment Analysis. With the purpose of understanding better the user perception about the select applications, we analyzed their Google Play[7] reviews, which can provide a general opinion from the users about them [13]. The perceived value of reviews on the Web is uncontested and constitute an attractive area for exploration [10]. Sentiment analysis systems are being applied in almost every business and social domain because opinions are central to nearly all human activities and are key influencers of our behaviors [15]. For example, consumer surveys show people cite product reviews as a top influencer in purchase decisions [19].

[6] http://az-screen-recorder.br.uptodown.com/android.

[7] The official app store for the Android operating system - https://play.google.com/store.

Unfortunately, reading through the massive amounts of product reviews available online from many e-communities, forums, and newsgroups is not only a tedious task but also an impractical one [31]. So, to analyze by classifying sentiment polarity of reviews at the document level is a consolidated solution to solve this problem and it has been used in some research, such as [13,23,29].

Sentiment analysis is a language processing task that uses a computational approach to identify opinionated content and categorize it as positive or negative. Also, it tries to define the expressions of opinion and mood of writers [10], and it can analyze evaluations, attitudes, and emotions from written language [15].

There are many tools, application programming interfaces (API's) and services that provide sentiment analysis, as SentiWordNet [2], AlchemyAPI [30], PowerReviews[8], or BuzzMetrics[9]. Saif et al. [26] evaluated popular tools for entity extraction and concept identification, and, based on this information, we chose the AlchemyAPI to apply.

AlchemyLanguage is a collection of natural language processing APIs that help you understand the sentiment, keywords, entities, high-level concepts and more [30]. Natural language processing APIs available through AlchemyLanguage add a high-level semantic information, and it is capable of performing: Entity Extraction, Sentiment Analysis, Emotion Analysis, Keyword Extraction, Concept Tagging, Relation Extraction, Taxonomy Classification, among others.

We developed a system to automatized the Sentiment Analysis by using the AlchemyAPI: SANGRIa (Sentimental ANalysis Googleplay RevIews)[10]. Data collected by SANGRIa help us understand better the user satisfaction.

2.3 Third Phase - To Improve

The goal of this phase is to propose improvements that can be adopted by the applications to enhance the user experience. Like the previous phases, we have defined questions to help guide this step:

1. What kind of improvements can be incorporated to enhance the user experience?
2. What is the opinion of users about the improvements?

To help answer these questions, we performed a brainstorming [12] session with the same team of experts who participated in the heuristic and user evaluations. The researchers conducted each of two sessions for approximately two hours. Sessions began with a recap of the biggest problems of each application, followed by discussions and sketching proposals for solutions to each problem. As a result, the meetings provided a set of improvements for each application. Then, we apply modifications in the application interfaces.

[8] PowerReviews - http://www.powerreviews.com/.

[9] BuzzMetrics - http://buzzmetrics.com/.

[10] The code of SANDRIa is available in https://github.com/GREatPesquisa/SANGRIa.

In order to validate the modifications, we developed an online questionnaire[11] showing them for the same users who participated in the evaluation. "Waze" questionnaire had seven questions and "Meu Ônibus" had nine questions. Questions were about the found issues and the improvements suggested. Therefore, for each issue was pointed out a solution. In this way, the user had to score on a five-point Likert Scale if that solution was meaningful to the user experience or not. According to the users' answers, it was possible to map which solutions best suited the users' needs.

3 Results

The following subsections describe the primary results from each method used.

3.1 First Phase - To Understand

The survey reached 17 Brazilian states, and Ceará was the state with the most replies counted (81%). The survey obtained 345 respondents, 59% male respondents and 41% female respondents. The main results achieved by this method was the definition of the user profiles, as shown in Fig. 2, and the identification of two most used applications.

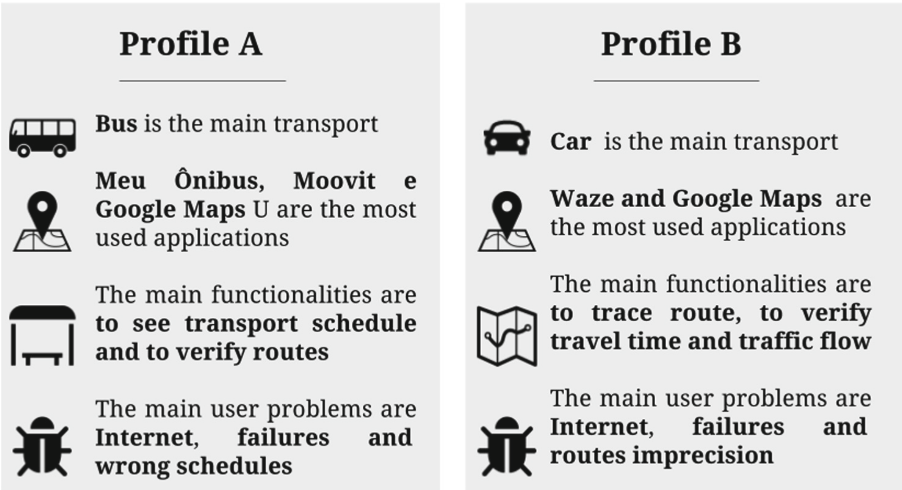

Fig. 2. Profiles A and B

The two profiles have the following characteristics in common: the user is in the 21–30 age group, does not feel safe using his means of transport, lives in a large city that has a vast network of public transportation and heavy traffic.

[11] The following link provides a copy of surveys used in the third phase: https://github.com/GREatPesquisa/2017-Research-Urban-Mobility-Applications.

The main difference between the profiles is in the means of transportation used in everyday life. Profile A uses car and profile B uses bus.

The user profiles and the ranking of the applications in the questionnaire guided the selection of the applications. The most chosen application to help bus mobility was "Meu Ônibus" (15%), selected to Profile A. To a particular transport, the most chosen was "Waze" (52%), selected to Profile B.

In "Meu Ônibus", the main features used are: see transport schedules; verify the route of public transport; check which transport lines pass through a certain place; and check stops of transport. The main features used in "Waze" application are to trace a route; search by address; check travel time between places, and check the flow of real-time traffic. Thus, these features guided the activities of heuristics and user evaluation in both applications of each profile.

Additionally, the questionnaire identified the main problems faced by each user profile. The main problems encountered by profile A are the Internet connection, application crashes and wrong schedules. The main problems faced by Profile B are route updating and accuracy, Internet connection, in safe routes and application crashes. Thus, this information was useful for evaluators to gain an understanding about the difficulties people face during interaction with the selected applications.

3.2 Second Phase - To Evaluate

Heuristic Evaluation. The heuristic evaluation was performed in the two applications chosen by the previous phase ("Meu Ônibus" and "Waze"). In this way, the presentation of the results is organized by evaluated applications. First, the results of "Meu Ônibus" are displayed and then those from "Waze".

Meu Ônibus

33 usability problems were found in total. The graph of Fig. 3 presents the distribution of the problems by severity and by heuristics. A problem can violate more than one heuristic. Thus the sum of the problems of each heuristic exceeds the total value of problems found for "Meu Ônibus".

From these problems, there are few catastrophic problems (four), but there are a considerable number of severe and simple severity problems (ten and ten, respectively).

Figure 3 shows all heuristics were violated, except Heuristic 6 (Recognition), since there were no identified situations in which users have to memorize information from one part of the interface to another part. The most problematic heuristic was *Visibility*, which shows a recurring problem with feedback. Figure 3 illustrates one of the problems encountered about this. This problem happens because map feedback messages to inform point locations are very fast, if the user is walking or distracted, he/she will not see the message.

Waze

In the evaluation of the "Waze" application, we consolidated fifteen usability problems. The graph of Fig. 4 shows the distribution of these problems by severity and by heuristic. From these problems, 2 are catastrophic, 7 are major, 6 are minor, and only 1 is cosmetic.

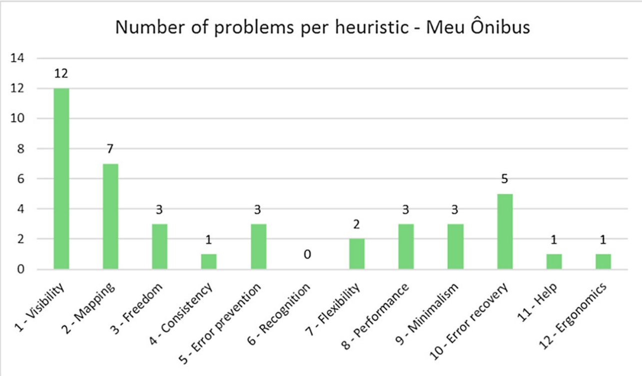

Fig. 3. Number of Problems per Heuristic - "Meu Ônibus"

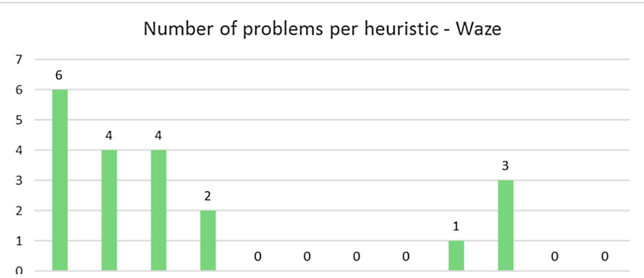

Fig. 4. Amount of problems per Heurístic - "Waze"

The heuristics Correspondence and Real World presented the most problems. For example, when the user adds a stop location between the beginning and the destination, displaying next to the final destination the expression "ETA". This expression is not clear, since it is not commonly used by to Brazilian users. "ETA" is English acronym for the expression "Estimated Time of Arrival'.

User Evaluation. The presentation of the user evaluation results is organized by applications. In this way, the results of "Meu Ônibus" are first described and then those of "Waze".

Meu Ônibus

The evaluation with all users of "Meu Ônibus" identified 28 interaction and interface problems. After eliminating duplicated problems, we consolidated 13 different problems. From the four activities performed in the evaluation, Activity 3 (to find bus stops) was the most problematic, since no user could perform it because they even could not find the feature. The activity that presented the highest completeness by the users was Activity 1 (to view transport schedules), which deals with the basic feature of the application, which is to check the time of transport. The feature checks how long the bus arrives at the user's current bus stop.

Waze
The evaluation with all users of "Waze" identified 30 interaction and interface problems. After eliminating duplicated problems, it found 11 different problems. From the five activities performed in the evaluation, Activity 3 (to verify travel time between location) was the most problematic since the users have difficult to find and understand how doing this in "Waze". The activities that presented the highest completeness by the users were Activity 4 (to create a route) and Activity 5 (to follow the route real-time). The completeness by activities four and five is good because they are one of main features of "Waze".

Sentiment Analysis. The sentiment analysis in the GooglePlay reviews has the focus on distinguishing between statements of fact vs. opinion, and on detecting the polarity of sentiments being expressed [13]. Sentiment analysis was used at a document level to classify whether a whole opinion document expresses a positive or negative sentiment [15]. This level of analysis assumes that each document expresses opinions on a single entity.

So each review has a score between -1.0 to 1.0, that means the sentiment rate indicates the sentiment polarity negative (<0.0), neutral(=0.0) or positive(>0.0). The API also provides if the review is a "mixed sentiment", which indicates that the sentiment is both positive and negative. To complement the sentiment results, we use AlchemyAPI to get the score to following emotions: joy, sadness, anger, disgust and fear. Each emotion has a score between 0.0 to 1.0. The sentiment analysis supports the Portuguese, but the emotion analysis does not. So, to apply the emotion analysis, it was necessary to translate to English by using a service GoogleTranslate[12].

Since the GooglePlay Reviews are available only for your owners, all the reviews were getting manually, what limited our data acquisition. GooglePlay platform classifies the reviews to show in three orders: newest, rating and helpfulness.

The sentiment analysis data results are an estimate, and this approach could have errors due to bad translations, the presence of unrecognized characters, like keyboard shortcuts for emoticons, swear words, popular terms not recognized to the dictionaries, even grammatical problems of writing. The last one frequently occurs in the reviews, which hinders a more accurate result. Besides having available services on the Internet that could insert artificial positive reviews to an application achieve a high level of the rank, known as Opinion Spam [15], e.g., ReviewApp4u[13]. So, the results are considering as complementary to others applied in the next phases.

Meu Ônibus
Figure 5 shows the results of "Meu Ônibus" application based on 189 reviews got in a three-month period classified by newest reviews. The sentiment rate found in the results was balanced with 38% negative reviews and 40% positive reviews. Comparing the emotions found in the reviews, sadness (0.28) and joy (0.34)

[12] Google Translate - https://translate.google.com.br.
[13] http://www.reviewapp4u.com.

highlighted, also balanced. Given these results, the general sentiment about this application is neutral, in which neither the positive sentiments, and emotions, nor the negative stand out.

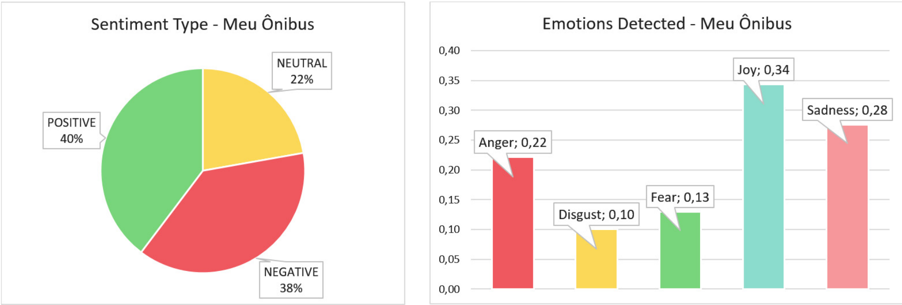

Fig. 5. Sentiment analysis of "Meu Ônibus"

Waze

Figure 6 shows the results of "Waze" application based on 293 reviews got in one-month period classified by helpfulness reviews. We chose these parameters to get the reviews due to a lot of reviews to "Waze" in GooglePlay platform every day. The sentiment rate shows the negative reviews (34%) are substantially over the positive reviews (24%). Comparing the emotions found, anger (0.29) and sadness (0.36) highlighted. As a result, the sentiment and emotions about this application are more negative than positive, different from "Meu Ônibus" application.

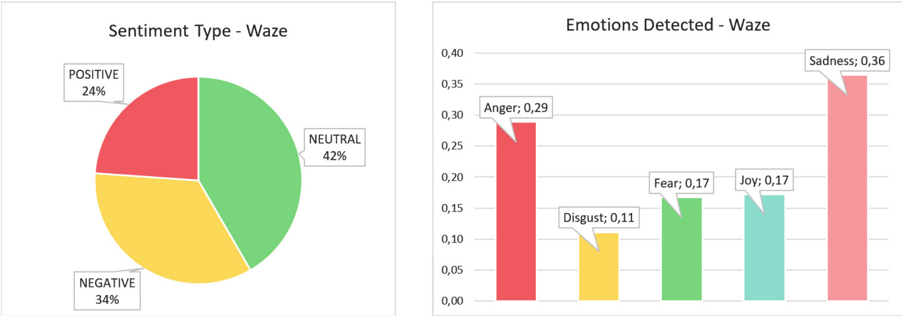

Fig. 6. Sentiment analysis of "Waze"

3.3 Third Phase - To Improve

As a result of brainstorming sessions, we planned improvements in the interface of each application. In the "Meu Ônibus" application, we made the following changes:

- Modifications in the main menu of the application to leave the features clearer, for example, the term "adjustments", which increases or decreases the search area, was modified to "fit in the search radius";

– Allow the user to search the itinerary of a bus or favorite line;
– Display the direction of the bus line clearly in the application to facilitate the user's understanding of this information;
– Informing how the user can visualize the stopping points of a bus line;
– Leave the search field always enabled when starting the application to facilitate the search activity of the application;
– Change the search area by holding and dragging interaction;
– And finally, improve the interaction with the application support.

All users who participated of "Meu Ônibus" user evaluation in phase two answered the survey applied to confirm the improvements. 60% of users evaluated the set of improvements as positive for the use experience, of which 40% considering the improvements would significantly modify the user experience. In this way, all the improvements were evaluated as positive and relevant to improve the user experience. The change in the menu labels, as shown in Fig. 7, was the less friendly item in the user evaluation. About this improvement, 20% thought it would not modify the user experience, 20% found a neutral modification and 60% scored as significant changing the user experience.

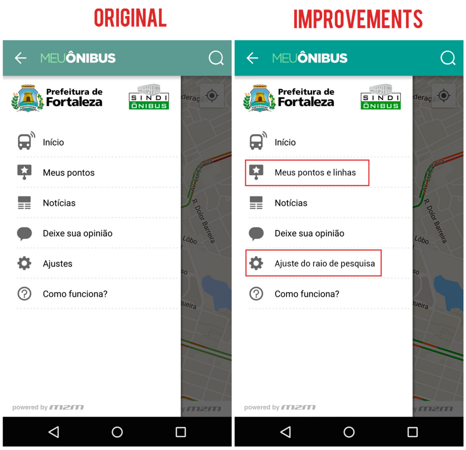

Fig. 7. Proposal of new menu for "Meu Ônibus"

The modification most well evaluated by the users was the search field enabled upon entering the application as showed in Fig. 8. Some users in user evaluation, phase two, were not aware of the application's search feature.

Improvements to "Waze" are related to:

– The presence of foreign terms misunderstood by users.
– The search result and the planned route overview in the application (Fig. 9).
– The difficult access of some features such as the conversion list.
– The way the application instructs users on the route.
– The possibility of signaling areas considered unsafe due to the occurrence of robberies.

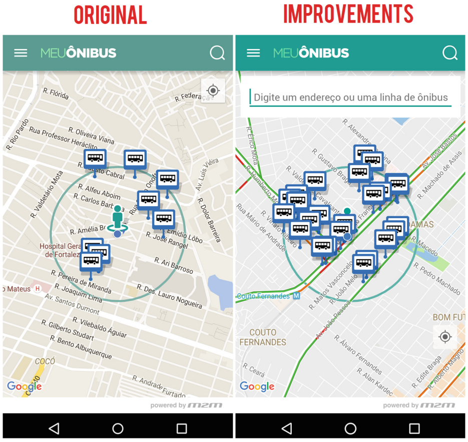

Fig. 8. Welcome screen proposal of "Meu Ônibus"

In the case of safety-related improvement of suggested routes, by flagging an area as unsafe, the application would avoid suggesting such parts on the route, or it would signal the dangerousness to users' pay attention as they navigate the stretch.

All users in "Waze" user evaluation participated in the evaluation of the improvement proposals. 60% of users evaluated the set of improvements as significant changes from users, 20% rated as good changes to user experience, and the other 20% believe modifications would not change the user experience. The most significant improvement for users was security-related (80%) as showed in Fig. 10. Modifying the search result to see the entire planned route on the map and the reduction of the steps was considered a positive change by 40% of users and a significant change by 40%, of which 20% believes that it would not improve the experience of use.

Users positively evaluated the rest of the improvements. The set of improvements was praised by users and considered relevant if the application implements them. The modifications vary in degrees of complexity, but by the users' willingness the security-related improvement, if implemented, would significantly transform the user experience. Even though the application can not guarantee the user that the suggested routes are safer, the users would feel more secure when using the application. It would slightly alleviate the tension of the users when they move between places in the city, improving their day to day.

4 Discussion

Mobility is a daily reality to great and medium-sized cities. Providing means of transportation that assist users in their most diverse destinations is a social duty of government and society. Technologies of urban mobility have the objective of helping the population to have a better displacement between places. Among

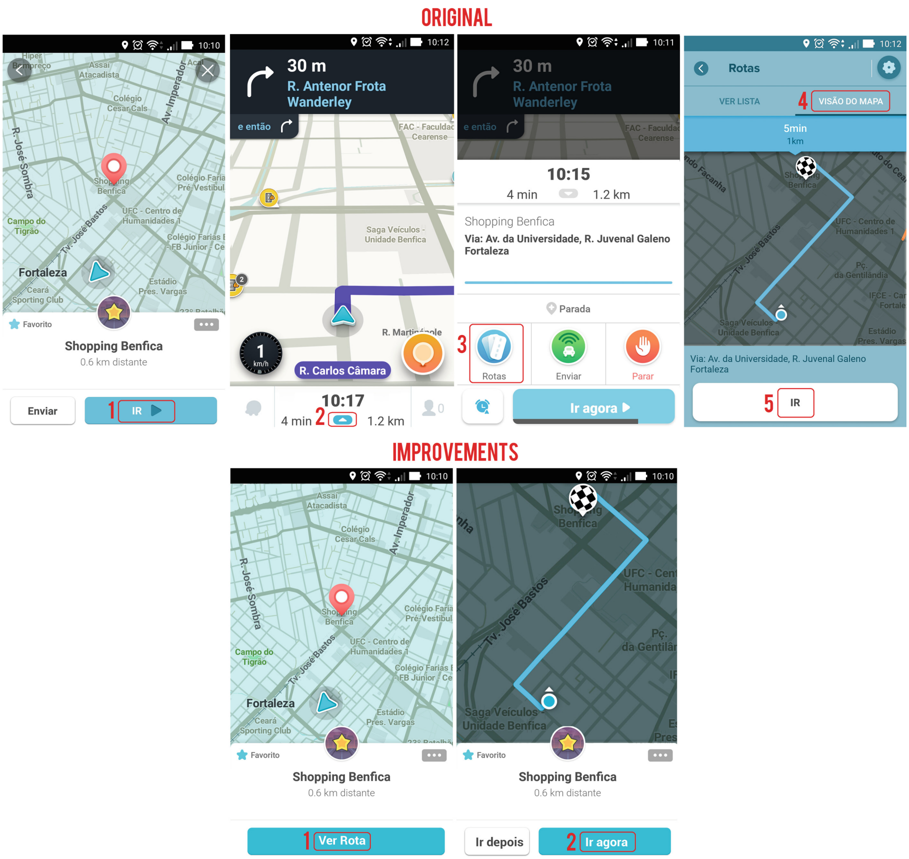

Fig. 9. New way to interact with "Waze" search result

Fig. 10. Security enhancements for "Waze"

these technologies, urban mobility applications have emerged as a resource for improving and optimizing users' lives.

Urban mobility applications need to address a multitude of information to suit to the reality lived by users. Thus, providing accurate, reliable and secure route and address information significantly influences user loyalty. Even though urban mobility applications are suited to the needs of the population, there are factors beyond the scope of applications.

Through the methodology and results of the evaluation, we identified four recommendations for urban mobility applications in the Brazilian context. The recommendation descriptions are in Table 1.

The recommendations are based on recurring principles of usability and human-computer interaction, and also on recurring themes identified in all phases of the methodology. In the third phase of the methodology, it was possible to perceive that some problems faced by the users are solved through the execution of simple ideas. In this way, providing applications that support urban mobility, which perform simple activities and always maintain a clear and objective *feedback* (recommendation 1 in Table 1) is a difficult but possible task. Such difficulty is set up due to the large amount of information that is needed to meet users' needs. The use of maps and detailed information on routes and places supplied the user needs but constitutes an interface overloaded with information. These features typically require interactions that can be cumbersome with elements with a large amount of information, such as zoom interactions and clicks on the map. To maintain a consistent informational level with the objective of the user in the interface is necessary. The improvements proposed mostly had the focus of minimizing the informational burden and transmitting it in a clear manner, do not keeping the information that lead users to unnecessary or incorrect interactions.

Another issue that significantly affected the user experience was the lack of synchronicity between the *information provided by the system with the reality lived by the users* (recommendation 2 in Table 1). Bad routes, wrong schedules, and false information in the context of urban mobility have a profound impact on the user's daily life.

A major problem faced by users of mobile applications is dependence on the Internet connection. Although extremely necessary, the quality of the connection is not always good to offer a good user experience. As a result, applications should *minimize network dependency* (recommendation 3 in Table 1), allowing in a short amount of time with the connection the user receive the most relevant information for the user context. In addition to connection, another essential component for systems that support mobility is GPS, that, attached to the connection, allows the location of addresses and the elaboration of routes. Problems with GPS and connection need to be treated in a proper and distinct way so that user actions are efficient and accurate.

When proposing an application of urban mobility, the social aspects (recommendation 4 in Table 1) related to the context of user use should receive attention. In our research, safety is the most important aspect. In other countries,

Table 1. Recommendations for urban mobility applications in Brazil

Recommendations	Description	Examples
Feedback	Pay attention to provide a clear feedback and keep the user informed about the actions made by the application	When using maps or providing routes to users, make clear the information for the means of transport that the user will use. By car name the streets, an overview of the route. If it is Bus, line name, a route to be traversed, stop points and landmarks if they exist
Reliability	Ensure the reliability of the information provided to users	Do not provide incorrect time and wrong information in real time. If there are unforeseen problems with public transport, prepare a way to notify them about it
Connection	Project well the dependency the connection of app	Provide off-line services to users such as routing and transport schedules. See intelligent ways of providing services in possible connection losses
User context	Consider factors related to the context of use and social aspects suggesting, itineraries and routes safety for users	Try to protect the user from unwanted situations such as dangerous areas or unwanted destinations resulting from the context of use

perhaps other social aspects are more relevant than safety. Applications need to be aware of the social context in which they operate.

Regarding the applications evaluated in phases 2 e 3 of the methodology, "Meu Ônibus" presented serious problems of *feedback* and visibility of the system's features, even that the sentiment analysis showing a neutral sentiment, maybe due to the importance and to the usefulness of the application in your daily activities. The application tends to not leave the user satisfied with their user experience. The problems of "Meu Ônibus" make day to day life of users more difficult. Thus, through phase 3 of the methodology, the simple ideas adoption could solve some problems. Such improvements have had a very positive reception by users. In addition to being mostly simple improvements, they are mostly functions that the application already presents, but difficult access or understanding to users. In this way, if "Meu Ônibus" adopts some of the evaluated solutions the users would better understand the dynamics of use of the application, providing a better user experience and greater loyalty in the utilization of the application by the users.

In the evaluation of "Waze", we encountered that external factors to the application have a profound impact on the user experience, also presented in the sentiment analysis, that showed a negative user perception. The application needs to handle with factors such as the drop in connection, lack of GPS synchronization, and *social factors* in the best way for the users. Security-related improvement was the most highly valued and most significant improvement for users. This result reinforces the users' concern about the issue. Besides security, the other improvements presented an optimization in the activities and features treated in the scope of the realized improvements. Improvements have some variations of complexity, but even the simple ones are essential for the user experience.

5 Conclusions

Urban mobility in Brazil is a big issue that needs a lot of discussion with users to reach a good evaluation. However, the evolution of mobile technologies helps to minimize some of the urban mobility problems. For example, there are several mobile applications to various types of transport (*e.g.,* Google Maps, Waze, Uber, Meu Ônibus, Moovit, Bicicletar, Strava) avaliable.

This work did an investigation into the experience of using urban mobility applications in a large Brazilian city. This work complements the results presented in the evaluation competition of the XV Brazilian Symposium on Human Factors in Computational (IHC 2016) [1]. We added in this paper more methods of evaluation (sentiment analysis and assessment of the improvements proposed by the HCI experts) that makes possible to deepen aspects related to user satisfaction and to visualize some solutions of the problems faced by the users. The adopted methodology provided a real insight into the context of the urban mobility experience by the users, their main problems and what kind of solutions could be adopted to improve the use experience of the applications. Thus, from this methodology, the results of the research allowed tracing two profiles of users that move around using mobile applications like "Waze" and "Meu Ônibus", which are the ones chosen to be evaluated in our work.

In the "Meu Ônibus" application, we found 33 usability problems in the heuristic evaluation, and the user evaluation found 13 problems. The problems encountered by these two methods mainly concern feedback. Failures with the Internet connection or user location were recurring, and these are not reported in an appropriate manner to the user or within a reasonable time. A set of improvements has been proposed by some of the experts in the user experience and also evaluated by the users. These improvements varied in degree of complexity and were well accepted by the users with the productive potential to improve the user experience of "Meu Ônibus".

"Waze" presented 15 usability problems in the heuristic evaluation and the user evaluation presented 11 problems. Results of the assessment with the users were more worrisome, because they identified external factors that have a profound impact on the dynamics of the application's functionalities and the user experience. Problems such as dangerous routes have been detrimental to the safety of the user. This scenario runs against to one of the principles of the national urban mobility policy. "Waze"'s security enhancement has been rated best by users. If "Waze" presented a similar feature, users would feel more confident and safety when moving around the city and using the app.

As a future work, we intend to evaluate more urban mobility applications and identify more issues to be explored. Also, we intend to conduct a new survey and recognize more user profiles for other means of transportation such as subway, taxi and bicycle. Based on that, further enrichment of the recommendations for urban mobility applications is necessary.

Acknowledgments. We would like to thank the CTQS/GREat team for the technical support for this work and also the "CAcTUS - ContextAwareness Testing for Ubiquitous Systems" project, supported by CNPq (MCT/CNPq 14/2013 - Universal) under grant number 484380/2013-3.

References

1. Almeida, R.L.A., Mesquita, L.B.M., Carvalho, R.M., Junior, B.R.A., Andrade, R.M.C.: Quando a tecnologia apoia a mobilidade urbana: Uma avaliação sobre a experiência do usuário com aplicações móveis. In: Proceedings of the XV Brazilian Symposium on Human Factors in Computer Systems (IHC 2016). Sociedade brasileira de Computação - SBC, Porto Alegre, Brazil (2016, in Portuguese)
2. Baccianella, S., Esuli, A., Sebastiani, F.: SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: LREC, vol. 10, pp. 2200–2204 (2010)
3. Behr, A., Corso, K.B., Nascimento, L.F.M.D., Freitas, H.M.R.D.: Mobilidade urbana sustentável e o uso de tecnologias de informação móveis e sem fio: em busca de alternativas para a cidade de porto alegre/rs. Gestão Contemporânea [recurso eletrônico]. Porto Alegre **10**(14), 61–90 (2013)
4. Bollini, L., De Palma, R., Nota, R., Pietra, R.: User experience & usability for mobile geo-referenced apps. A case study applied to cultural heritage field. In: Murgante, B., et al. (eds.) ICCSA 2014. LNCS, vol. 8580, pp. 652–662. Springer, Heidelberg (2014). doi:10.1007/978-3-319-09129-7_47
5. Carvalho, C.M., Rodrigues, C.A., Aguilar, P.A., de Castro, M.F., Andrade, R.M.C., Boudy, J., Istrate, D.: Adaptive tracking model in the framework of medical nursing home using infrared sensors. In: 2015 IEEE Globecom Workshops (GC Wkshps), pp. 1–6. IEEE (2015)
6. Carvalho, C.H.R.D.: Desafios da mobilidade urbana no Brasil. Instituto de Pesquisa Econômica Aplicada (IPEA) (2016)
7. Carvalho, R.M., Santos, I.S., Meira, R.G., Aguilar, P.A., Andrade, R.M.C.: Machine learning and location fingerprinting to improve UX in a ubiquitous application. In: Streitz, N., Markopoulos, P. (eds.) DAPI 2016. LNCS, vol. 9749, pp. 168–179. Springer, Cham (2016). doi:10.1007/978-3-319-39862-4_16
8. Carvalho, R.M., Andrade, R.M.C., Oliveira, K.M., Sousa Santos, I., Bezerra, C.I.M.: Quality characteristics and measures for human-computer interaction evaluation in ubiquitous systems. Softw. Qual. J. **24**, 1–53 (2016)

9. Gabrielli, S., Forbes, P., Jylhä, A., Wells, S., Sirén, M., Hemminki, S., Nurmi, P., Maimone, R., Masthoff, J., Jacucci, G.: Design challenges in motivating change for sustainable urban mobility. Comput. Hum. Behav. **41**, 416–423 (2014). http://www.sciencedirect.com/science/article/pii/S0747563214003045
10. Ghorbel, H., Jacot, D.: Sentiment analysis of french movie reviews. In: Pallotta, V., Soro, A., Vargiu, E. (eds.) Advances in DART. SCI, vol. 361, pp. 97–108. Springer, Heidelberg (2011). doi:10.1007/978-3-642-21384-7_7
11. Inostroza, R., Rusu, C., Roncagliolo, S., Rusu, V., Collazos, C.A.: Developing smash: a set of smartphone's usability heuristics. Comput. Stand. Interfaces **43**, 40–52 (2016)
12. Ivanov, A., Cyr, D.: Satisfaction with outcome and process from web-based meetings for idea generation and selection: the roles of instrumentality, enjoyment, and interface design. Telematics Inform. **31**(4), 543–558 (2014)
13. Kim, S.M., Pantel, P., Chklovski, T., Pennacchiotti, M.: Automatically assessing review helpfulness. In: Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, EMNLP 2006, pp. 423–430. Association for Computational Linguistics, Stroudsburg (2006). http://dl.acm.org/citation.cfm?id=1610075.1610135
14. Leite, D.F.B., Rocha, J.H., Batista, C.D.S.: Busão: um sistema de informações móvel para auxílio à mobilidade urbana através do uso de transporte coletivo. IX Simpósio Brasileiro de Sistemas de Informação, pp. 170–181 (2013)
15. Liu, B.: Sentiment analysis and opinion mining. Synth. Lect. Hum. Lang. Technol. **5**(1), 1–167 (2012)
16. Maia, M.E.F., Fonteles, A., Neto, B., Gadelha, R., Viana, W., Andrade, R.M.C.: LOCCAM - loosely coupled context acquisition middleware. In: Proceedings of the 28th Annual ACM Symposium on Applied Computing, SAC 2013, pp. 534–541, ACM, New York (2013). http://doi.acm.org/10.1145/2480362.2480465
17. Maia, M.E., Andrade, R.M., de Queiroz Filho, C.A., Braga, R.B., Aguiar, S., Mateus, B.G., Nogueira, R., Toorn, F.: Usable-a communication framework for ubiquitous systems. In: 2014 IEEE 28th International Conference on Advanced Information Networking and Applications, pp. 81–88. IEEE (2014)
18. Marx, R., de Mello, A.M., Zilbovicius, M., de Lara, F.F.: Spatial contexts and firm strategies: applying the multilevel perspective to sustainable urban mobility transitions in Brazil. J. Cleaner Prod. **108**, 1092–1104 (2015)
19. McGlohon, M., Glance, N.S., Reiter, Z.: Star quality: aggregating reviews to rank products and merchants. In: ICWSM (2010)
20. Nielsen, J.: Usability Engineering. Elsevier, Amsterdam (1994)
21. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 249–256. ACM (1990)
22. Nielsen, J., Clemmensen, T., Yssing, C.: Getting access to what goes on in people's heads? Rreflections on the think-aloud technique. In: Proceedings of the Second Nordic Conference on Human-Computer Interaction, pp. 101–110. ACM (2002)
23. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? Sentiment classification using machine learning techniques. In: Proceedings of the ACL 2002 Conference on Empirical Methods in Natural Language Processing - Volume 10, EMNLP 2002, pp. 79–86. Association for Computational Linguistics, Stroudsburg (2002). https://doi.org/10.3115/1118693.1118704
24. Pupi, S., Pietro, G., Aliprandi, C.: Ent-it-UP. In: Stephanidis, C. (ed.) HCI 2014. CCIS, vol. 435, pp. 3–8. Springer, Cham (2014). doi:10.1007/978-3-319-07854-0_1

25. Rogers, Y., Sharp, H., Preece, J., Tepper, M.: Interaction design: beyond human-computer interaction. netWorker: Craft Netw. Comput. **11**(4), 34 (2007)
26. Saif, H., He, Y., Alani, H.: Semantic sentiment analysis of Twitter. In: Cudré-Mauroux, P., et al. (eds.) ISWC 2012. LNCS, vol. 7649, pp. 508–524. Springer, Heidelberg (2012). doi:10.1007/978-3-642-35176-1_32
27. Santos, R.M., Oliveira, K.M., Andrade, R.M.C., Santos, I.S., Lima, E.R.: A quality model for human-computer interaction evaluation in ubiquitous systems. In: Collazos, C., Liborio, A., Rusu, C. (eds.) CLIHC 2013. LNCS, vol. 8278, pp. 63–70. Springer, Cham (2013). doi:10.1007/978-3-319-03068-5_13
28. Silva Junior, D.P., Souza, P.C., Maciel, C.: Establishing guidelines for user quality of experience in ubiquitous systems. In: Streitz, N., Markopoulos, P. (eds.) DAPI 2016. LNCS, vol. 9749, pp. 46–57. Springer, Cham (2016). doi:10.1007/978-3-319-39862-4_5
29. Turney, P.D.: Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, ACL 2002, pp. 417–424. Association for Computational Linguistics, Stroudsburg (2002). http://dx.doi.org/10.3115/1073083.1073153
30. Watson, I.: Alchemyapi. https://www.ibm.com/watson/alchemy-api.html. Accessed 02 Feb 2017
31. Zhang, R., Tran, T.T.: Helping E-commerce consumers make good purchase decisions: a user reviews-based approach. In: Babin, G., Kropf, P., Weiss, M. (eds.) MCETECH 2009. LNBIP, vol. 26, pp. 1–11. Springer, Heidelberg (2009). doi:10.1007/978-3-642-01187-0_1