A Wireless Infrastructure for Delivering Contextual Services and Studying Transport Behavior

Tiago Camacho, Vassilis Kostakos, and Claudio Mantero

Abstract— **We present an intelligent transport infrastructure for passenger sensing and data dissemination. The infrastructure consists of multiple distributed components and uses Bluetooth to sense mobile devices and by implication the passengers who carry these devices. By analyzing real-time sensed data the infrastructure provides a set of contextual services to passengers. The novelty of this work is the associated distributed architecture, and a specification language for defining contextual services. This paper presents tests of the performance of the infrastructure for both content dissemination and environment characterization, and the results highlight the usefulness of Bluetooth monitoring for studying and enabling inference of population behavior.**

I. INTRODUCTION

In this work we present an infrastructure used for data collection and context-aware content dissemination. Context-aware dissemination, and contextual services, refer to information services that are executed and[1] contact passengers only when a series of conditions are met. These conditions may refer to concepts such as the passenger's location, time, preferences, history, as well as transportation conditions. The system described here is composed of several distributed components that provide the basis for development and deployment of contextual services.

This paper describes our system architecture and details each component's functionality. We demonstrate that our system is a feasible and low cost solution for data collection and dissemination, and that it can be deployed at several locations. We show how we can use a specification language to define the context in which services are to be executed, and present results that illustrate the usefulness of the system in environment characterization. Moreover, our results suggest that Bluetooth data collection is useful for inferring aggregate passenger behavior.

This paper is structured as follows. In section II we describe some related work. In Section III we present our system architecture, describing individual components and the specification language. System test deployment is presented in section IV, while results from our pilot tests are shown in section V and discussed in Section VI. Our concluding comments are presented in Section VII.

Manuscript received April 12, 2010. This work is funded by the Portuguese Foundation for Science and Technology (FCT) grant CMU-PT/HuMach/0004/2008 (Sustainable Interaction with social Networks, context Awareness and Innovative Services), and the European Union project INTERVIR+.

Tiago Camacho is with the Madeira Interactive Technologies Institute at the University of Madeira, 9000-390 Funchal, Portugal (email: tiago@m-iti.org).

Vassilis Kostakos is with the Madeira Interactive Technologies Institute at the University of Madeira, 9000-390 Funchal, Portugal (email: vk@m-iti.org).

Claudio Mantero is with Horários do Funchal, Transportes Públicos S.A., 9020-242 Funchal, Portugal (email: claudiomantero@horariosdofunchal.pt).

II. RELATED WORK

A market study conducted in 2007 with 1000 subjects, indicates that about 22% of the Portuguese population aged 10 or more years, uses the Bluetooth/Infrared wireless technologies, while 88% owns a mobile phone [7]. The use of these wireless technologies increased about 2% between 2006 and 2007.

According to the same study, men (26%) are more prone to use these technologies than women (18%). Also, and according to Portugal's age structure [13], usage varies according to age group, with 13% of the users belonging to the 10-14 group, 36% to the 15-24, 14% to the 35-44, 5% to the 45-55, 1% to the 55-64, and 1% to the 65 or more. These numbers suggest a higher tendency of use by the younger age groups.

Bluetooth technology does allow users to disable the technology or even enable it in a non-discoverable mode. As such, a large number of Bluetooth-capable devices remain undetectable while on the move. Several studies have been made to assess the percentage of Bluetooth devices in discoverable mode. Results indicate different percentages according to location, with 2% for Bremen (Germany), 6% for San Francisco (USA), and 7% for Bath (UK) [10, 11]. More importantly, however, these studies indicate a positive linear correlation between the number of observable pedestrians and the number of discoverable Bluetooth devices.

A mobile device equipped with a Bluetooth transceiver that is set to discoverable mode allows for relative positioning determination. With the deployment of dedicated equipment that act as *scanners*, determination of the presence of a Bluetooth device is straightforward and practically imperceptible to the user.

Proximity sensing can be seen as the capability of sensing an object presence in relation to another object - such as a scanner. Bluetooth has been used previously to achieve this effect, both at a general level [6, 9, 10, 11], as well as in relation to public transportation passenger counting [4] and vehicle counting [14]. Besides this feature, Bluetooth also acts as a means for enabling interaction with users.

Aalto et al. developed a push-only dissemination system entitled *B-MAD* which makes use of localization information [1]. After a user registers its Bluetooth unique address (BD_ADDR) and cell phone number, the Bluetooth stations/scanners trigger a WAP push message over SMS when the device is detected nearby. The need for users to give away their phone number can lead to privacy issues, and the use of SMS involves higher operational costs.

In the context of public transportation, LeBrun & Chuah deployed what they term as Content Distribution Stations [5]. The idea behind this work is to take advantage of the idle time which passengers experience as they commute in

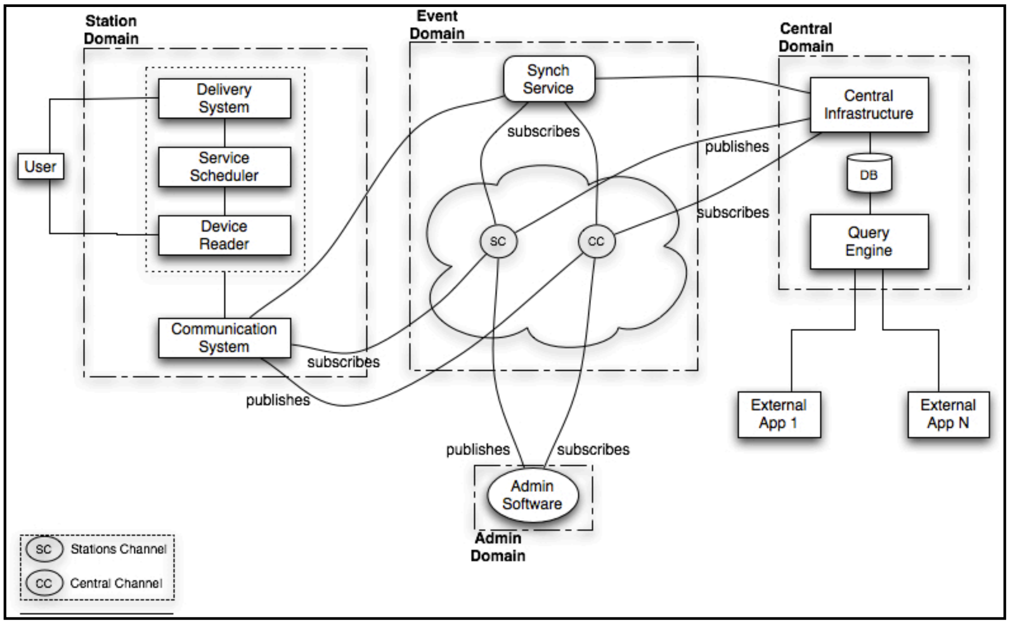

Figure 1. Overall System Architecture

their daily lifestyle. The stations are capable of delivering various types of content, ranging from textual news to more elaborate multimedia content. The system works using Bluetooth to establish a connection with the users, but also supports Wi-Fi. As the stations have no connection with the outside world, they need to synchronize with central servers which can be located at major public transit stops. The lack of synchronization between elements may lead to unexpected results, especially if the synchronization is not frequent.

Further work exploring public transit passengers commutation idle time is presented by McNamara et al., which enables peer-to-peer file sharing between users' Bluetooth-capable devices [8]. The authors describe an algorithm to determine the "capability" of other peers to transmit content of interest based on past history and content size. Making use of Bluetooth's Service Discovery Protocol (SDP) to determine appropriate devices' presence, users can receive content sent by other users as long as they are in the vicinity long enough. This peer-to-peer architecture allows great flexibility and removes unique points of failure, but crucially the work fails to address security and trust issues that may arise in comparison with a centralized content distribution scheme. Moreover, this system requires the installation of custom software on users' devices, which may lead to compatibility issues and to overall lower acceptance rates.

III. SYSTEM DESCRIPTION

Our system consists of multiple distributed components that communicate via a wireless Internet connection. Multiple stations are scattered across a city (e.g. at bus stops and bus terminals) where they collect data about passenger movement and also deliver contextual services to passengers via Bluetooth. Stations use a communication channel to route the collected data to the central infrastructure, which is the component that stores the information in a database. In addition, administrators can remotely construct new services, tweak existing services, and troubleshoot the infrastructure. The infrastructure is flexible and robust: new technologies (e.g. SMS, WiFi) can be incorporated, the effects of networking downtime have been minimized, faulty stations can be replaced with minimal effort, and the system is scalable. To describe how this has been achieved, we provide here a short technical description of the system.

I. Event domain

The event domain is the message bus of the system. It consists of event channels, a synchronization service, and provides a publisher/subscriber mechanism that the other components of the infrastructure use to communicate. The system is developed to support both XML-based (XMPP) and object-oriented based (CORBA) communication.

Two different event channels are used to send and receive messages asynchronously - the stations channel and the central channel. The former is used by stations to gather information directed at them, and the latter is used by the central infrastructure to update the database. The administration software also makes use of the event domain to control individual stations.

Although the publisher/subscriber mechanism provides flexibility and component decoupling, it hinders synchronization as it lacks acknowledgments of message delivery. To overcome this, a synchronization service located within the event domain listens to all published messages and stores them in an internal buffer. Upon initialization, and also periodically during run-time, both the stations and the central infrastructure contact the synchronization service to ensure all messages have been received.

II. Station domain

Stations are the components that actually interact with users. The purpose of the stations is two-fold: 1) continuously scan for devices within the vicinity, and 2) disseminate content to users. Several hardware configurations are possible for the stations components, but typically they are equipped with two Bluetooth transceivers:

a class 1 device with a 100m range, and a class 2 device with a 10m range.

The device reader component within stations acts as a Bluetooth scanner that continuously tries to detect devices in range. This information is then fed into the service scheduler, the component responsible for service execution and content delivery issuing. As the scheduler receives real-time contextual information, it examines its internal service queue to determine if any service object needs to be executed. If it does, then it contacts the delivery system whose responsibility is to execute commands that typically disseminate content to users.

Stations' communication with the outside is solely handled through the communication system, thereby shielding all internal components from the external world. The communication system forwards internal requests to the event domain. Additionally, it forwards remote requests received through the event domain to the correct local component. Most types of information forwarded to the event domain are buffered to improve transmission efficiency.

III. Central domain

The central infrastructure acts as a single repository that maintains detailed information about all stations. It stores all information sent to the event domain by the stations. It has a domain-specific structured database that holds all station related data. This includes information about which services were executed by which stations, which devices were detected, and status updates from stations.

Crucially, the central infrastructure provides a query engine interface so that external applications can access station-related information without contacting the stations directly. This enables for server-side applications such as real-time visualization of the network activity.

IV. Administration domain

The administration software can control most aspects of individual stations. It is possible, for example, to control the device reading settings, including options such as Bluetooth friendly name retrieval or device signal determination (RSSI). It can also be used to retrieve and manipulate the stations' service queues. An administrator may construct new services by using the Service Specification language described next.

V. Service Specification

Services are objects that define the context within which a specific piece of information will be delivered to end-users. As many authors state, context is a multi-faceted concept [2, 3, 12]. For this reason, flexibility in defining context is appropriate. Our Service Specification Language (SSL) is a way to describe the contextual requirements that must be satisfied in order for a service to be executed.

Figure 2 provides an overview of the properties of the service object. Service objects are differentiated through their ID and checksum fields. This latter field uniquely identifies services so that repetition is avoided, as services may be present in multiple stations.

The service objects also have two sets of flags associated with them: overall flags and trigger flags. The former define the source of the content (e.g. text or a URI), how to deliver the content (e.g. via Bluetooth), and if the service is to maintain state by remembering to which devices it already delivered content to. The trigger flags define the actual service execution. They allows us to specify the list of destination devices, the list of locations (i.e. stations) where content is to be delivered, the class of devices to deliver content to, and the concrete date of execution. Also, the trigger flags allow to define additional rules such as co-presence of devices, minimum and maximum time in range of the scanner, and the service execution frequency.

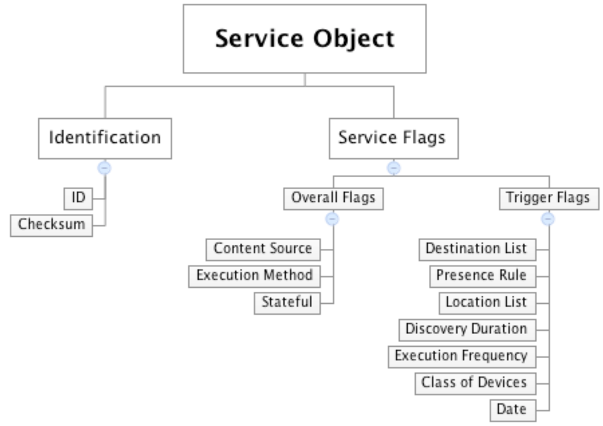

Figure 2. Service Object Structure Overview

VI. Sample Services

The flexibility of our Service Specification Language allows for a plethora of contextual services to be defined. For instance, a service can be defined to send someone a happy birthday message on the day of their birthday. Other services may be defined to send electronic vouchers to anyone who visits a particular location, to send the news headlines to every passenger at 5pm every day, or to send a link to a website to all bluetooth-enabled laptops. The following two scenarios describe potential uses of our system.

Mary is a public transit passenger that makes daily use of public transit buses to commute from home to work and vice-versa. Mary is a regular commuter and already knows the approximate schedules of the buses, and dislikes when they are not on time. Having heard of a new point award winning service that lets her obtain discount on her monthly boarding pass when buses are late, Mary went to the public transit company website to perform registration. Every time Mary now waits for the bus, she sets her device's Bluetooth to discoverable mode. The system can then identify her Bluetooth address every time she is waiting at the bus stop, routing this information to the central infrastructure. The public company can now cross-reference their buses positioning information with Mary's collected data to determine if she should be credited with award points.

Mark uses the bus several times a day to commute between places. He likes to be updated along the day about sport news, but sometimes it is difficult to do so as the only way is by using its mobile operator's Internet access. Knowing that the 3G wireless access is expensive, Mark avoids to make use of it and only accesses updated news

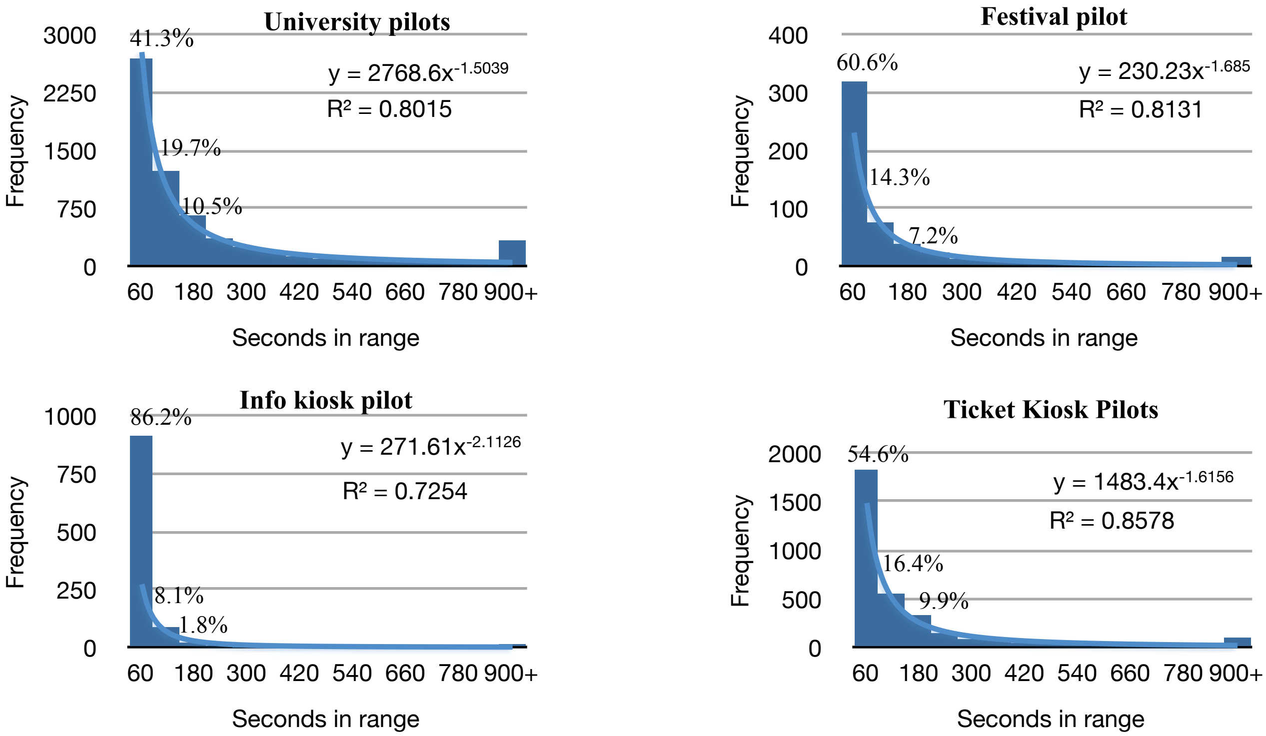

Figure 3. Devices continuously seen in range of the scanner for the university, festival, information kiosk, and ticket vending kiosk pilots respectively. Percentages on top of columns represent the total percentage of seen devices for the first three time frames

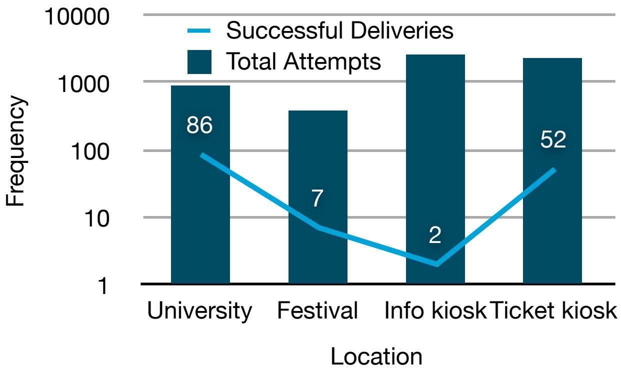

Figure 4. Successful deliveries in relation to total attempts for each of the test locations

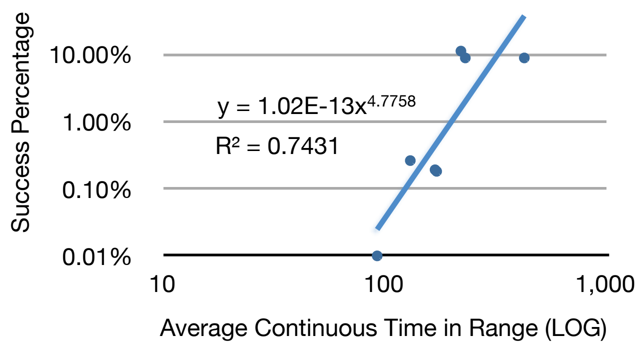

Figure 5. Correlation between pilots' average continuous device time in range and percentage of successful deliveries

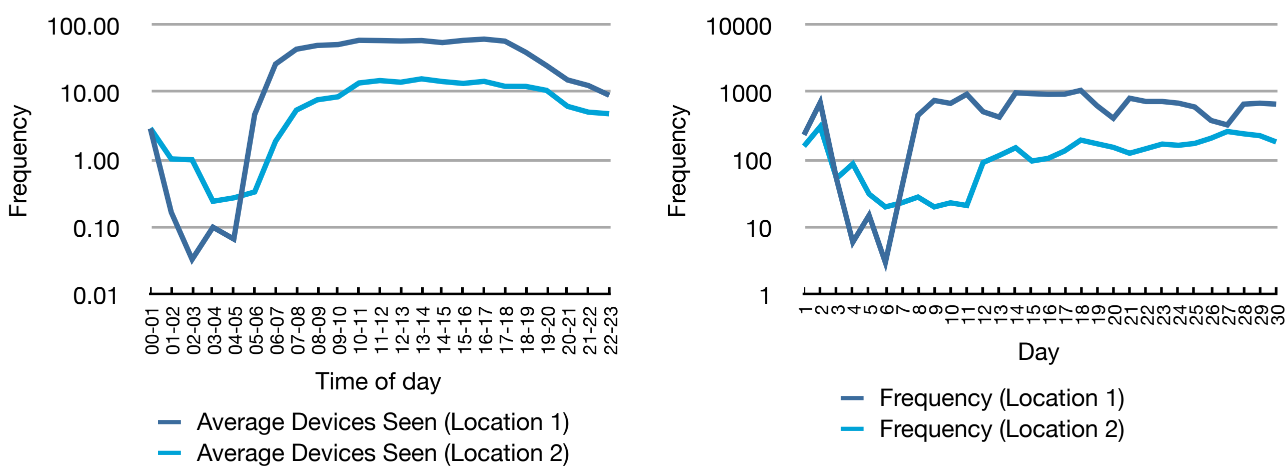

Figure 6. Left: Average number of devices seen according to time of day for both locations. Right: Number of devices seen per day of pilot. NOTE: On day 3 both locations were damaged by floods and landslides.

when he reaches home. Noticing that a new service that lets public transit commuters to access updated news for free is being offered, Mark goes to the public transit website to perform the registration process. After indicating his personal information, Mark now enables his device's Bluetooth capabilities every time he is near a bus stop or while traveling. As the system already has the knowledge of Mark's preferences, every time that his device is identified by the scanners, the system triggers the execution of a service that fetches content from an on-line service, formats it according to Mark's presentation preferences, and finally delivers the information to Mark.

IV. DEPLOYMENT

Our system is built in collaboration with Horários do Funchal, the public transportation company at Funchal, Portugal. The organization manages over 160 buses, serving about 130 million passengers per year, across more than 1400 bus stops. The company already employs a set of technologies on-board buses to facilitate fleet management and localization, such as GPS, GPRS connections, odometers, door sensors, and RFID readers.

For the system functionality assessment two sets of tests were conducted. The first set was carried out to test the dissemination capabilities of the system. On this set of tests two kinds of services were used. One of the services disseminated static content (e.g. that does not change over time) and the other disseminated dynamic content (e.g. that changes over time). Both services were "stateful", hence the station remembered the devices to which it delivered content, and would not send the information again to that device. In these tests only one station was used at a time.

The second set of tests was performed to determine the feasibility of our communication infrastructure. For this set of tests two stations were installed in two distinct points of the city and no dissemination occurred. The stations solely recorded all seen devices and sent this data back to the central infrastructure for storage and further analysis.

The stations used in all tests consisted of a small factor computer equipped with two Bluetooth transceivers and a 3G wireless Internet connection. One transceiver would perform repeated Bluetooth inquiries, and the other transceiver would - when applied - be responsible for pushing content to users using the OBEXPUSH protocol.

V. RESULTS

Our pilot tests consisted of two sets. The first was performed to test the dissemination infrastructure and the second set to test the communication infrastructure. In these results a "hit" is the successful delivery of information to an end-user's device.

In the first set, a total of nine pilots were executed. The first five pilots were conducted at the local university campus, the sixth pilot at a festival, the seventh at a tourist information kiosk, and the last two at a public transit ticket vending kiosk. Results have been grouped together according to location of execution. In Figure 3 we have the distribution of devices continuously seen in range according to time. It is observed that all locations follow an approximate power law distribution, with coefficients of determination ranging from 0.72 to 0.85.

Dissemination information seen in Figure 4 indicates that a higher number of successful deliveries was achieved at the university pilots followed by the ticket kiosk pilots, with 86 and 52 successful hits respectively. On the other hand, the lowest success rates were seen in the festival and information kiosk pilots, with only 7 and 2 successful hits respectively. Also, Figure 5 shows a correlation between pilots' average continuous device time in range and percentage of successful deliveries (R-squared = 0.74).

The second set exclusively consisted of information collection about Bluetooth enabled devices in the vicinity of the stations. Two stations ran for about one month, silently recording and routing information to the central infrastructure. Results for the second set can be seen in Figure 6, where the hourly and daily flow of devices at the two locations is shown.

VI. DISCUSSION

The results presented in Figure 3 suggest that independently of the location being observed, the time that a device is continuously seen by the scanner follows an approximate power-law distribution. This indicates that the majority of the devices spend a relatively short time in the same place. Nevertheless, all distributions also exhibit a small peak for the last category (900+ seconds), suggesting that a small percentage of devices are rather motionless regardless of the location. This could indicate that these are people working in the vicinity of the scanner.

Furthermore, inspection of the percentages presented on top of the first three categories reveals a variation according to location. We verify that a high percentage of devices (86.2%) was seen for less than one minute on the information kiosk pilot, contrary to the 41.3% observed in the university pilots. Therefore, such indicators may be useful in characterizing the flow in the observed environment. Comparing the data for the information kiosk pilot and the university pilots, we verify that the former environment is of a higher dynamic nature, as devices contact durations are short-lived in comparison to the latter environment. Additionally, the number of devices seen could be used to provide additional characterization of the environment. For example, by averaging the number of seen devices per hour, we can develop a concept of the volume of devices simultaneously present in a location.

The results also suggest that the nature of the environment influences the rate of successful delivery hits. For instance, Figure 4 shows that the environments where there was a higher number of devices spending less than one minute in range of the scanner corresponded to the environments with a lower rate of successful hits. This is also suggested by the results seen in Figure 5, where we observe that a power-law correlation exists between the average continuous time in range of devices and the delivery success percentages. This shows that we observed higher success rates in less dynamic environments. A hypothesis that explains this trend is that due to mobile devices' inappropriate feedback and user notification during Bluetooth message exchange, users normally do not notice that content is being sent to them at the first attempt.

In relation to the second set of tests, we observed that even though stations were deployed at two different locations in the city approximately one kilometer apart, results follow a similar trend. Looking at the left side of Figure 6 we observe that the average number of seen devices per hour peaks at approximate rush hours (9am-5pm), and that it follows an ascendent pattern with some fluctuations from 06:00 onwards. These results indicate that Bluetooth users may be representative of the general population's daily rhythm.

In addition, the number of seen devices per day of the pilot is shown in the right side of Figure 6. Besides verifying that both stations again follow a similar trend, these results reflect the floods that recently devastated some parts of Funchal. On the third day of testing, a flood and landslides reached several parts of the city severely damaging buildings and vehicles and consequently reducing the number of people in the center of the city. The figure not only reflects this fact, but also the subsequent recovery. Location 2 was more affected and showed slower recovery. The figure demonstrates the pattern of recovery seen as we can observe that the number of scanned devices slowly increased until reaching previously seen values.

From our findings we conclude two things. First, environment characterization is made possible using the Bluetooth inquiry procedure. This characterization may help us in determining the best approach for content dissemination. Results suggest that using the same method of dissemination independently of the nature of the environment leads to below par successful delivery rates (using push-only dissemination). As such, for more dynamic environments techniques such as increasing the number of simultaneous connections or decreasing the size of content to disseminate may lead to better results. Also, as stated in [8], data transfers in high density environments may see an increase in latency up to 50%. Being able to adjust the dissemination approach according to environment may prove to be an efficient way to tackle poor dissemination results, and this is something that the flexibility of our system enables.

The second implication is that Bluetooth is a feasible technology to infer behavior of the general population. Studying both the number of seen devices per day and the average number of seen devices per hour, we have observed that for two physically separated stations results follow a similar trend. Furthermore, both of the stations were able to reflect actual real-world accidents and the consequent pattern of recovery. We believe that the use of environment characterization can enable novel transport applications. These could be used, for example, to determine the approximate number of people waiting in a determined bus stop or infer the waiting patterns of passengers at specific bus stops. Applications also extend beyond the public transportation domain.

VII. CONCLUSION AND ONGOING WORK

In this paper we present a distributed infrastructure used for Bluetooth-based data collection and data dissemination. We describe our system architecture and demonstrate how we use a flexible specification language to define the context in which services are executed. Moreover, we describe possible applications of the infrastructure at a service level, as these can also be deployed at the server-side of our system.

The results demonstrate that our system is a feasible alternative not only for content dissemination, but also for environment characterization. Furthermore, we present data that highlight the benefits of using Bluetooth sampling to infer general behavior.

VIII. REFERENCES

1. Aalto, L., Gothlin, N., Korthonen, J. and Ojala, T. Bluetooth and WAP Push Based Location-Aware Mobile Advertising System. *Proc. of the 2nd International Conference on Mobile Systems, applications and services*. Boston, MA, USA, (2004), 49-58.
2. Chen, G. and Kotz, D. A Survey on Context-Aware Mobile Computing Research. (2000). Technical Report.
3. Dey, A. and Abowd, G. Towards a Better Understanding of Context and Context-Awareness. Georgia Institue of Technology, College of Computing (June, 1999). Technical Report.
4. Kostakos, V. Using Bluetooth to capture passenger trips on public transport buses. Arvix e-prints, Jun, (2008).
5. LeBrun, J. and Chuah, C. Bluetooth content distribution stations on public transit. Decentralized Resource Sharing in Mobile Computing and Networking, (2006), 63-65.
6. Leguay, J., Lindgren, A., Scott, J., Friedman, T. and Crowcroft, J. Opportunistic Content Distribution in an Urban Setting. *Proceedings of the 2006 SIGCOMM workshop on Challenged networks*. Pisa, Italy, (2006), 205-212.
7. Marktest study, 2007, http://www.marktest.com. Last accessed 1 Apr. 2010.
8. McNamara, L., Mascolo, C., and Capra, L. 2008. Media sharing based on colocation prediction in urban transport. *In Proceedings of the 14th ACM international Conference on Mobile Computing and Networking.* San Francisco, California, USA, (2008), 14-19.
9. Nicolai, T., Yoneki, E., Behrens, N. and Kenn, H.. Exploring Social Context with the Wireless Rope. On the Move to Meaningful Internet Systems 2006, 874-883.
10. Nicolai, T. and Kenn, H. About the relationship between people and discoverable Bluetooth devices in urban environments. Mobility (2007), 72-78.
11. O'Neill, E. Kostakos, V., Kindberg, T., Schieck, A. F. gen., Penn, A., Fraser, D.S. and Jones, T. Instrumenting the City: Developing Methods for Observing and Understanding the Digital Cityscape, *Proceedings of UbiComp 2006,* Lecture Notes in Computer Science 4206, Springer, Berlin / Heidelberg, (2006), 315-332.
12. Schilit, B., Adams, N. and Want, R. Context-Aware Computing Applications. *Proceedings of IEEE Workshop on Mobile Computing Systems and Applications.* Santa Cruz, California, USA. (1994), 85-90.
13. US Census. http://www.census.gov/ipc/www/idb/country.php. Last accessed 26 Fev. 2010.
14. Young, S. The Continuing Evolution of Travel Time Data Information - Collection and Processing: Utilization of Bluetooth Traffic Monitoring. *Transportation Engineering & Safety Conference*. Pennsylvania, USA, (2009).