Social and Technical Pitfalls Designing a Tourist Guide System

Christian Bornträger
Computing Department
Lancaster University,
Bailrigg,
Lancaster
Lancs., LA1 4YR
christian@borntraeger.net

Keith Cheverst
Computing Department
Lancaster University,
Bailrigg,
Lancaster
Lancs., LA1 4YR
kc@comp.lancs.ac.uk

ABSTRACT

In recent years, considerable effort has been made to explore several aspects of context-aware tourist guide system. Research about mobile guides is split into 2 main categories. The first category deals with technical solutions, the seconds deals with social and usability aspects. There is still a big need for detailed information how tourist interact and how a tourist guide system has to be designed. Therefore it is still necessary to do field trials with real test persons. This process is ongoing and several teams are planning tests or deployments of guide systems. Unfortunately, it is not a trivial task, to design a tourist guide prototype and to design an appropriate test for the social aspect of mobile guides. We made a field trial in November 2002 in Lancaster City to research the use of media types depending on the context of the user. This short paper does not concentrate on the primary results of our field trial but the side effects and information we got besides the focused topic that might help doing a better trial. Therefore, we deal with possible pitfalls and problems that occurred during the design phase and during our field trial. We discuss technical and social problems researchers should be aware of. Furthermore, we give hints and instructions to circumvent these problems to enable other researchers to do further field trials more easily.

Keywords

Context-Awareness, Tourist Guide, Design, Test run, Pitfalls

1. INTRODUCTION

In the last years, big effort has been made to build better mobile guides [1, 10, 12, 13, 3]. Even today, there is ongoing research in the topic of guidance system. This research has two main aspects. Some of the research projects are dealing with technical aspects, others with the human computer interaction. With no doubt, there will be further experiments in the future.

A lot of mobile tour guides are dealing with tourism, as tourism is a very interesting topic with quite good market chances. Furthermore, there seems to be a high acceptance rate among tourists. Though dealing with the technical aspect is not trivial, the understanding of the interaction between humans and computers is more complex and is still not very good.

In November and December 2002 we made a field trial with a multi-modal context-aware tourist guide system in Lancaster, UK. We were interested in the use of media types and the dependency on the context. Our prototype is based on a Compaq iPAQ PocketPC combined with a GARMIN Etrex GPS receiver. We consider a field trial as the best way to evaluate design decision of tourist guides. Real tourist seemed to be the best evaluators for our prototype. We acquired them in the Tourist Information Centre in Lancaster and near the Lancaster Castle. The tourist guide prototype application was designed to answer our question about context and its influence on different modes and medias. It was neither designed as a new base for future long lasting experiments nor for a deployment.

Unfortunately, running a field trial is not a trivial task. It is not possible to design a prototype and start the field trial immediately after the prototype is ready. Every supervisor should run test trials with usability experts before starting tests with real tourists.

There are known problems we already faced in former tests. Nevertheless, during our latest field trial, we faced several new problems that slowed down our trials. Therefore, this paper does not report the results of our field trial but gives an overview about pitfalls designers of tourist guides might fall in. We also try to give hints to circumvent these pitfalls and give general rules for running a field trial. Therefore, the guidance given in this paper should enable researcher in this area to run tests more easily.

In the next section we discuss the human aspects supervisors have to take care of. In section 3 we discuss technical aspects

and problems using state of the art but available technology, followed by a view on methodology in section 4. Finally there is a summary in section 5.

2. HUMAN ASPECTS

Acquisition of people

The most challenging part of a field trial is the acquisition of test persons. There are several ways of solving this particular problem. One possibility is to ask people in the laboratory and run the tests with these persons. Though it is quite easy to get these people, the major problem of this approach is the biased view of experts. This possibility should be used only to prepare the tests and to find flaws in your test setup.

Another way of getting test persons is to ask test persons offering money. This way is a quite common approach. It is the easiest way of getting people and there is probably a low rate of experts within that test group. Nevertheless, a test group with paid participants is different to your target group - tourists in our case.

Though we tested a tourist guide prototype in a university city it was not sufficient to ask several students of Lancaster University to participate. While newly arrived first year students have a limited knowledge about the city, the results will be biased nevertheless. These students do not mirror the target group.

Not only the social but also the age profile of a field trial should mirror the typical profile of the target audience. Therefore we decided to run a field trial with tourists we acquired in the Lancaster Tourist Information Centre (TIC) and nearby the Lancaster Castle. This results in a subsample of exactly the group we are interested in.

The help of the TIC was indispensable for our tests. The members of the TIC helped us not only getting and correcting the data of the guide system, they also helped to ask possible tourists. This is especially important as people have worries if an unknown persons approaches. The members of the TIC have a nimbus of an official person representing the state. Therefore, tourists are more open to participate in a field trial if the test is supported by official personal working in the area of interest.

Fears

We not only asked people in the TIC but also near the castle. In this area, there is no official TIC member giving your test the necessary respect and reputation.

One problem we faced during our field trial is the danger of being recognised as a sales agent, a survey person or another time-consuming agent. Therefore some people just pass by with a *"No, Thank you!"*. Fortunately, only a minority passes by. Most people listen to the introduction. One important aspect are the first two sentences. An introduction like "Hello, we are from the Lancaster University..." seems to give a good reputation as well. Unfortunately, not all people listening to the first sentences are going to participate in the field trial.

The most given excuse was *"Sorry, Unfortunately I don't have time"*. From our perspective it seems quite hard to convince these people to participate in a field trial. We believe, that this statement was a quite common excuse if people just do not like to participate.

Nevertheless, it is not trivial to convince open minded tourists as well. Quite often people asked *"Do I have to pay for this device?"*. It is essential to design the test setup as a free offer to circumvent this worry. But not only the design but also the necessary advertisement is important to convince people. Therefore, a test supervisor has to have a prepared list of several reasons why to participate. This list should be merged into a good and short introduction. Tourists should get the impression of not paying but getting something back. **Guideline: A test designer should prepare good reasons why the tourist should participate in this test. Examples: (1) You will learn more about this city in a quite short time. (2) You will get offers where to go next**

Another question we were asked was *"How long does it take?"*. This statement is different from the *"I don't have time"* as it shows interest in the test. Some tourists seem to consent spending some minutes but are unwilling to spend longer times, like half an hour. After using the tourist guide and seeing the advantages even people firstly being *"in a hurry"* might reconsider using the device for a longer period. Therefore a good answer to the question *"How long does it take?"* is "as long or as short as you like. If you want to stop, you can stop whenever you like". This opens a mental escape route for the tourists to stop the test without being worried causing a problem. **Guideline: Possible test candidates should not worry of spending too much time. A supervisor should emphasise the fact that the test person can go whenever he likes to**

The third important worry deals with the test itself. Our tourists asked, what will happen and *"What do I have to do"*. There are several reasons why people ask about the test before it starts. One worry is the necessary attention that people need. "Is this a test that costs half an our of my time and I have no additional use?" The arguments given in the previous section should solve this problem as well.

Another worry related to this question is the feeling of being unable to do what the supervisor is expecting. The youngest person in our test was 19, the oldest person 71. Some of them are agnostic in terms of technology. Therefore, we had to mention that they don't have to learn a complex interface. Furthermore some are quite happy hearing that it does not matter if something brakes. We asked them to walk around and use the simple interface to satisfy their needs. The simplicity and avoidance of complex tests in a laboratory convinces some people to participate. **Guideline: If your test is possible with simple and easy tests you should avoid creating a lab ambiance. Instead you should embed the test in their visit. This will give more authentic results than tests with payed "tourists"**

This worry is closely related to the Hawthorn effect. It was first discovered measuring employees of a company. If the employees know that they are measured, they change their behaviour. The same effect happens in other evaluations -

the test persons try to please the evaluator. It is not always easy to prevent this effect. Just keep in mind that people behave different if they know they are evaluated.

In some former tests, made during the Lancaster GUIDE project [6, 4] the test was supervised by more than one supervisor with at least on female and one male person. During the latest test [3] we were sometimes unable provide more than one supervisor doing the test and interview with the tourists. Comparing this experience with the former one, we have to highly recommend having two or more representatives asking for test persons.

No matter how good your preparation is, you will not be able to convince all people. There will always be a group of people that will not participate. It is quite hard to guess what kind of impact these people have on the final results.

Getting in Touch

During the test run there was always a supervisor which followed the tourists in a distance. The supervisor should help in case of technical problems as well as observing typical patterns of behaviour.

One interesting problem we faced during our trials is the fact, that tourist tried to get in touch with the supervisor. They started to talk, asking questions what this is for, what the supervisor is doing and so on. This might be a special case of a behaviour patter mentioned in [14]. Woodruff states, that the attention is split between the object, the guide book and the companions. As the supervisor is a human being, some tourists see the supervisor as a companion or just feel uncomfortable not knowing who is following them. Keeping in mind that the target audience (tourists in our case) does probably not participate for scientific enthusiasm but for their own pleasure, this behaviour is understandable.

The problem with this interaction is the influence on the result. It can be all right if a group member starts to talk to the supervisor, while the other members use the tour guide system. If this is a permanent interaction, account this group as a group with one person less.

What methods are available to prevent the interaction? First of all the supervisor should mention in the introduction, that he will follow in a **distance** and has no intention to influence the behaviour of the tourist. It should be made clear that the tourist should give the lead.

Furthermore, it can be helpful to increase the distance, if the first test runs show this behaviour. If the distance is more than about 5 meters, this behaviour should happen less often or never.

3. TECHNICAL PITFALLS

GPS as Position Detection

Designers of context-aware guide system, which are planning to use GPS as input for the positioning system should evaluate their needs for accuracy and the given location for their test.

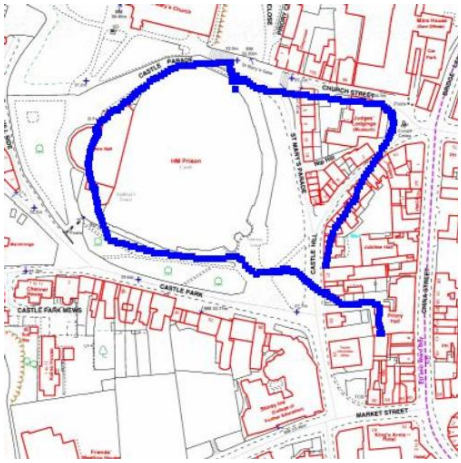

Figure 1: A good GPS trace

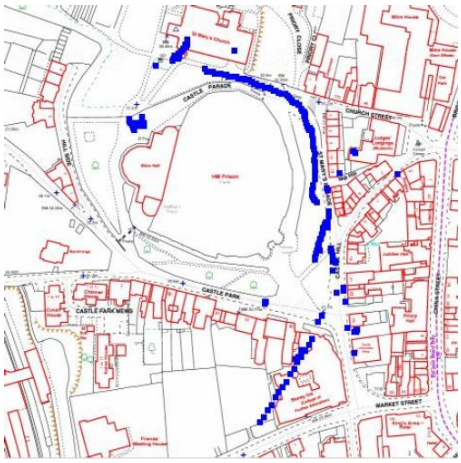

Figure 2: A bad GPS trace

As we designed an outdoor tourist guide, we used GPS and faced some drawbacks of this system. The selected availability (SA) was recently deactivated which increased the accuracy of GPS drastically. Nevertheless, GPS in not able to work indoors and is problematic in city canyons and under trees. The performance of GPS also depends on the time and the date as the amount of "spare" satellites can increase the accuracy.

Figure 1 shows a trace under good conditions and figure 2 a trace with very bad conditions.

Our experience showed that GPS needs a warm up period. Starting immediately after finding the first satellites gives bad results. This is especially problematic when we asked tourists in the TIC. Every time we started the test, walking outside the GPS-receiver took some time to finally getting the position. According to [11] the worst case scenario for reaching full accuracy lasts 20 minutes. This time is necessary to get all ephemeris, that means the coordinates and orbit of the satellites.

Apart from the experiences we made with the acquisition

of people this is another reason for having at least 2 supervisors. One waiting outside, keeping the GPS working and one inside the building.

Nevertheless, most of the time the GPS system was sufficient for our test area and our interests.

Size of the GPS receiver

To use a wireless LAN link in our Compaq iPAQ we used a 802.11 network card in the extension slot of our prototype. Therefore we used an external GARMIN GPS receiver plugged into the serial port. Some test persons mentioned, that the "final" product should include this external device and combine both devices. Nevertheless, in our test there was no major problem related to this downside of our solution. Newer versions of some PDAs offer build-in Wave-LAN and/or GPS. This problem will disappear within the next years. Nevertheless, as we mentioned, that GPS is not optimal for every guide system this gives another reason to look for alternatives. Possible ideas are Bluetooth cells or Wave-LAN triangulation [2].

Effect of Lags

In some special locations the GPS receiver looses contact to the satellites. Therefore the positioning can be disabled for some seconds. Most of the test people using the map interface realised that the cursor stopped moving. All of them expected a GPS receiver to work everywhere and perfect. If they became aware of this short time lag they became unsure and quite often asked the supervisor if something is broken. Even short periods caused this effect. Nobody complained about these lags afterwards in the interview - the amazement happened quite spontaneous but seems to go away. Nevertheless, future tests should evaluate possible interface metaphors to minimise the amazement.

Usage of Headphones

We already stated in [3] that our test person avoided the usage of headphones. Nevertheless we think that we should mention this fact again as this design decision is another possible pitfall. No person that participated in our 20 test runs wanted to have head phones but preferred the speaker instead. There was even one test person which used a Walkman with headphones inside the Tourist Information Centre. This test person rejected the offer to use his head phones instead and insisted on using the speaker. There was only one hard of hearing test persons that resigned oneself to use headphones because the speaker was not loud enough.

This result was the biggest surprise in our test. We, as the designers expected to have a normal consumption of head phones, finally we used only one. Some former mobile test projects are based on head phones indoors or outdoors [8, 7]. With our results in mind, we have to conclude, that their design decision might be problematic, at least for the outdoor guides. The situation might be completely different for indoor systems. As we have not evaluated an indoor guide recently this should be evaluated in the future.

There are several reasons, why people rejected head phones. Tourists claimed in the interview, that *headphones look strange* and they *don't want to be recognised as tourists.* This explanation is quite interesting because wearing a PocketPC with a speaker stating facts about the area will qualify them as tourists nevertheless. Other aspects are the fear of the unknown and privacy. As our prototype guide is a PUSH-based guide [5] there is a problem that tourist do not exactly know, when audio is pushed.

Furthermore, before the test the tourists do not even know what they will hear. Therefore it seems safer for them to not use the headphones. Having an unknown sound near the head might violate the personal privacy of the users.

Another aspect is the feeling of being isolated from the soundscape around the tourists or their companions. Quite a lot of people feel uncomfortable if they talk to somebody wearing a head phone. Besides closed and semi-open headphones there are open headphones that don't block ambient noise, but most people are not aware of this technical possibility.

4. ASPECT OF METHODOLOGY

Flexibility During the Test

In our first design we implemented an accurate highly detailed map. It was a classic map, and not a sophisticated 3d map shown in [9]. The choice was not intentionally but we just used an available digital map. During the first test runs, the tourists gave comments about the map, which is too hard to read. Therefore we added a symbolic map afterwards to please these wishes. Both maps were available and the tourists could choose the map.

In the following test runs most tourists preferred the symbolic and simple map. Therefore, it was a good decision to not stick to the old design but being sufficiently flexible during our test run.

Besides the unexpected avoidance of headphones this is the second example where a design decisions did not conform to the wishes of the test persons. Therefore it is necessary to keep flexible all the time, not only in the semi-structured interview, but also about the test design and the used tests.

Unfortunately, changing the test setup might have an influence on the results. This methodological problem can cause incomparable data. One solution is to do these changes quite early and only count the results after the changes. The results of the first test runs can then be evaluated separately.

Fortunately, the results of the log files and the opinions from the interviews are not always heavily affected.

Guideline: Assumption may be wrong. Sometimes the test has to be changed to reflect the new knowledge. The test designer has to evaluate the influences of the changes. It might be necessary to discard or split some results.

Designing the interview

A field trial should always be augmented with an interview of the test persons. We decided against a written paper feedback form and used a semi-structured interview. It is nice to know that most people do not mind of the interview

is recorded with a Dictaphone. This helps not only getting more information from the interview but also to recheck if the supervisor should focus on special areas in more detail. Though our interview is not written, a structure or list of questions is necessary.

In our tests, tourists asked to give comments about the device answered not the questions we were interested in, but evaluate your guide system *("it was good..", "it was easy to use...",)* Therefore this question was only a small introduction for a more detailed semi-structured interview.

The most important aspect of the interview is of course the aspect you are researching. Nevertheless, we experienced good results by not sticking to a strict order of questions. Instead the order and choice of question was dynamic, depending on the former answers. Just pay attention to focus the users to the questions you are interested in.

As we recorded the interview we were able to get some interesting results. Getting these results would have been impossible by only evaluating log files. Especially wanted features are a typical result from an interview. Sometimes a test person had no clue, what additional feature should be included. In this case we gave small examples (like a self rotating map). Afterwards the tourist came up with more ideas what kind of feature should be added. **Guideline: In an interview tourist might have a black out regarding some questions. Giving small examples often leads to new ideas**

5. SUMMARY

In this paper we discussed some aspects designers of field trials should be aware of. Furthermore we gave instructions and hints for future tests. Test supervisors should prevent special pitfalls before and during the test. If possible, experienced test designers should participate in the field trial.

To acquire test persons you should have a good offer. Even more important, there should be some answers for the most common worries of tourists. Designers should prepare good and convincing answers for possible questions like "How long does it take?", "Do I have to pay for it?" and "What do I have to do? These answers will influence the decision if people will participate in the test run.

During the test run consider the Hawthorn effect and the urge to communicate with the supervisor. Furthermore, your system and the technology should be stable, otherwise the test persons might get confused.

Test designers should be flexible enough during the test and between the test runs. Our trials showed that the design decision based on the opinion of the designer is not always in coherence with the opinion of the test persons. If you change the setup, make sure to not spoil the results.

Additionally to a test run with logging you should always do an interview. We think that a semi-structured interview is the best compromise between completeness and cooperation.

6. REFERENCES

[1] G. Abowd, C. Atkeson, J. Hong, S. Long, R. Kooper, and M. Pinkerton. Cyberguide: A Mobile Context-Aware Tour Guide. *ACM Wireless Networks*, 3:421–433, 1997.

[2] P. Bahl and V. N. Padmanabhan. RADAR: An In-Building RF-Based User Location and Tracking System. In *Proceedings of The IEEE Infocom 2000*, pages 775–784, 2000.

[3] C. Bornträger, K. Cheverst, N. Davies, A. Dix, A. Friday, and J. Seitz. Experiments with multi-modal interfaces in a context-aware city guide. In *Proceedings of Mobile HCI 2003*. Springer.

[4] K. Cheverst, N. Davies, K. Mitchell, A. Friday, and C. Efstratiou. Developing a context-aware electronic tourist guide: some issues and experiences. In *Proceedings of CHI'00*, pages 17–24, 2000.

[5] K. Cheverst and G. Smith. Exploring the notion of information push and pull with respect to the user intention and disruption. In *Proceedings of The International workshop on Distributed and Disappearing User Interfaces in Ubiquitous Computing*, pages 67–72, November 2001.

[6] N. Davies, K. Mitchell, K. Cheverst, and G. Blair. Developing a Context Sensitive Tourist Guide. In *Proceedings of The First Workshop on Human Computer Interaction for Mobile Devices*, 1998.

[7] G. Eckel. Immersive Audio-Augmented Environments. In *Proceedings of the 8th Biennial Symposium on Arts and Technology*, 2001.

[8] S. Holland and D. R. Morse. Audio GPS: spatial audio in a minimal attention interface. In *Proceedings of The Third International Workshop on Human Computer Interaction with Mobile Devices*, 2001.

[9] M. Kulju and E. Kaasinen. Route guidance using a 3d city model on a mobile device. In *Workshop Mobile Guides, Mobile HCI*, 2002.

[10] R. Oppermann and M. Specht. A nomadic information system for adaptive exhibition guidance. In *ICHIM*, pages 103–109, 1999.

[11] P. Röbke-Doerr. Navigation mit satelliten. *c't magazin für computer technik*, 1:150–151, 2003.

[12] N. Sawhney and C. Schmandt. Speaking and Listening on the Run: Design for Wearable Audio Computing. In *Proceedings of The International Symposium on Wearable Computing*, pages 108–115, 1998.

[13] B. Schmidt-Belz, S. Poslad, and A. Zipf. Creation of user-friendly mobile tourism services. In *Workshop Mobile Guides, Mobile HCI*, 2002.

[14] A. Woodruff, P. M. Aoki, A. Hurst, and M. H. Szymanski. Electronic Guidebooks and Visitor Attention. In *Proceedings of The International Conference on Hypermedia and Interactivity in Museums*, pages 437–454, 2001.