ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Interacting with Computers 17 (2005) 343–366

Interacting with Computers

www.elsevier.com/locate/intcom

Mobile agents for mobile tourists: a user evaluation of Gulliver's Genie

M.J. O'Grady[a,*], G.M.P. O'Hare[a], C. Sas[b]

[a]*PRISM (Practice and Research in Intelligent Systems and Multimedia) Laboratory, Department of Computer Science, University College Dublin, Belfield, Dublin 4, Ireland*
[b]*Computing Department, Lancaster University, Lancaster LA1 4WA, UK*

Received 8 March 2004; revised 10 January 2005; accepted 25 February 2005
Available online 25 February 2005

Abstract

How mobile computing applications and services may be best designed, implemented and deployed remains the subject of much research. One alternative approach to developing software for mobile users that is receiving increasing attention from the research community is that of one based on intelligent agents. Recent advances in mobile computing technology have made such an approach feasible. We present an overview of the design and implementation of an archetypical mobile computing application, namely that of an electronic tourist guide. This guide is unique in that it comprises a suite of intelligent agents that conform to the strong intentional stance. However, the focus of this paper is primarily concerned with the results of detailed user evaluations conducted on this system. Within the literature, comprehensive evaluations of mobile context-sensitive systems are sparse and therefore, this paper seeks, in part, to address this deficiency.

Keywords: User evaluation; Mobile computing; Context-sensitive service delivery; Intelligent agents

1. Introduction

At present, mobile computing offers a vision of applications and services that is generally unobtainable for the average person. The reasons for this vary, and

* Corresponding author. Tel.: +35 317162922; fax: +35 312697262.
E-mail address: michael.j.ogrady@ucd.ie (M.J. O'Grady).

0953-5438/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.intcom.2005.01.003

paradoxically, have little to do with the technology per se. Issues as disparate as cost, software development environments for mobile devices, lack of off-the-shelf packages as well as popular misgivings about security and privacy all hinder the widespread deployment of wireless applications. While research into device integration and interoperability continues unabated, nevertheless, one critical issue remains—how best to design and implement software that will perform reliably and efficiently within the limited computational constraints of such devices while simultaneously meeting the expectations of a demanding user group.

Within this paper we introduce Gulliver's Genie (O'Hare and O'Grady, 2003), a PDA-hosted electronic tour guide typical in many respects of a broad class of location-aware or context-aware services. Though the literature is sprinkled with prototypes of context-sensitive applications, comprehensive and systematic user evaluations are noticeably absent. In an effort to improve this situation, the results of a user evaluation are presented and analysed. In contrast to all but a few such systems however, in our design and implementation we have adopted an approach based on the deployment of *intelligent agents*. Indeed, the use of mobile software agents on a PDA forms a key constituent of Gulliver's Genie and places it at the forefront of a new generation of intelligent services. Furthermore, market analysis has consistently revealed a need for wireless handheld devices that function as personal tour guides and which act as an all-in-one device, subsuming and replacing the myriad of pre-existing devices, (PC, cellular phone, electronic diary).

Most research to date in the area of electronic guides has focused on the needs of indoor users, for example, visitors to museums and art galleries. Bellotti et al. (2002) provide a recent and useful overview of research in this area. However, the outdoor environment offers its own unique set of challenges. Two of the most widely documented systems for the outdoor environment are CyberGuide (Abowd et al., 1997) developed by Georgia Tech, and GUIDE (Davies et al., 2001), a system developed for the city of Lancaster. While Gulliver's Genie shares some of the objectives of these systems, it differs radically in its use of intelligent agents as a means for achieving its objectives. In the following sections, the state-of-the-art in intelligent agents is described. Following this, we provide a brief but thorough description of Gulliver's Genie. We then present and discuss the results of the user evaluations. Finally, we present our conclusions.

2. Intelligent agents

Gulliver's Genie adopts an *agent-oriented paradigm*. The advantages of this approach include modular design, system intelligence, enhanced system reliability through graceful degradation, expandability and load balancing via agent migration. Gulliver's Genie commissions mobile intelligent agents from the Belief Desire Intention (BDI) family (Kinny et al., 1996). Such agents exhibit all the characteristics of the notion *of strong agenthood* (Wooldridge and Jennings, 1995) with the agent mental state comprising of an aggregation of beliefs, commitments, and commitment rules. This mental state drives all future agent actions and the adoption of commitments further influences future activities. The mental state of an agent evolves with respect to time, with its belief set being updated

according to both its perception of its environment and the receipt of communications from fellow agents with whom it is acquainted. The *Agent Factory* system (Collier et al., 2003) is an environment developed in part by one of the authors which supports the fabrication of agent-based applications. Agent Factory comprises an integrated environment for the rapid development of agent-based systems, providing a methodological framework together with an accompanying set of software tools that support the various stages in the design, specification, implementation, debugging and visualisation of agent behaviour. In order to support agent deployment on PDAs, a lightweight implementation of Agent Factory consisting primarily of the agent runtime environment was developed in Java thus enabling the deployment of intelligent agents on devices possessing limited computational resources.

Mobility is a desirable characteristic of agenthood and Agent Factory naturally supports process migration. Mobile agents move, or in agent parlance, migrate, by transmitting their code from one computer platform to another. Generally speaking, mobile agent systems provide one of two kinds of migration:

- *Strong migration* where the agent's object state, code and control state is captured. Upon migration this allows the agent to resume execution on the new machine from the exact point that it left off.
- *Weak migration* where only the agent's object state and code is captured. Upon migration the system calls a known entry-point in the code to restart the agent on the new machine.

The deployment of agents in the delivery of context-aware services has begun to find some favour, though agents in such systems are invariably singular, not multi-agent in nature and do not fully exploit features characteristic to *strong* agents. One of the pioneering mobile systems to utilise agent technology was TravelMATE (Cheyer and Julia, 1995), a prototype system developed by SRI International as part of their research into multi-modal user interfaces. More recently, several agent-based location-aware services have emerged, for example, Impulse (Youll et al., 2000), Mia (Buester et al., 2000) and ComMotion (Marmasse and Schmadt, 2000). Specifically CRUMPET (Poslad et al., 2001) attempts to address the issues involved the integration of diverse services in the effective delivery of location-aware services to the mobile tourist community.

It is instructive to reflect briefly on why agents are being considered as a basis for context-aware applications such as the Genie. Classic agent attributes include proactivity, reactivity, mobility, social ability and autonomy. Though conventional software can be explicitly endowed with these characteristics, they are inextricably linked to the agent paradigm. BDI agents are augmented with additional characteristics, namely a reasoning mechanism based on mentalistic notions as described previously. Such constructs offer designers alternative and intuitive mechanisms for modelling software and can be particularly useful in complex and unpredictable environments where conventional software approaches are severely challenged. Supplying context-aware services to an end-user whose context is inherently dynamic, for example a roaming tourist, is an obvious example of such a scenario.

3. Gulliver's genie

Meeting the navigation, cultural and information needs of tourists is the raison d'être of Gulliver's Genie.[1] Given the reluctance of people in general to embrace new ideas, unless of course some advantage can be manifestly demonstrated, tourists are unlikely to adopt systems such as the Genie unless they palpably offer some value added services. Therefore, we have identified navigation support and the dynamic delivery of personalised cultural information as two areas in which such services might be usefully based. While acknowledging that standard location-aware services are important to tourists, we see their future incorporation into the Genie as essential, though not an immediate priority at this time.

The modus operandi of the Genie is straightforward. As a tourist, equipped with a PDA hosting the Genie software, wanders a city, they can see their current location and orientation at all times on an electronic map. On approaching various tourist attractions, they are automatically presented with multimedia-rich presentations on the attraction. In so far as possible, the presentation has been adapted to reflect the tourist's interests. The tourist may view as much or as little of the presentation as they choose. On moving away from the attraction, the default interface containing the electronic map is restored.

3.1. Design objectives

The key research goal of Gulliver's Genie was the adoption of strategies for minimising, and, when possible, overcoming the difficulties inherent in operating within the confines of a computationally limited device whilst delivering context-sensitive information to mobile users. Therefore a number of design criteria were identified:

(I) Dynamically adapt content to both the perceived interests of the tourist and to their spatial context.
(II) Facilitate intuitive and simple interaction with the Genie.
(III) Exploit the autonomous and proactive nature of intelligent agents to anticipate and provide for tourists' content requirements in a timely manner.
(IV) Ensure that the services provided by the Genie could be easily duplicated in new geographic areas.
(V) Maximise the deployment potential of the Genie through a judicious selection of enabling technologies.
(VI) Place minimum restrictions on the tourist's selection of device by ensuring that the Genie is portable.

Reconciling these criteria and ascribing priority to each required careful consideration and occasional compromise. The first and third items are essentially concerned with software design. In order to fulfil the second criterion, two interaction modalities were incorporated:

[1] Gulliver's Genie will henceforth be abbreviated to 'the Genie'.

(I) *Implicit*. Tourists can interact with the Genie by their behaviour and activities within an environment. For example, by facing a sculpture for some time, it may be inferred that a tourist is indicating some level of interest in it.
(II) *Explicit*. Tourists can explicitly interact with the Genie by means of a stylus, for example, they can browse the map for items of interest within their vicinity.

The fourth criterion is addressed by designing a toolkit for assembling the relevant multimedia content and associating it with the appropriate attractions. Indeed, this toolkit is also essential in achieving the first design criterion. Some spatial logic is delivered through the incorporation of a Geographic Information System (GIS) component. In addition, a component for defining user profiles and correlating such prefabricated profiles with the appropriate multimedia content is also provided. Both of these components are essential for realising the adaptive nature of the Genie.

The final two criteria are concerned with the selection of the appropriate hardware and software. In essence, we did not have the luxury of designing and deploying a specialised networking infrastructure totally dedicated to meeting the needs of tourists. Likewise, the option of designing a handheld device optimised for the use of tourists was non-existent. In short, we had to work within the constraints of pre-existing and commonly available technological solutions as we intended that the Genie would operate satisfactorily in the real world.

To host the Genie, we identified the IPAQ as an archetypical PDA device. In particular, it could display various multimedia types and its functionality could be expanded via a dual-slot expansion sleeve. Two extra features were required by the Genie, namely, a wireless communications facility and a mechanism for determining position. In the latter case, GPS is the only practical option at present. Support for an electronic compass is also provided for. Unfortunately, hardware limitations excluded its use. As a substitute, the GPS bearing parameter was used. In essence this is a vector of the tourist's movement; thus while it coincides with the direction the tourist is facing in most circumstances, it cannot be assumed to do so in all cases. Nevertheless, it forms a key component of the tourist's spatial context.

For enabling wireless data communications, the situation was more complicated than with position determination. As yet, there is no single worldwide standard for accessing mobile telecommunications networks. That being the case, we chose to work with the current prevailing standard here in Europe, namely the General Packet Radio Service (GPRS). Java was selected as the development and deployment environment. In reality, there was little choice given that we wanted to ensure maximum portability. Though it is presently unrealistic to expect every tourist to own a PDA supporting this configuration, it is plausible to assume that this situation may change over time. It is also likely that the future PDA will be seamlessly integrated with a wireless communications facility and some positioning component of either the satellite-based or cellular network-based variety.

3.2. Architecture

Gulliver's Genie follows a standard client/server architecture. The Genie client is hosted on a PDA, which is augmented with GPS and GPRS facilities (Fig. 1). The Server

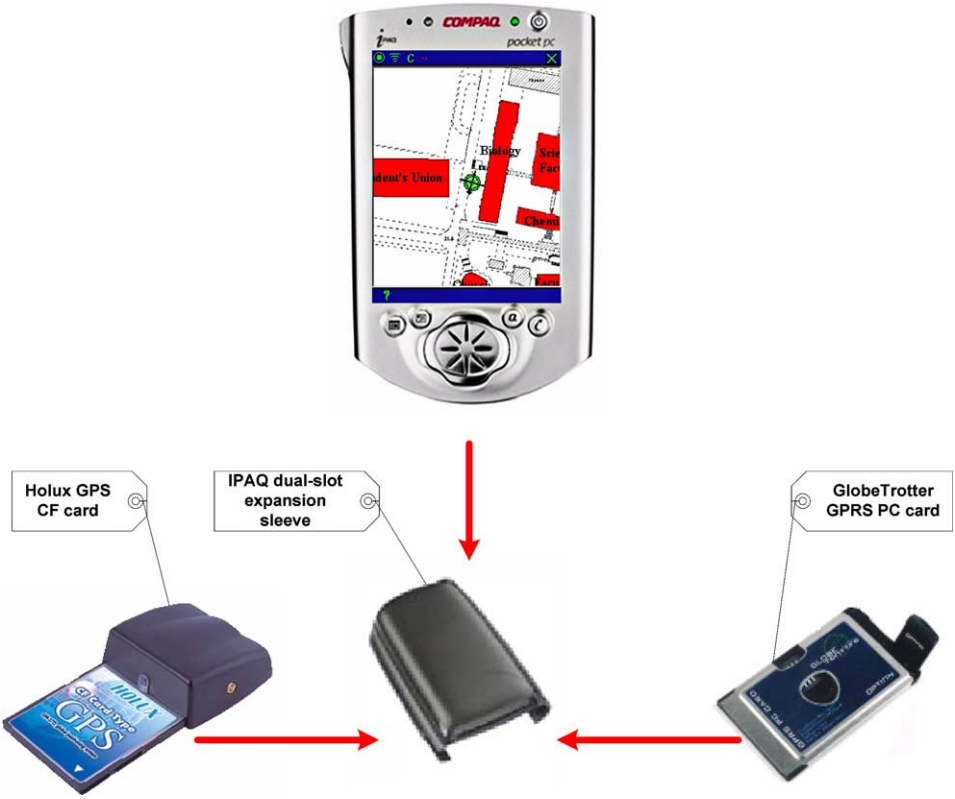

Fig. 1. The hardware that comprises the Genie client.

is hosted on a normal workstation that has Internet access. HTTP is the protocol over which all communications occurs. The logical structure of both the client and server is outlined in Fig. 2 and the critical components are now described.

3.2.1. Genie client

- *Spatial Agent*. The Spatial Agent continuously monitors the GPS sensor. From this it extracts both position and orientation readings. After verifying that the readings are accurate and consistent, it broadcasts them to all other interested components.
- *Cache Agent*. This agent collaborates with the other agents (including those on the server) to ensure that the cache on the client always contains information that is pertinent to the tourist's immediate location and reflects any recent changes to their interest profile. In addition, it decides if and when to inform the tourist about any attractions within their immediate vicinity.
- *PDA Controller*. The PDA Controller is responsible for initialising and terminating Genie sessions. Nominally, it controls the Genie, but for the most part, acts under the direction of both the Spatial and Cache Agents, respectively.

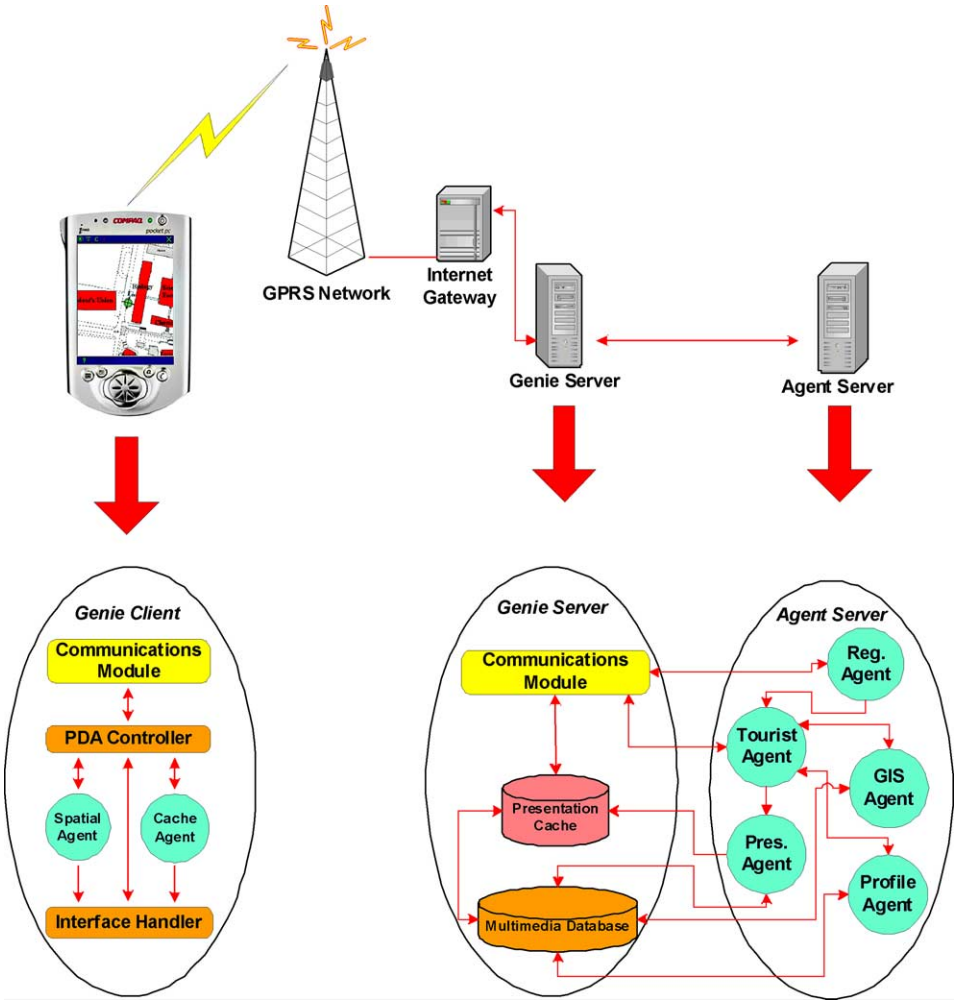

Fig. 2. Architecture of Gulliver's Genie.

- *Interface Handler*. This component is primarily responsible for managing the user interface as well as capturing any explicit interaction the tourist may initiate via the stylus or the PDA's 5-way navigation pad.
- *Communications Handler*. Schedules the dispatch of messages to the server and handles any responses accordingly.

3.2.2. Genie server

- *Communications Module*. Routes incoming requests from Genie clients to the appropriate agent or component and returns all responses to the corresponding client.
- *Presentation Cache*. Controlled by the Presentation Agent, this cache contains

pre-build presentations of all attractions within the tourist's immediate vicinity. It is continuously updated in light of ongoing tourist movement and changes to the tourist's profile.

- *Database*. Underpinning Gulliver's Genie is a sophisticated database, which, in essence, defines the Information Space supported by the Genie. As well as containing information of a geospatial nature, it also contains the individual multimedia snippets from which the presentations are assembled.

3.2.3. Agent server

- *Registration Agent*. After verifying and authenticating the tourist, the Registration Agent assigns an individual Agent (from the Tourist Agent template) to the tourist. On disconnecting, any resources that are in use are reclaimed.
- *Profile Agent*. Responsible for deriving inferences from the tourist's interaction with the Genie and updating the tourist's profile accordingly.
- *GIS Agent*. The GIS Agent handles cache update requests from the client. It queries the database and, based on the current position of the tourist, constructs a model of the tourist's immediate environment that includes details of those attractions that are within the tourist's immediate vicinity before returning it to the Cache Agent for closer observation. Note that it is the tourist's own agent that actually co-ordinates this process!
- *Tourist Agent*. On registering with the Genie Server, all tourists are assigned their own individual agent. This agent acts as an interface between the client and the server. Based on information received from the agents on the client, it maintains a snapshot of the tourist's activity at any given time. It co-operates with the GIS Agent and the Profile Agent to ensure that the contents of the client's cache are always appropriate to the tourist's position.
- *Presentation Agent*. Given a tourist's profile and a model of their environment, the Presentation Agent builds presentations for all attractions in the tourist's immediate vicinity and stores them on a cache on the server. In this way, requests from tourists are anticipated and no time is lost assembling the required presentation before dispatching it to the client. This cache is continuously updated in response to changes in the tourist's environmental model and amendments to their profile.

3.3. The tourist's experience

A prospective tourist must first register with the Genie Server during which time a number of cultural domains of interest to them, for example, art, architecture, literature, folklore, history and religion must be specified. They may then utilise the Genie in any area for which an Information Space has been designed and assembled. Its operation is quite simple. As tourists explore an area, they can see their current position and orientation on the electronic map at all times (Fig. 3). In addition, they may also browse the map for items of interest in their vicinity. In this way, their essential navigational requirements are addressed. In the meantime, they are being continuously monitored by their agent on the server, which keeps the Cache Agent briefed on what exhibits are in their immediate

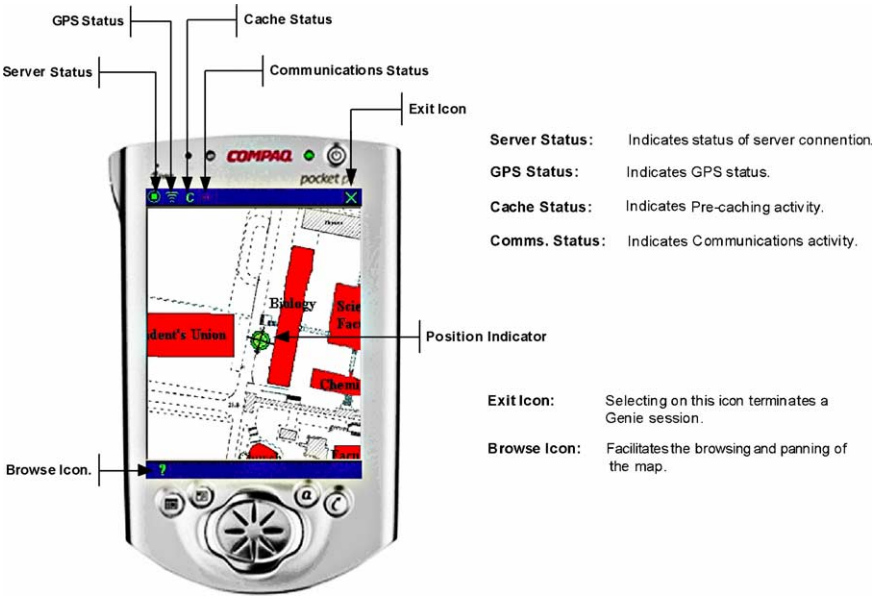

Fig. 3. An illustration of the standard navigation screen with the tourist's current position and orientation indicated by the green icon. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

vicinity. Also, presentations are being continuously pre-cached on the server in anticipation that a request may be forthcoming at any time. The tourist, of course, is completely oblivious to all this activity.

As the tourist approaches some attraction, the Cache Agent, sensing this, will request the appropriate presentation from the server. This process, which we have termed *intelligent precaching*, seeks to anticipate where the tourist is heading and to pre-arrange that the appropriate content is delivered to their device in an almost *just-in-time* basis (O'Grady and O'Hare, 2004). On receiving the presentation, it confirms that it is still valid for the tourist's position and, if so, proceeds to display it. All presentations conform to the template in Fig. 4. Using the stylus, the tourist may listen to other information snippets of interest. A record of their interaction is returned to the Profile Agent for analysis after which the tourist's interest model may be further refined. On moving away, the Genie restores the default navigation screen.

If you recall, the Genie was required to support both explicit and implicit interaction modalities. We will now illustrate how this is achieved. As can be seen from the previous discussion, the tourist indicates certain preferences through their behaviour. In moving towards a known tourist attraction, the tourist indicates a possible interest in it and the Genie infers that they would welcome some information on it. If the tourist confirms this implicitly by continuing their trajectory and encountering the attraction, this serves as a confirmation to the Genie to display a presentation. This it does, and by convention, briefly introduces the tourist to the attraction in question. Should the tourist walk away or ignore the presentation, their implicit interaction is interpreted as having no interest in

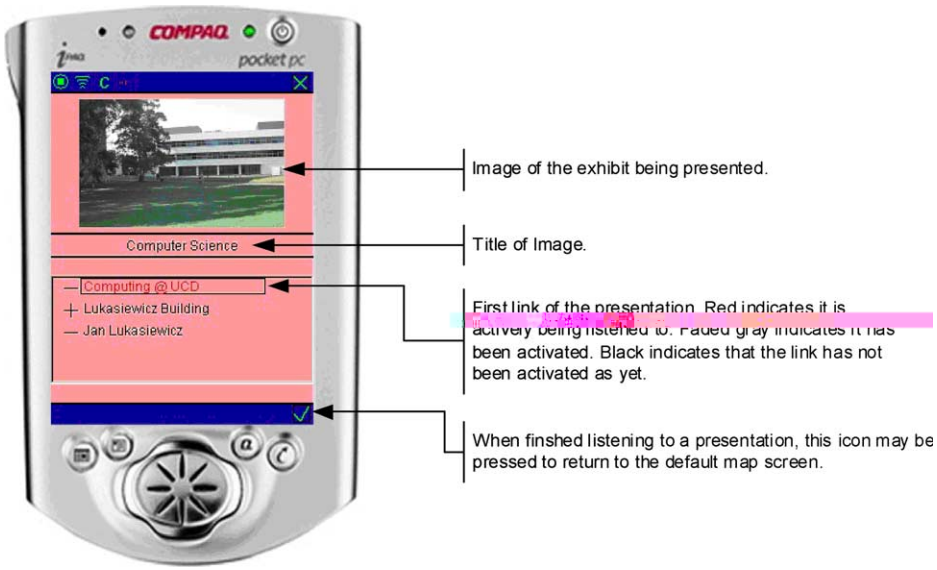

Fig. 4. An illustration of the standard Genie presentation screen. In this case the tourist is listening to a presentation about the Computer Science Department and has the option of listening to a presentation about Jan Lukasiewicz, a famous theoretical computer scientist.

the attraction in question and the Genie automatically restores the default navigation screen. However, if they use the stylus to interact with the presentation, they are explicitly interacting with the Genie and are indicating certain preferences by what they explicitly chose to hear and, implicitly, by what they chose to ignore. Interpreting such interactions is fundamental to the refinement of the user model.

Intelligent agents are fundamental to the realisation of the implicit interaction modality and some of their inherent characteristics make them particularly suitable for this task. Their autonomous nature is essential for the ongoing monitoring of tourist activities. They proactively seek to anticipate the tourist's future behaviour and arrange the availability of multimedia presentations in a just-in-time basis. Furthermore, should the tourist desire to explicitly interact with the Genie, the reactive nature of the agents enables them to also cater for this scenario.

4. Evaluation

The evaluation was undertaken on the campus of University College Dublin (UCD) during late autumn. Weather conditions were cold but dry and the campus itself was somewhat hectic as normal college activities were ongoing.

4.1. Procedure

For this particular evaluation, 40 subjects actively participated. Each participant in the trial was first asked to complete a profile questionnaire. As well as some standard personal

data, it elicited information concerning the subject's behaviour and experience as a tourist as well as their knowledge and expectations of technologies such as PDAs, GPS and so forth. After registering their interests with the Genie, each subject was given a PDA and asked to make a tour of the campus in any direction and to follow any path of their choosing. An observer accompanied each subject during the tour, which lasted on average about 40 min. The purpose of the observer was twofold: to ensure that the subject actually understood what was expected of them and to verify that the subject had in fact completed the necessary tasks so that they could complete the questionnaires. Once the observer was satisfied that the subject understood what was expected, the subject was asked to make a brief tour of the campus. The observer did not interfere in this process, unless they suspected the system might have crashed, but recorded their observations at the end of the evaluation. While potentially limiting the effectiveness of the evaluations to some degree, for example, the subject may have felt more confident with somebody obviously on hand, or alternatively, more nervous knowing they were being observed, it was considered that the benefits accruing from ensuring the evaluation was conducted adequately outweighed the possible limitations for this particular evaluation. On completion, the subject was asked to complete a usability questionnaire. Finally, all participants were invited to freely express their opinions of the Genie, including what worked, what did not and what features they would like to see included together with any other observations they would care to make.

4.2. Goals of the usability study

International standards relating to usability delineates three conceptual dimensions of *usability* (ISO, 1997):

- *Effectiveness* as 'the accuracy and completeness with which specified users can achieve specified goals in particular environments.'
- *Efficiency* as 'the resources expended in relation to the accuracy and completeness with which the users achieve goals.'
- *Satisfaction* as 'the comfort and acceptability of the work system to its users and other people affected by its use.'

There are two modalities of measuring usability (Nielsen, 1999): *objective methods* consisting of performance measurement, in terms of users' behaviour for accomplishing tasks, and *subjective methods* consisting in users' attitude measurement, regarding their interaction with the system.

This study is primarily an exploratory one and its main goal consists of employing subjective methods for investigating the way in which potential users would receive such a system and perceive its usability. No performance criteria were designed for this study, though there was scope for a number of objective measurements of the Genie's performance. A further iteration of the Genie's design will incorporate the results of this study, after which a new evaluation is planned. It is expected that the device configuration used for this future evaluation will be more reflective of that which a future tourist will possess. In particular the use of 3G telecommunications technologies is envisaged thus improving the raw computational performance of the Genie. Thus, it is anticipated that

a number of objective performance measurements will be carried out, leading to more meaningful results which will be used for augmenting the Genie still further.

Gabbard and Hix (1997) identified the following user characteristics which lead to individual differences on the perceived system usability:

- user experience;
- age;
- gender;
- technical aptitudes.

User prior experience with a specific type of task leads to the acquisition and training of a set of skills required by the tasks thus impacting on their completion. The direct benefit of this is increased performance for particular types of tasks. Traditional computer experience and, in particular, previous exposure to PDAs will potentially impact on the perceived usability. As Waller (2000) noted, the level of prior experience in working with computers and feeling comfortable with them could have a significant impact on computer task performance. How this general computer experience can be transferred to the usage of an electronic tour guide is an open question. Because all our subjects had prior experience of working with computers, and given the system characteristics, we hypothesised that previous experience with technology and wireless communication would impact on system usability. In particular, any previous exposure to PDA systems may impact on the perceived usability, since it may shape users' expectations with respect to what Genie has to offer.

Both age and gender are user characteristics often mentioned as impacting on task performance and experienced level of satisfaction, particularly in the context of spatial tasks (Lawton, 1994; Waller, 2000). Most of the study participants were rather young (65% younger than 30 years of age, while the remaining 35% being younger than 50 years of age). No valuable age-based hypothesis can be tested on this sample. Age-based impact on system usability remains however, an idea to be investigated in further studies. This is of particular interest given the increased segment of elderly tourists who could significantly benefit from such travel guides. Gender is considered a predictor variable, which clearly impacts on spatial tasks. Studies indicate that, generally, these differences tend to favour males (Lawton, 1994). Men appear more accurate at pointing at spatial cues (Montello and Pick, 1993; Kozlowski and Bryant, 1977; Cornell et al., 2003). In addition, there is a difference in the attitude towards technology which appears to be socially learned and favours males as well (Campbell and McGabe, 1984). To conclude, there are several gender-related individual differences in spatial cognition (Self and Golledge, 2000) and attitude towards technology and it seems reasonable to expect a gender impact on the perceived experience, with males experiencing probably a higher level of satisfaction with the system.

Given the context of its use, any electronic tourist guide relies to a certain degree on a set of technical aptitudes (i.e. spatial orientation, spatial memory, and spatial visualisation) which support performance on navigational tasks (Egan, 1988). For the purpose of this work, we consider sense of direction. This construct covers different meanings such as the ability to identify and locate landmarks, the ability to develop and use strategies

for learning routes, or the ability to perform well with navigation and orientation tasks (Cornell et al., 2003). First attempts to measure sense of direction belong to Kozlowski and Bryant (1977) who used one question: 'How good is your sense of direction?' on a Likert scale. Like any self-rating questionnaire, this measurement tool is based on the intuitive understanding of what sense of direction is and on people's ability to assess it. Cornell et al. (2003) investigated whether people's ability to self-rate their sense of direction relates to their ability to orient and find their way in large-scale environment. Their findings suggest that the self-rating of sense of direction correlated negatively with the mean absolute error of pointing, and positively with recall of direction of travel and landmarks, scenes recognition along the route, correct choices of paths during route reversal, and latency of pointing to the ends of the routes. Within this study, sense of direction was likewise measured through a self-rated question. Based on these theoretical aspects, the following hypotheses have been formulated and tested through the evaluation study.

H1 Users having previous experience with technology and wireless communication would find the Genie easier to use.
H2 Users with a better sense of direction would be more satisfied with the spatial feedback of the Genie than those with a below-average sense of direction.
H3 Gender dimensions would shape the user's experience.
H4 Users would perceive a need for such guides.
H5 Previous experience using maps would impact upon satisfaction with the system.

4.3. Analysis of the participant's profile

Three aspects of each participant's profile were obtained. The first included standard factual data. The second aspect of interest was their behaviour and experience as tourists. For example, details of how information concerning tourist attractions was obtained, when such information was obtained and the frequency and type of guides used during a visit were ascertained. Finally, we tried to quantify their experience of mobile computing, telecommunications and GPS as well as assessing their understanding and expectation of such technologies. Our reasons for collecting all this data was to identify the impact of these independent variables upon the level of the subject's satisfaction with the Genie as well as identifying any prior expectations they might possess.

4.3.1. Results

The sample group consisted of 40 subjects, including students, employees and visitors to the campus. Among them, 56% were males and 44% were females, with 65% of the sample group being younger than 30 years of age. Our findings indicate that 66% of participants regularly collect information about their destination prior to visiting it. The preferred type of guide is a book, used by 88% of the participants, while a human guide is used by 41%. Indications are that women tend to use human guides somewhat more than men. When we analyse the temporal dimensions of guide usage, we see that the most important is the present with 85% of participants using guides during the visits, followed by the past with 58% of participants obtaining a guide before the visit. When we

Table 1
Concept operationalisation

Dimensions	Variables
D1. Satisfaction with Gulliver's Genie in its entirety	Interaction, intuitive interface, system responsiveness, enriching visiting experience
D2. Satisfaction with navigation support	Position, direction, disorientation, electronic map
D3. Satisfaction with ability to meet cultural requirements	Content, format, quality, duration, relevance to position and orientation, responsiveness
D4. Satisfaction with system facilities	PDA, stylus, HELP
D5. Market potential for a system such as Gulliver's Genie	Perceived need, intention to rent, intention to subscribe, intention to buy

considered the collective experience of the sample group with computing technologies, we found that the vast majority (97%) claimed to have some experience of general computing. However, only 44% had prior experience of PDAs. Over 38% of the group had some experience of GPS while 20% had actually used GPRS.

4.4. Usability analysis

Though there are several questionnaires available for assessing software usability (Chin et al., 1988; Kirakowski and Corbett, 1993), none have been explicitly developed for a class of system such as the Genie. Thus, for the purpose of this evaluation, a questionnaire was specifically designed to measure user satisfaction with the Genie[2]. Recalling that navigation support and addressing a tourist's cultural expectations were the prime motivations for developing the Genie, Table 1 presents a brief description of the various dimensions and associated variables which encode the concept of user's satisfaction for each of these aspects of the Genie.

The questionnaire was administered immediately after the subjects had finished their tour. The questionnaire itself contained 41 items measured with a 7-point Lickert scale. In essence, the subject was given a number of questions and asked to give their response using a Likert scale where a '1' was labelled 'Not at all' and a '7' was labelled 'Completely'. For example, one question was 'To what extent were you satisfied with the system?'. If a subject found the system incomprehensible or unusable, they would be expected to score it with either a '1' or a '2'. Alternatively, if they thought the Genie was wonderful, they would presumably score it with a '7'. All questions were formulated in such a way that this scheme would seem intuitive to the subjects. Some further examples of questions, including one for each dimension, may be seen in Table 2. Other standard questions were used to obtain the subject's profile, for example, age, previous experience with technology and so forth. Some open ended questions were also included to encourage the subjects to share their views of the Genie in a less structured way.

[2] The questionnaire may be obtained upon request from the corresponding author.

Table 2
Sample questions from the questionnaire

Question	Dimension	Likert Scale						
		(Not at all)						(Completely)
		1	2	3	4	5	6	7
To what extent were you satisfied with the system?	D1	☐	☐	☐	☐	☐	☐	☐
To what extent did you become disorientated?	D2	☐	☐	☐	☐	☐	☐	☐
To what extent was the content easy to understand?	D3	☐	☐	☐	☐	☐	☐	☐
To what extent would you use an online HELP?	D4	☐	☐	☐	☐	☐	☐	☐
To what extent would you purchase such a system if it were available option?	D5	☐	☐	☐	☐	☐	☐	☐

4.4.1. Results

In assessing the system usability, we started by processing the results provided by the user satisfaction questionnaire. Using the criterion of 0.33 as the cut-off point for retaining variables, 14 items were deleted. The Cronbach's alpha coefficient, $\alpha = 0.84$, indicates the reliability of the questionnaire. The results, structured according to the dimensions and variables developed for questionnaire construction, and including the Cronbach's alpha coefficients for each of the subscales, are presented in Table 3. It should be mentioned that the 'overall satisfaction' dimension appeared not to offer a consistent picture, thus a further analysis of the item-total statistics suggested the deletion of one of its items. This resulted in a more acceptable reliability coefficient (from $\alpha = 0.46$ to $\alpha = 0.62$). The new subscale was used in the subsequent analyses.

As Table 3 shows, all dimensions imply a good level of satisfaction (recall that rank 7 means completely satisfied). Fig. 5 presents these results in a more detailed manner. Thus, 85% of participants were quite satisfied with the system in general, their average scores for items composing the dimensions of overall satisfaction being within the 5–6 range (Fig. 5a). The satisfaction with respect to the spatial orientation achieved through the system was high (within the 6–7 range), since 55% of participants were very satisfied with it (Fig. 5b). The level of satisfaction with respect to the presentation of visited exhibits was also good, 84% of participants scoring it with the 5–7 range (Fig. 5c).

Table 3
User satisfaction results

Dimensions	Median	Mean (SD)	Cronbach's coefficients
Overall satisfaction	5.0	5.0 (0.77)	0.62
Satisfaction with navigation support	5.6	5.6 (0.87)	0.60
Satisfaction with ability to meet cultural requirements	5.3	5.3 (1.1)	0.87
Satisfaction with system facilities	5.0	5.0 (1.2)	0.68

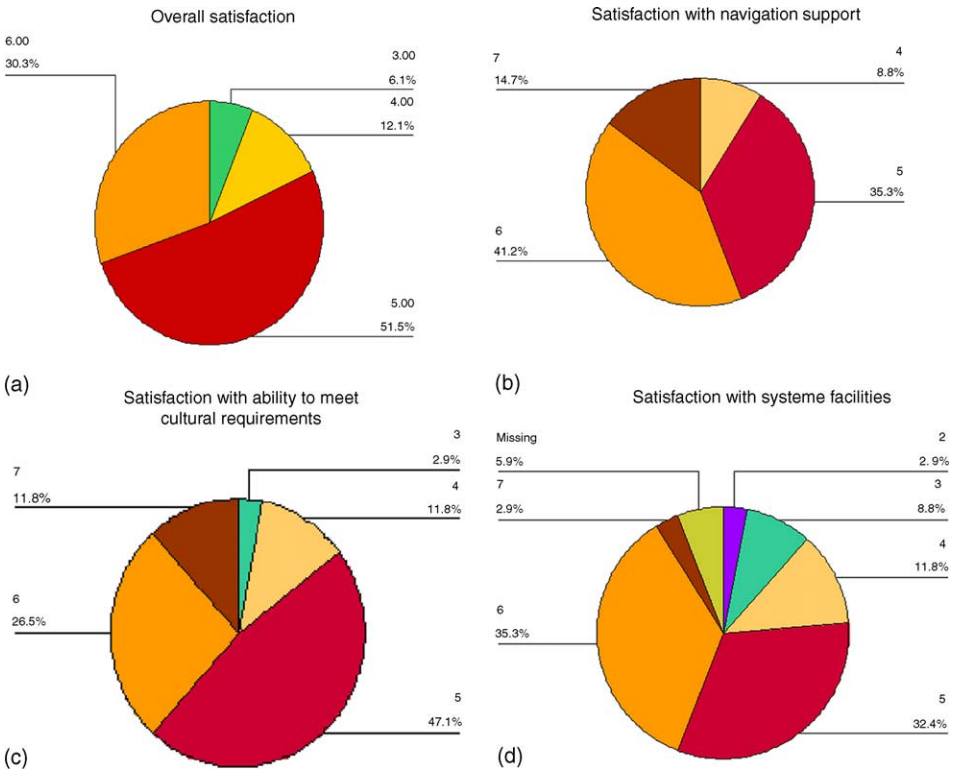

Fig. 5. Graphical depiction of the key results of the usability test.

Ultimately, more than 70% of participants were satisfied with the system functionality, scoring it within the 5–7 range (Fig. 5d).

A detailed analysis of each item suggested that aspects such as the ease with which a presentation could be understood, (Mean = 6.3), and the perceived coherency of the presentations (Mean = 6.2) are those which induced maximum satisfaction. On the other hand, the minimum level of satisfaction (just about average) was associated with the system responsiveness when using stylus (Mean = 4.06), with the interacting via the stylus (Mean = 4.24) and with the system responsiveness during the presentation (Mean = 4.15). Though not registering as serious deficiencies, addressing these issues is one obvious way of improving the Genie's usability.

4.5. Hypotheses revisited

Having analysed the results, we were then in a position to test our working hypotheses.

Hypothesis HI. Users having previous experience with technology and wireless communication would find the Genie easier to use.

We found a significant impact of previous technological and wireless communication experience upon the perceived 'ease-of-use' of the Genie. Positive correlations (some significant, some marginally significant) were found between the perceived ease of use and users' previous experience with computers ($r(39)=0.45$, $p<0.01$), experience in using PDAs ($r(39)=0.32$, $p=0.01$), experience in using GPRS ($r(39)=0.33$, $p=0.05$), experience in using a stylus ($r(39)=0.33$, $p=0.06$) and with some experience in using the older GSM data communications facility—HSCSD ($r(39)=0.34$, $p=0.05$). These findings suggest that the system appears significantly easier to use for people having experience with computers, PDAs, stylus, GPRS or HSCSD. This indicates a positive transfer of knowledge and skill previously acquired during interaction with such kinds of technology. This leads to a sense of comfort and greater satisfaction with the Genie system. In conclusion, the needs of users with limited experience of technology should be specially addressed, possibly through the utilisation of more adaptive interfaces.

In addition, the visiting experience was significantly enriched for subjects who had no previous experience with PDA devices ($t(2, 34)=3.85$, $p<0.05$) (Fig. 6). A possible reason for the high satisfaction rating amongst those with little PDA experience could be the novelty of using cutting-edge technologies. In contrast, experienced users may have some pre-conceived ideas about how the Genie should operate and find the concept of implicit interaction radically different from how they normally interact with a PDA. A solution to this might be to provide more facilities for explicit interaction thus allowing such users to feel more in control of the Genie.

Hypothesis H2. Users with a better sense of direction would be more satisfied with the spatial feedback of the Genie than those with a below-average sense of direction.

This hypothesis suggests a positive relationship between users' sense of direction and their satisfaction with respect to the spatial orientation achieved during the visit.

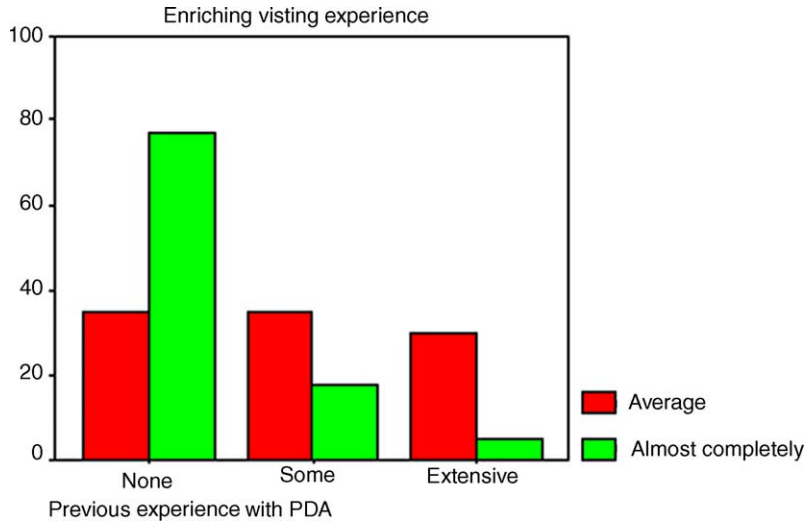

Fig. 6. Those with no prior experience with PDAs found the experience was enriching.

Users were asked to rank their sense-of-direction as poor, average, good or excellent. Satisfaction with the spatial orientation facilities offered by the Genie was obtained from the usability questionnaire. We found that satisfaction with respect to the general spatial orientation acquired and maintained during the evaluation was greater for those subjects having a sense of direction above average ($t(35)=2.46$, $p<0.01$). A number of other items were identified which contributed to this significant difference. For example, matching the real world landmarks against those on the map was significantly easier for subjects with a sense of direction above average ($t(35)=3.36$, $p<0.01$) while satisfaction with the level of detail offered by the map was significantly higher for such subjects ($t(35)=2.22$, $p<0.01$). While confirming this hypothesis, it indicates that the Genie requires further refinement for those users whose spatial abilities are below average. For example, icons representing landmarks could be designed more realistically and the use of 3D images potentially considered. A facility for rotating such images could also be investigated. Likewise, rotating the map automatically so that it consistently coincides with the tourist's orientation may also provide some benefit.

Since any previous exposure to the university campus, where the evaluation study took place can impact on spatial feedback needed for orientation, the analysis should take this variable into account. Therefore, we investigate the combined impact of familiarity with the campus and sense of direction on users' satisfaction with the spatial orientation enabled by the Genie. Without being significant, ANOVA analysis suggest a moderate interaction effect between familiarity with the campus and sense of direction ($F(4, 38)=2.14$, $p=0.08$).

In the case of users with an average sense of direction, the more familiar with the campus, the more satisfied with the spatial orientation enabled by the system. In contrast, the familiarity with campus of users with a good or excellent sense of direction does not lead to any increased satisfaction with spatial orientation provided by the system (Fig. 7). In other words, the system support for navigation is highly acknowledged by those unfamiliar users whose sense of orientation is above the average or by familiar users with a below average sense of direction. Both categories of users rely on different factors for orienting, such as a good sense of direction or the familiarity with the campus. Not surprisingly, the lowest satisfaction comes from users with average or below average sense of direction and little knowledge of the campus. Since this segment represent almost 25% of users in the study sample, future work should be carried out in order to investigate ways in which the system can address this issue.

Hypothesis H3. Gender dimensions would shape the user's experience.

We found that during their visits, women use maps significantly more than men ($t(35)=2.03$, $p<0.05$). Likewise, women use a guide of any kind significantly more than men do ($t(35)=2.65$, $p<0.05$) thus validating also the hypothesis H5. As gender is considered an independent variable with a clear impact upon spatial tasks, studies indicating that generally these differences favour men (Lawton, 1994; Montello and Pick, 1993; Self and Golledge, 2000), this result is not surprising. It places an onus on the Genie to ensure that its navigation features are particularly oriented towards the female group. Without reaching significance, males appear to be slightly more satisfied than

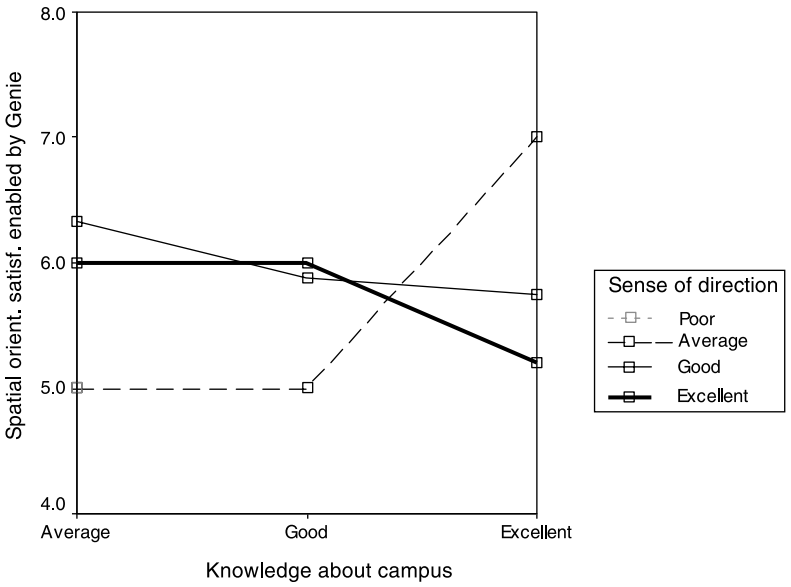

Fig. 7. User satisfaction with the spatial orientation enabled by Genie, as a function of user sense of direction and familiarity with the campus.

females with the Genie usability. Looking at the individual items, the findings suggest that females rotate the PDA significantly more in order to match the map's orientation with their own ($t(35)=2.09$, $p<0.05$), while males found the system easier to use ($t(35)=2.07$, $p<0.05$). Without reaching significance, males consider the system more intuitive and are more inclined to purchase such a device, while females consider that Genie made their experience during the experiment slightly richer.

Hypothesis H4. Users would perceive a need for such guides.

With the exception of the item concerned with the intention of buying a system like the Genie, all the other items, that is, the perceived need for such a system, the intention to use such a system if one were available and the intention to rent such a system have a mean above 5.5 (Fig. 8). We performed one-sample *t*-tests, testing the hypothesis that our sample was not drawn from a population with a mean of 5. The findings indicate that there is a significant difference between the means for these items obtained by the subjects from our sample and a theoretical population for which the means of these items is 5. Thus we obtained significant difference for the perceived need for such a system ($t(36)=6.65$, $p<0.01$), the intention to reuse the system if it were available ($t(36)=4.76$, $p<0.01$) and the intention to rent the system ($t(36)=2.37$, $p<0.05$). In other words, the market potential of the system is significantly higher than 5, which on a scale from 1 to 7 is already high. This result validates the hypothesis that the users perceive a need for an electronic tourist guide.

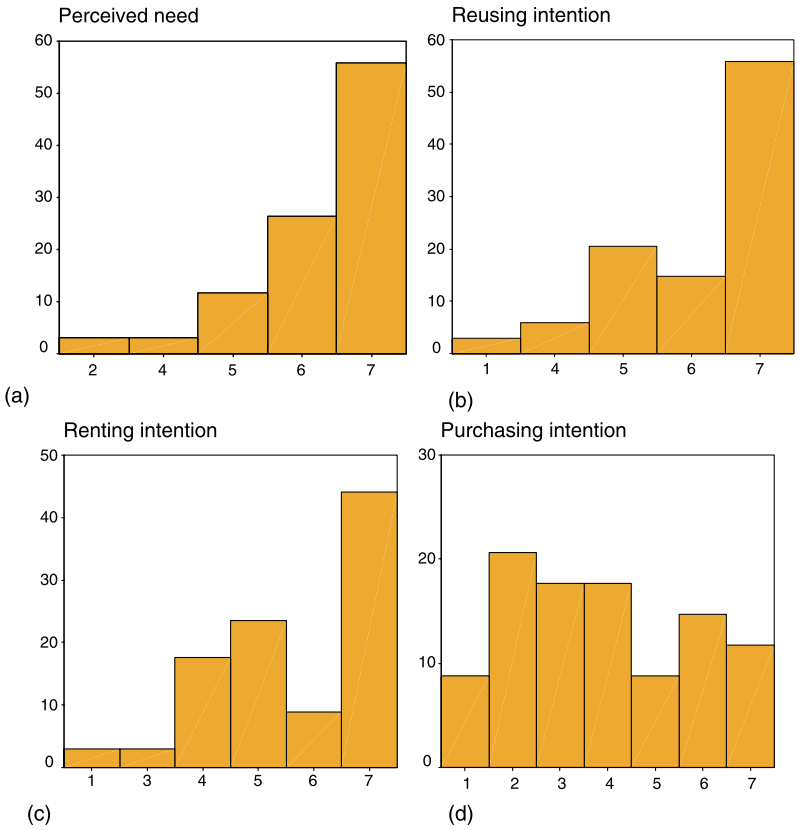

Fig. 8. The sample distribution of those items measuring the market potential of the system.

Hypothesis H5. Previous experience using maps would impact upon satisfaction with the system.

Users were asked to quantify their use of maps when visiting new places in terms of: always, often, sometimes or never. Satisfaction with respect to the extent with which the user's position on the electronic map coincided with their position on the ground was significantly smaller for users who use maps on a regular basis as distinct from those who use maps sparingly ($t(35) = -6.97$, $p < 0.01$). This validates the hypothesis. To conclude, users more experienced in using maps have higher expectations regarding system accuracy in delivering proper and timely position updates while using the Genie.

4.6. Subjective user analysis

The usability questionnaire ends with three open questions regarding system strengths, limitations and suggestions for improvement. We present the results obtained through a qualitative content analysis of these answers. Users' answers were grouped within clusters

based on their similarity. This post-codification process was carried out independently by two of the authors, and the correlation between the number of answers within each cluster was statistically significant ($r(12)=0.94$, $p<0.01$). We present the clusters, together with the frequency of their appearance, expressed as a percentage of the total answers within each of the considered topics, that is, strengths, weaknesses and improving suggestions. These figures are an indicator of the prevalence of the associated clusters within the previously mentioned topics.

4.6.1. Strengths

The system's strengths as the users identified them coincide with the two essential issues that the Genie was designed to address. Navigation support was identified as the Genie's most important feature (51%). Ensuring that users always know their current position through the 'you are here' icon helps in maintaining good orientation and position awareness. In addition, it also affords a psychological comfort to some, thus facilitating more adventurous exploration by reducing the anxiety of getting lost in an unfamiliar area. The next strength consists of the Genie's user interface (30%). As participants observed, it is easy to use, interactive and intuitive. As such, users perceived that the Genie could indeed form the basis of an electronic pocket guide or perhaps serve as a suitable substitute for a traditional guidebook. The third strength identified was the potential of the Genie to deliver information about relevant attractions at the right place and time (19%). This also verifies the relevance of the content delivered and the adequacy of the presentation format.

4.6.2. Limitations

The weaknesses of the Genie as perceived by the users was an important goal of this usability study as an awareness of them is essential if the system is to be improved. Responsiveness was the most frequent limitation identified (62%). Users felt that delays in updating their position could lead to confusion. It was also pointed out that the downloading of presentations was occasionally slow as presentations were displayed after the attraction in question had been passed. Another weakness consists of the interaction via the stylus (21%). Some users found it difficult to use thus reducing the perceived responsiveness of the Genie. In the case of the multimedia presentations, some users claimed that the volume was not always adequate (10%). Finally, some users observed that the electronic map was over simplified and could be usefully augmented with other relevant features (7%).

4.6.3. Suggestions for improving

A recurring suggestion for system improvement was to make the Genie faster and more responsive (44%) as well as providing some additional feedback indicators on the Genie's activity. Regarding navigation, several users thought the level of detail on the map should be increased, as should the frequency of the position and orientation update (30%). It was also suggested that the map be made more dynamic through the availability of a zoom function and, more interestingly, through the automatic rotation of the map to match the user's current orientation.

A small number of users (17%) suggested that before a presentation is downloaded, the Genie should highlight the attraction in question and provide a summary before asking

the user if they wish to download the full presentation. In doing this, they felt they would have more control of the Genie and could decide whether they wanted to view the presentation or not. In a similar vein, it was suggested that such a feature should be available for all attractions as this would aid the tourist decide whether they wanted to visit it or not and, in this way, conserve their energy and use their limited time more effectively. Surprisingly, a number of users observed that audio was not adequate on its own and would be better if could be augmented by an equivalent scrolling text commentary. Circumstances where this would be useful would include situations where there was a lot of background noise. Interestingly, such a facility would also be beneficial to people who might have hearing difficulties. With respect to the problems encountered when using the stylus (9%), it was suggested that the standard buttons on the PDA be used instead. In which case, the Genie could be operated using only one hand, a scenario that a few users thought desirable, as one observed, 'hands are a scarce resource for a tourist'.

5. Conclusion

This paper has introduced Gulliver's Genie and depicted the system's operation and the associated user experience. We have presented and analysed the results of a detailed user evaluation conducted by subjects on the campus of our university. Specifically we used these results to establish the validity or otherwise of an initial set of working hypotheses. Prior literature reporting detailed user evaluations of context-aware PDA hosted systems is scant and this paper seeks in part to address this gap.

This research constitutes a case study for which the results are merely preliminary. Study limitations consist of the non-random sampling procedure for selecting participants and the relatively small sample size, which reduce the generality of the obtained findings. However, the strength of study outcomes does not reside in their generalisation power but rather in their exploratory potential for identifying questions (Lynn, 1991). They also suggest how a computer-based system for navigation support is received by users in terms of perceived usability.

Based upon the user evaluations, several enhancements to the system present themselves. In terms of the general usability analysis, whilst user satisfaction overall was high, several improvements were identified. In terms of system responsiveness, the stylus as an interaction device seemed less favoured with users touching the screen multiple times, not holding the stylus perpendicular to the screen or not contacting the screen region sufficiently long for the interface to register the user choice. Thus an interface revision has been designed to support user selections via the navigation pad. System responsiveness in terms of content retrieval from the server is largely constrained by the bandwidth of the network. While GPRS offers advantages over 2G, only with the rollout of 3G networks will significant improvements occur. The system currently offers intelligent pre-caching of content, and perhaps, this needs to be augmented in some way.

The hypotheses similarly provide insights as to system enhancements. In particular, the importance of adaptive user interfaces is confirmed, reflecting the needs and perceived needs of constituent user groups. For example, users who are less technologically experienced may need a more supportive interface, which tracks their system interaction

proficiency and adapts accordingly. Female users may need enhanced navigational support with perhaps a map rotation or alignment feature to ensure that at all times the map displayed is relative to their current direction of movement. Furthermore, users would like the possibility of a richer set of interactions supported within the explicit interaction modality, demanding more user-directed activities rather than the predominance of system-directed activities. Prudent use of the implicit interaction modality may improve end-user satisfaction provided that the issue of perceived user control is considered. A suitable strategy may be to allow users to dictate their preferences by allowing them to specify default system behaviours and allowing them to opt in or out as they deem appropriate.

Overall the evaluations were broadly supportive and users did perceive a need for Gulliver's Genie as well as exhibiting a favourable disposition to incur costs for its usage. As the essential functionality, namely intelligent pre-caching, customised content delivery, adaptivity and user profiling, was realised using intelligent agents, we feel our adoption of an agent-based approach has been vindicated.

Acknowledgements

Michael O'Grady gratefully acknowledges the support of the Irish Research Council for Science, Engineering and Technology (IRCSET) though the Embark Initiative postdoctoral fellowship programme. Gregory O'Hare acknowledges the support of Science Foundation Ireland through Grant No. 03/IN.3/1361.

References

Abowd, G., Atkeson, C., Hong, J., Long, S., Kooper, R., Pinkerton, M., 1997. Cyberguide: a mobile context-aware tour guide. ACM Wireless Networks 3 (5), 35–41. ACM.

Bellotti, F., Berta, R., De Gloria, A., Margarone, M., 2002. User testing a hypermedia tour guide. Pervasive Computing 1 (2), 33–41. IEEE.

Buester, G., Thomas, B., Wolff, C., 2000. MIA—a ubiquitous multi-agent web information system, Proceedings of International the International ICSC Symposium on Multi-Agents and Mobile Agents Symposium in Virtual Organizations and E-Commerce (MAMA'2000), Wollongong, Australia 2000.

Campbell, P., McGabe, G., 1984. Predicting the success of freshman in a computer science major. Communications of the ACM 27, 1108–1113. ACM.

Cheyer, A., Julia, L., 1995. Multimodal maps: an agent-based approach, Proceedings of the International Conference on Cooperative Multimodal Communication (CMC'95), Eindhoven, The Netherlands, May 1995, vol. 1374. Springer LNCS, Berlin pp. 111–121.

Chin, J.P., Diehl, V.A., Norman, K.L., 1988. Development of an instrument measuring user satisfaction of the human–computer interface, Proceedings of the SIGCHI conference on Human factors in computing systems, Washington DC, USA. ACM, New York pp. 213–218.

Collier, R.W., O'Hare, G.M.P., Lowen, T.D., Rooney, C.F.B., 2003. Beyond prototyping in the factory of agents, Proceedings of the Third International Central & Eastern European Conference on Multi-Agent Systems (CEEMAS 2003), Prague, Czech Republic. Springer LNCS, Berlin.

Cornell, E.H., Sorenson, A., Mio, T., 2003. Human sense of direction and wayfinding. Annals of the Association of American Geographers 93 (2), 399–425.

Davies, N., Cheverst, K., Mitchell, K., Efrat, A., 2001. Using and determining location in a context-sensitive tour guide. Computer 34 (8), 35–41. IEEE.

Egan, D., 1988. Individual differences in human–computer interaction, in: Helander, M. (Ed.), Handbook of Human–Computer Interaction. Elsevier Science, Amsterdam, pp. 543–568.

Gabbard, J., Hix, D., 1997. Taxonomy of usability characteristics in virtual environments. Final Report to the Office of Naval Research. Virginia Polytechnic Institute, Blackenburg.

ISO (International Organization for Standardization), 1997. ISO 9241: ergonomic requirements for office work with visual display terminals. Part 11: guidance on usability specification and measures.

Kinny, D., Georgeff, M., Rao, A., 1996. A methodology and modelling technique for systems of BDI agents, Proceedings of the Seventh European Workshop on Modelling Autonomous Agents in a Multi-Agent World (MAAMAW'96), LNAI, Eindhoven, The Netherlands, vol. 1038. Springer, Berlin, pp. 56–71.

Kirakowski, J., Corbett, M., 1993. SUMI: the software usability measurement inventory. British Journal of Educational Technology 24 (3), 210–212.

Kozlowski, L., Bryant, K., 1977. Sense of direction, spatial orientation, and cognitive maps. Journal of Experimental Psychology 3, 590–598.

Lawton, A., 1994. Gender differences in way-finding strategies: relationship to spatial ability and spatial anxiety. Sex Roles 30, 765–779.

Lynn, D., 1991. The application of case study evaluations. ERIC/TM Digest 71, ERIC.

Marmasse, N., Schmadt, C., 2000. Location-aware information delivery with commotion, Proceedings of the Second Symposium on Handheld and Ubiquitous Computing (HUC 2000), Bristol, UK. Springer LNCS, Berlin pp. 157–171.

Montello, D.R., Pick, H.L., 1993. Integrating knowledge of vertically-aligned large-scale spaces. Environment and Behavior 25, 457–484.

Nielsen, J., 1999. Designing Web Usability. New Riders Publishing, Indianapolis.

O'Grady, M.J., O'Hare, G.M.P., 2004. Just-in-time multimedia distribution in a mobile computing environment. IEEE Multimedia 11 (4), 62–74. IEEE.

O'Hare, G.M.P., O'Grady, M.J., 2003. Gulliver's Genie: a multi-agent system for ubiquitous and intelligent content delivery. Computer Communications 26 (11), 1175–1185. Elsevier Press.

Poslad, S., Laamanen, H., Malaka, R., Nick, A., Zipf, A., 2001. Crumpet: creation of user-friendly mobile services personalised for tourism, Proceeding of the Second IEE International Conference on 3G Mobile Communication Technologies, London, UK 2001.

Self, C.M., Golledge, R.G., 2000. Sex, gender, and cognitive mapping, in: Kitchin, R., Freundschuh, S. (Eds.), Cognitive maps: past present, and future. Routledge, London, pp. 197–220.

Waller, D., 2000. Individual differences in spatial learning from computer-simulated environments. Journal of Experimental Psychology: Applied 8, 307–321.

Wooldridge, M., Jennings, N., 1995. Intelligent agents: theory and practice. The Knowledge Engineering Review 10 (2), 115–152.

Youll, J., Morris, J., Krikorian, R., Maes, P., 2000. Impulse: location-based agent assistance, in software demos, Proceedings of the Fourth International Conference on Autonomous Agents, Barcelona, Spain 2000.