Social networks and collective intelligence applied to public transportation systems: A survey

Ana Paula Chaves  
Coordination of Technology Systems for Internet Federal Technological University of Paraná (UTFPR) Campo Mourão - PR anachaves@utfpr.edu.br

Igor Steinmacher  
Coordination of Technology Systems for Internet Federal Technological University of Paraná (UTFPR) Campo Mourão - PR igorfs@utfpr.edu.br

Vaninha Vieira  
Department of Computer Science Federal University of Bahia (UFBA) Salvador - BA vaninha@ufba.br

ABSTRACT
Road traffic in big cities is becoming increasingly chaotic. The use of public transportation is an alternative to improve this scenario, since it diminishes the number of private vehicles on the roads. However, to improve its quality is important to understand the problems people are facing when using this kind of transportation. This paper presents the results of a survey conducted with public transportation users which investigates how collaborative systems based on social networks and collective intelligence can support sharing information with the passengers. The results show that there is a scarcity of ways to obtain real-time information related to public transportation and that the use of social network applications and collective intelligence is an interesting way to share and obtain this kind of information.

Keywords
Intelligent Transportation System, collective intelligence, social network

1. INTRODUCTION
Big cities are increasingly suffering from the problem of urban mobility. Several problems, such as constant traffic jam, lack of investments in road infrastructure, and poor quality of the services provided by public transportation makes the citizens daily life stressful and complicated. Estimulating the use of public transportation is essential to support urban mobility, because it reduces the number of cars on public roads [5]. Besides, since Brazil will be hosting important sport events in the coming years, it is important to investigate how information technology can be applied to public transportation to better support users’ needs and decisions.

According to [3], the use of information technology for planning, managing, operating and monitoring public transportation has become a feasible alternative. Intelligent Transportation Systems (ITS) aims to investigate how to apply well established information and communication technologies to improve the quality of transportation systems [12]. A classification for ITS solutions, presented in [12], considers the area of Advanced Public Transportation Systems (APTS), which aims to propose solutions to support information provisioning to passengers, such as the current status and timetables of the public vehicles, the indication of the route lines and the stops attended by a line, incidents on public transportation with the indication of possible delays, and others. When notified of such information, passengers can make their decisions easily, planning their route and activities, which results in less stress and anxiety and more satisfaction with the public transportation services. According to a recent article in Wired magazine [1], people are willing to leave their cars or motorcycles, if they can get information in real time, about the public transportation.

The popularization of social media (such as blogs, social networks and wikis) changed the way people deal with their daily problems by increasingly using virtual communities to share and seek for information. As indicated in [10], it enables communities to solve more problems than it would be possible individually. Collective intelligence involves combining knowledge (which may include behaviors, preferences, ideas, and so on) provided by a group of people to produce novel information or insight [11]. Some applications [14, 6] that make use of collective intelligence capabilities were proposed to collect knowledge based on what is being discussed on the Internet.

In this context, we are investigating how collaborative systems, specially social media and collective intelligence, can increase the awareness of public transportation users reducing the lack of information related to the route conditions, the lines availability and the waiting time at bus stops. This research is part of a project called UbiBus [13] whose objective is to propose, specify and implement technological solutions to easy users’ access to public transportation information in real time, based on dynamic context-sensitive information, available on multiple devices. A module of UbiBus, called UbiBus-Net, aims to investigate and implement collective intelligence solutions and tools that analyze textual information produced by users on social networks about the public transportation conditions, allowing the access to the generated collective knowledge.
To support the requirements elicitation for UbiBus, we designed and conducted a survey with transportation users, in different Brazilian cities, through the application of a questionnaire. This survey aims to identify deficiencies and needs related to the information provided to public transportation passengers. The main advantage of conducting a field survey is that we can have the opinion of who actually use the public transportation and know the real services conditions in their city or region. Thus, we can develop applications closer to users’ expectations.

This paper presents the results obtained in this survey and analyzes the problem of providing information to public transportation users based on collective intelligence and social networks. To accomplish it, the paper is organized as follows. Section 2 presents some related work in the area of ITS and Collective Intelligence. Section 3 brings the methodology used to perform the survey. In Section 4, we discuss the results obtained with the survey, and in Section 5 we present our conclusions about the survey findings.

2. BACKGROUND

2.1 Collective Intelligence

Currently, collaboration among people has been gaining attention due to the popularity of social media such as forums, weblogs, wikis and social networks. These media are collaborative systems that enable users to share their knowledge, skills and other information. As a result, they are becoming important sources of collective intelligence. The science of collective intelligence, proposed from the discussions of Pierre Lévy [8], tries to harness the potential of social networks as a means to exercise the citizenship. It assumes that individual intelligences are summed and shared across society.

Collective intelligence led to the rise of a new business model known as crowdsourcing. The term, coined by Jeff Howe [7], describes a model that takes advantage of several creative solutions that people can propose. The idea of crowdsourcing is to send a task to the crowd instead of running it using their own resources.

There are some recent initiatives that make use of collective intelligence to benefit public or private efforts. WikiCrimes [6] enables people to share information about the occurrence of criminal acts in a particular location. Vivacqua and Borges [14] propose the usage of collective intelligence capabilities for emergency response systems. In some countries, there are initiatives to share information about corruption, as the sites “I paid a bribe” (India) and “No les votes” (Spain). The “Map of Corruption in Brazil” consists of a site hosted on Google Maps for people to provide and seek for information about corruption in Brazil. In Panama, the site “Mi Panama Transparente” uses collective intelligence to gather information about various aspects relevant to society such as corruption, crime and violence.

The increasing demand and use of collective intelligence applications indicates that people tend to collaborate in order to build a collective knowledge, since such knowledge is beneficial to themselves and the group they belong. Thus, collective intelligence principles can offer solutions to many sectors in the society, including the improvement of public transportation systems.

2.2 Advanced Public Transportation Systems

Advanced Public Transportation Systems (APTS) is a subarea of Intelligent Transportation Systems that aims to use technology to provide real time information to assist users of public transportation on planning their trips [12]. This information includes the types of available transportation, transportation’s estimated departure and arrival times, their itinerary, among others. A better accessibility to information concerning the public transportation in a city makes this option a more attractive mean of transportation to its potential users [4]. This interaction can occur by using different devices, such as information centers, displays inside a bus or at a bus station, the web, smartphones or interactive TV.

Several applications have been developed to assist users on being aware about public transportation means in a city (e.g. OneBusAway [5], Your City on Time [13], TOTRansit5, OnibusRecife6). Tools for mobile devices like OneBusAway7 and Waze8 provide information according to the user’s geographic location, achieved by GPS (Global Positioning System). Waze generates maps according to the route informed by the user, and presents traffic data collected in real time from other people who are also using the software [9]. OneBusAway is directed to bus users from the city of Seattle and vicinity, and presents all bus stops closer to the user’s location, the lines associated to each stop and their current schedules, updated in real time by the bus company. OneBusAway also provides a web interface that displays a map with the bus stops’ locations and the timetables for all vehicles associated to those stops.

Other available tools are based on the principle of visualizing information on maps, including: OnibusRecife which displays the bus stops closer to a given (typed) location in the city of Recife, Brazil, and the possible routes to a given destination; ToTransit displays the map location of streetcars in Toronto, Canada, with information provided, in real time, by the transportation companies; and Bus Maps London9, which shows static information about bus stops, routes and schedules for the city of London. Google Maps10 also presents APTS features for some cities, enabling people to plan their routes using public transportation system. Moreover, Google Maps APIs are used to compose most part of the applications mentioned on this section.

Despite those initiatives, it is difficult to obtain dynamic information about public transportation conditions, since many unexpected events may occur in the route (e.g. traffic

---

5http://totransit.ca/
6http://www.onibusrecife.com.br/
7http://www.onibusaway.org/
8http://www.waze.com/
9http://www.tfl.gov.uk/tfl/gettingaround/maps/buses/
10http://maps.google.com
jam, accidents or incidents). Most solutions also strongly depend on the information provided by the official transportation companies. A feasible alternative to obtain real-time information is to create an intelligent transportation system that makes use of the power of collective intelligence.

3. METHODOLOGY
The research presented in this paper was based on a questionnaire answered by public transportation users in Brazil. Results were analyzed to check if they could indicate some requirements to build an APTS to provide information to passengers and if the usage of social networks and collective intelligence could create better APTS. The following steps were performed:

- planning;
- conduction;
- data preparation; and
- data analysis and results.

During the planning activity, we defined the research hypotheses, which were used to formulate the questionnaire. The hypotheses are:

- **Hypothesis 1**: public transportation passengers have very few ways to get information about the lines they need to use, and when available, the information is incomplete.

- **Hypothesis 2**: passengers tend to share information about events that affect public transportation, so collaborative systems based on collective intelligence can be useful to improve the quantity and quality of available information.

- **Hypothesis 3**: passengers have more confidence on information provided by people in their social network (friends, acquaintances, or other ties) than on information achieved from unknown sources.

- **Hypothesis 4**: applications for mobile devices can be used to provide information to passengers, to ease information sharing about occurrences on public transportation at the time they happen and at any place the user is situated.

The questions in the questionnaire were divided in two main groups. The first one included twelve questions (Table 1) related to the user profile, the public transportation they use, and their knowledge about available information. The second group comprised nine questions (Table 2) related to the use of social networks by passengers, and the use of online applications to share information about public transportation. Most questions are multiple choice type and it was always possible for the respondents to point out “Other” as an alternative, in case none of the available options were adequate.

### Table 1: Group 1: You and the Public Transport

<table>
<thead>
<tr>
<th>Question</th>
<th>Response Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  City you live:</td>
<td></td>
</tr>
<tr>
<td>2  State you live:</td>
<td></td>
</tr>
<tr>
<td>3  How often do you use public transportation?</td>
<td></td>
</tr>
<tr>
<td>4  When you need to use public transportation, how do you get information about the time that the vehicle will be arriving at a given stop?</td>
<td></td>
</tr>
<tr>
<td>5  How often does the transportation arrive at the bus stop at the expected time?</td>
<td></td>
</tr>
<tr>
<td>6  When delays occur in the line you use, what kind of problems can be considered as causes of delay in most cases?</td>
<td></td>
</tr>
<tr>
<td>7  Suppose you are at the bus stop, waiting for the vehicle of a given line. If there is any delay, do you have a way to become aware of what happened, before the vehicle reaches the stop?</td>
<td></td>
</tr>
<tr>
<td>7.1 (If the answer to question 7 was “Yes”) How do you get information about the reasons for the delay?</td>
<td></td>
</tr>
<tr>
<td>7.2 (If the answer to question 7 was “No”) Do you consider important that this kind of information was available?</td>
<td></td>
</tr>
<tr>
<td>8  When you need to take a bus from a particular line for the first time, do you try to get information about this line?</td>
<td></td>
</tr>
<tr>
<td>8.1 (If the answer is to question 8 was “Yes”) What information sources do you usually consult?</td>
<td></td>
</tr>
<tr>
<td>8.2 (If the answer is to question 8 was “Yes”) What kind of information can you find?</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2: Group 2: You and the Social Networks

<table>
<thead>
<tr>
<th>Question</th>
<th>Response Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  Do you use Internet on your cellphone?</td>
<td></td>
</tr>
<tr>
<td>2  Do you use GPS on your cellphone?</td>
<td></td>
</tr>
<tr>
<td>3  Do you participate in social networks?</td>
<td></td>
</tr>
<tr>
<td>3.1 (If the answer to question 3 was “Yes”) In which social networks do you participate actively</td>
<td></td>
</tr>
<tr>
<td>4  Do you post events that occur in traffic on social networks?</td>
<td></td>
</tr>
<tr>
<td>4.1 (If the answer to question 4 was “Yes”) Do you wish there was an application to process and display the information posted on social networks?</td>
<td></td>
</tr>
<tr>
<td>4.2 (If the answer to question 4 was “No”) Would you post this kind of events if there was an application to facilitate this?</td>
<td></td>
</tr>
<tr>
<td>5  If there was an application to display real-time information about the bus lines you use, would you consult it?</td>
<td></td>
</tr>
<tr>
<td>6  Do you trust information about bus lines posted on the Internet?</td>
<td></td>
</tr>
</tbody>
</table>

The questionnaire was created using LimeSurvey 11 and made available electronically12. We conducted the survey from March 24th to April 29th, 2011, using web-based sampling methods. According to Caulfield and O’Mahony [2], this method consists on sending the questionnaire to a initial group of people with common attributes. The first group of respondents complete the survey and then pass the survey on to their network. The process repeats and gathers momentum as the number of responses increases and the snowball effect occurs [2]. For this survey, we sent the questionnaire URL via email and published it on social networks (Twitter13, Facebook14 and Orkut15), addressed to some public transportation users and communities. We requested them to answer the questionnaire and forward the URL to their contacts.

The data preparation activity was performed to categorize the responses and standardize the open ended questions.

12 [http://goo.gl/B3UJM](http://goo.gl/B3UJM)
13 [http://www.twitter.com](http://www.twitter.com)
14 [http://www.facebook.com](http://www.facebook.com)
15 [http://www.orkut.com](http://www.orkut.com)
(questions 1 and 2 of group 1), which had different spellings for the same answer. We have collected a total of 683 responses. During the data preparation, 13 responses were dismissed as they presented some kind of inconsistency, resulting in a total of 670 valid responses. The reasons of the exclusions are presented in Table 3.

The data analysis activity was performed to verify the hypotheses. The results of this activity will be described in next section.

4. DATA ANALYSIS AND RESULTS

To perform the data analysis activity, we considered the 670 valid responses, involving 94 cities in 21 different Brazilian states. Six states appeared in the results with a single response (Acre, Amapá, Espírito Santo, Pará, Pernambuco and Piauí). States that had the most significant number of responses were Bahia (140 responses, 132 from the city of Salvador), Paraná (128), São Paulo (121 responses, 83 from the state capital) and Rio de Janeiro (66 responses, 57 from the state capital). The distribution of responses per state can be found in Table 4.

When asked about how often they used public transportation, 61% of respondents said they use every day and 20% reported that they use few days a week. Only 19% said they occasionally use public transportation (Figure 1).

4.1 Analysis of Hypothesis 1

Regarding Hypothesis 1, we wanted to know if the public transportation passengers have means to get information about the lines they use. In addition, we intended to verify, when information are available, if they are sufficient and complete.

We asked respondents how they become aware of the time a vehicle will arrive at a given stop/station. It was possible to check more than one choice among: (i) via Internet, (ii) by telephone, (iii) printed timetables or guides, (iv) asking other people, (v) there is no way to be aware, I go to the stop and wait for the vehicle, and (vi) other. The number of responses for each option is shown in Figure 2.

Despite the apparent large number of people who reported “There is no way to be aware” (47%), only 202 people (30% of total) marked only this option. The remaining 114 respondents checked this option in conjunction with other(s). This may occur due to the fact that the information is not always available in one way or another. Thus, when the information is not found, the passenger go straight to the stop/station. The other 468 respondents (70%) reported they have some way to obtain this information.

Although people can get the timetables for the buses associated to a given stop, when we asked about the punctuality of vehicles (Figure 3) only 32 respondents (5%) said that the transportation always comes on time. Out of these 32, only seven people live in capitals. We believe that the countryside suffer less with the phenomena that cause delays, such as accidents and traffic jams, due to the smaller number of vehicles on the streets. Other 240 people said that the bus rarely or never arrives at the stop on time. In addition, 180 people said they do not know if the bus is delayed because they are not aware of the time it was expected to arrive. This shows that information provided on the timetables are not enough to make the passenger aware of the time the vehicle should arrive at the stop.

We asked respondents what were the problems that often cause delays in vehicles (Figure 4). The answers showed that the major issues are traffic jams (506), car accidents (411) and robbery (200). Flooding were reported 121 times and...
29 other people pointed to other security incidents. 106 people reported that other (not listed) problems cause delays. Among the cited problems, 33 refer to mechanical problems, 24 to uncommitted drivers, 17 to transportation companies planning failures and 11 to the lack of transportation vehicles.

As it can be noted, some of the problems can be solved by administrative measures, such as planning issues and lack of vehicles. However, most of them are unexpected events that occur along the route and, therefore, should be communicated to passengers who are waiting for the transportation. To find out if this information is available, we asked the respondents if they know ways to become aware of a delay or event that cause a delay before the vehicle reaches the stop (Figure 5). 648 people (97%) said “No”. We then asked these people if they consider it important to have this information and 633 said “Yes”. It shows that the available information about public transportation are generally based on estimations and do not represent the current real-time information. Furthermore, it shows that users require more accurate information regarding public transportation.

Out of 670 responses, 510 said that they seek for information on the their first time usage. Among them, 289 said they can find the departure time from the line origin and 279 answered they can check the location of the intermediate stops. 139 people said they can find the arrival time to the final destination. The other alternatives were the timetables for intermediate stops (64 answers), statistics of delays and other events (22 answers), and other information (33 answers). Among other information, the most cited was the line route or itinerary (18 answers). Other 40 respondents said they cannot find any information, and 26 of them marked only this option. This number can be explained because people are not aware of the available resources to access these information. Therefore, it is not enough to make the information available, but it is also necessary to make them easily accessible to the different categories of people who need it.

4.2 Analysis of Hypothesis 2
The goal of the second hypothesis was to determine whether passengers make use of the knowledge shared by other people to obtain information about public transportation lines. We aim to verify if there is a tendency of sharing and searching this information via Internet.

In this regard, we asked about the way they get information
about bus timetables (Figure 2). 267 people (40%) said they ask other people and 321 people (48%) reported they obtain information via Internet. This shows that the Internet is the most used way to obtain information, but collective knowledge is also extensively used. By analyzing the people who checked both options, we come to a total of 139 passengers (21% of the total) who use the Internet and information from others.

Another question that confirms the data presented above concerns the way users look for information when using a line for the first time (Figure 7). From the 510 people that seek for information, 314 said they obtain it through acquaintances who use the same line, and nine said they go to the stop and ask other people. To the same question, 376 people answered they use the Internet as a way to obtain information. Only 41 people seek for the company services and other 30 reported other options. When analyzing the group of people that ask acquaintances or other people and use the Internet, we have a total of 197 people (39% of the 510 people that seek for information).

Among all 670 respondents, only 22 (3%) answered that there are ways to become aware of delays before the vehicle arrives at a stop. When asked about the way this information is obtained, ten people said they check on terminals available at the stops or ask the transportation company, three of them answered they can check delays via Internet and 10 people answered “other options” (ask relatives by telephone, talk to other passengers at the stop, ask vehicle conductors or traffic inspectors).

We also observed the tendency of using social networks to share and seek for information about incidents that occurred in public transportation (Figure 8). We asked if, currently, they post events related to the traffic on their social networks. 225 respondents (34% of the total) said “yes”. Among the 445 users who answered “no”, 248 use Facebook, 408 use Twitter, 33 use FourSquare and 55 reported they use other (LinkedIn, MySpace, Flickr, Youtube, and others).

We also asked the tendency of using social networks to share and seek for information about incidents that occurred in public transportation (Figure 8). We asked if, currently, they post events related to the traffic on their social networks. 225 respondents (34% of the total) said “yes”. Among the 445 users who answered “no”, 248 use Facebook, 408 use Twitter, 33 use FourSquare and 55 reported they use other (LinkedIn, MySpace, Flickr, Youtube, and others).

Regarding the use of information shared via Internet, we asked passengers whether they would use an application that displays information about bus lines in real time (dynamically) (Figure 9). Most respondents - 646 (96%) - said “yes”, they would use such an application.

**Figure 8: Information sharing on social networks**

![Figure 8: Information sharing on social networks](image)

**Figure 9: About the use of an application to obtain real time information**

According to the presented data, it is possible to say that most passengers get information by asking other people (i.e. from the collective knowledge). In addition, there is a tendency to share and search for information due to the high amount of passengers that already use the Internet to do it. Confirming this tendency, many passengers said that they already inform events that occur in public transportation or would inform it if there was an application to support this, and also that they would use an application that displays real-time information about bus lines.

### 4.3 Analysis of Hypothesis 3

The goal of the third hypothesis was to verify whether passengers trust on information provided by unknown sources or if they have more trust on information provided by people in their social network. To do so, respondents were asked if they participate in any social network (Figure 10). Out of all 670 responses, only 24 (4%) said “no”. Among the 646 who answered “yes”, 539 said they use Orkut social network, 448 use Facebook, 408 use Twitter, 33 use FourSquare and 55 reported they use other (LinkedIn, MySpace, Flickr, Youtube, and others).

People were also asked about their trust on information about bus lines posted on the Internet (Figure 11). From 670 responses, only 27 (4%) said they do not rely on information posted on the Internet. From the remaining 643 responses, 111 said they trust on the information, regardless of who posted it, without considering the information source. 196 of the respondents only trust on information posted by the transportation company. 56 only trust on information posted by people in their social network. And 280 people answered they trust on information posted on social networks and on information posted by the company.
The analysis of these data shows that, in general, people prefer to have information supplied by transportation companies. However, if we consider the people who relies "only on social networks" and those that indicated they trust both on "social networks and transportation companies", we have a total of 336 respondents, which shows that the information published on social networks has a high acceptability. Moreover, these numbers indicate that some people do not care about the information source (17 %), but most prefer to know the entity that provided it.

Figure 11: Trust on information posted on social networks

4.4 Analysis of Hypothesis 4
The focus of Hypothesis 4 was to check whether applications for mobile devices can be used as a way to share information about public transportation. To retrieve this information, we firstly asked passengers if they make use of Mobile Internet. 246 of the respondents answered "yes" and 424 said "no", as presented in Figure 12. The amount of people who answered "yes" is apparently small, but represents 37% of all respondents. Considering the high cost of devices that support web applications and data packages offered by mobile phone companies in Brazil, the proportion is very significant. We further noted that 178 out of 246 people who said "yes" live in the state capitals and other 34 live in cities with more than 200,000 inhabitants, which totals 212 responses of people living in big cities. This indicates that people using mobile technologies for Internet access live in cities that present major traffic problems.

Another important aspect is that only 49 respondents who answered "yes" are occasional public transportation users.

Figure 12: Amount of people that use mobile Internet

The other 197 people use public transportation “some days a week” or “every day”, showing that a substantial amount of people that use mobile devices to access the Internet are frequent public transportation users. It is estimated that over the coming years, this scenario will be increasingly common, as the cost of devices and data packages is reducing.

Thus, the answers indicate that the use of mobile applications is a trend and that, for this reason, it is important to focus on the development of such applications. Eventually, more people that use public transportation will have access to this technology and will take advantage of these applications.

5. CONCLUSIONS
This paper presented the results of a survey applied to public transportation users. The goal of this survey was to identify deficiencies, needs and trends related to providing information to passengers related to the public transportation real-time conditions. We could conclude that:

- passengers need information about public transportation and, although there are several information available, they are, in general, static and based on estimations. In most cases, information do not correspond to the real conditions at the time users need it;
- people often feel the need to seek for information about public transportation and they use more often the Internet and other people to achieve that information. Thus, it could be relevant to combine these methods into a single application in the web, which enable people to build a collective knowledge about public transportation conditions, making it available to all interested parties;
- using social networks to share information is an interesting alternative, since many people that use this kind of social media, share and seek for information through them;
- using mobile devices to access the Internet is a trend, and it enables information about public transportation events to be shared in real time, due to passengers’ ubiquity. This resource is interesting because it increases the power of collective intelligence.

From the results of this research, we are able to elicit the requirements that address the needs of the UbiBus project...
end users. The users’ feedback indicate that it is desirable to build systems that use collective intelligence and that enable passengers to share and seek for information about public transportation through social networks applications via Internet. By using these applications, the citizens will be able to use the Internet, mobile devices and social networks to indicate, in real time, events that help to map the routes situation and the public transportation conditions. From the information provided by several passengers, the system will present, for example, lines delays, vehicle problems, best or worst routes at a given moment, the best or worst lines and bus companies, more or less dangerous regions, and so on. These requirements are being implemented and will be integrated in the UbiBus system. The results of these applications will be presented to the community in future works.

A limitation of this paper is that the survey was answered by people in a subjective way, and there may be inconsistencies. It is not possible to certify the truthfulness of the information provided.

Acknowledgments
Authors thanks CNPq (process 560135/2010-6) for the financial support. This work was partially supported by National Institute of Science and Technology for Software Engineering (INES).

6. REFERENCES

16http://www.ines.org.br