Title	Just-in-time multimedia distribution in a mobile computing environment
Authors(s)	O'Grady, Michael J.; O'Hare, G. M. P. (Greg M. P.)
Publication date	2004-10
Publication information	IEEE Multimedia, 11 (4): 62-74
Publisher	IEEE
Item record/more information	http://hdl.handle.net/10197/4423
Publisher's statement	Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
Publisher's version (DOI)	10.1109/MMUL.2004.30

O'Grady, M. J. & O'Hare, G. M. P., Just-In-Time Multimedia Distribution in a Mobile Computing Environment, IEEE Multimedia, Vol 11, No. 4, pp. 62-74, 2004.

Just-In-Time Multimedia Distribution in a Mobile Computing Environment

M. J. O'Grady & G. M. P. O'Hare

Practice & Research in Intelligent Systems & Media (PRISM),
Department of Computer Science,
University College Dublin (UCD),
Belfield,
Dublin 4,
Ireland.

{ michael.j.ogrady, gregory.ohare}@ucd.ie

O'Grady, M. J. & O'Hare, G. M. P., Just-In-Time Multimedia Distribution in a Mobile Computing Environment, IEEE Multimedia, Vol 11, No. 4, pp. 62-74, 2004.

Just-In-Time Multimedia Distribution in a Mobile Computing Environment

Abstract

Disseminating multimedia content to users in a mobile computing environment such that they receive it in an appropriate and timely manner is fundamental to the success of mobile information systems. Too often, however, this endeavour is hindered by the poor data rates supported by wireless telecommunications networks and by the limited computational resources available on mobile devices. We describe an approach to overcome these limitations, which is based on extremely dynamic and proactive precaching. This approach, which we have termed intelligent precaching, is realised through the innovative deployment of intelligent agents on mobile devices. To illustrate this concept, the design and implementation of an archetypical mobile computing application is provided, namely that of an electronic tourist guide. This is augmented with a description of the salient points derived from a user evaluation, from which emerging avenues for further research are identified.

Keywords:
Intelligent agents, distributed multimedia, intelligent precaching, mobile computing, context-sensitive service delivery

Mobile computing is widely anticipated to become the next major computer usage paradigm in the coming years. Such a paradigm shift will have profound implications for practically all aspects of the computer industry. Quite how hardware and software will evolve to address this new scenario remains to be seen. Not least amongst the issues that must be addressed is how to effectively distribute information to such users; information that may very well have a substantial multimedia component.

The physical environment is by its very nature extremely heterogeneous. Likewise, the electronic infrastructure available to mobile users can differ considerably even within small geographic regions. Catering for the complexity and differences inherent in such environments is a critical task facing those seeking to deploy applications and services to mobile users. One particular example that illustrates this heterogeneity is the varying quality and range of data communications services available to mobile users at present. Enabling the timely delivery of multimedia information to mobile users under such constraints is the primary focus of this article.

Mobile Computing

Mobile Computing has evolved rapidly over the last decade, aided by the explosive growth in wireless telecommunications. A popular vision of mobile computing involves a small handheld device of the PDA genre, connected to a fixed network via a wireless connection. Though this is the predominant vision at present, a number of alternatives have been proposed:

Location-aware Computing
By incorporating a location-sensing component, for example GPS, into a PDA and augmented it with a wireless network, location-aware computing becomes a possibility. Only information that is relevant to the user's location is presented. A classic example is that of an electronic yellow pages that sorts its entries according to their distance from the user. Indeed, location-aware computing has grasped the imagination of many in the business community and the market for such services is expected to grow substantially in the coming years. However, how best to realise location-aware computing remains the focus of much attention [1].

Ubiquitous Computing
The Ubiquitous Computing concept was first articulated by the late Mark Weiser just over a decade ago[2]. This envisages an environment saturated with sensors and other computing devices which the user can interact with in a graceful and seamless manner. This was a radical vision and one that cannot as yet be implemented in its totality. The term Pervasive Computing[3] is frequently used interchangeably with the term ubiquitous computing. In each case, the core concept is simple: the user can access the necessary computer resources anywhere and at anytime.

Wearable Computing

Wearable computing[4] envisages the user actually wearing the computer about their person. Philosophically, it is diametrically opposed to ubiquitous computing in that it envisages users having all the necessary computational resources with them.

References

1. C.A. Patterson, R.R. Muntz and C.M. Pancake, "Challenges in Location-Aware Computing," IEEE Pervasive Computing, vol. 2, no. 2, April-June 2003.
2. M. Weiser, "The Computer of the 21st Century", Scientific American, vol. 265, no. 3, September, 1991.
3. M. Satyanarayanan, "Pervasive Computing: Vision and Challenges, IEEE Personal Communications, vol. 8, no. 4, August 2001, pp. 10-17.
4. S. Mann, "Wearable Computing as Means of Personal Empowerment", 1st International Conference on Wearable Computing (ICWC'98), Fairdax, Va, USA, 1998.

Mobile Cellular Telecommunications

These are interesting and turbulent times in the history of mobile cellular telecommunications. Despite being one of the success stories of the last decade, cellular networks were, and to a greater degree, remain severely constrained in their ability to handle wireless data. In an effort to address this, the International Telecommunications Union (ITU) launched their IMT2000 initiative[1] which defined the specifications for so called Third Generation (3G) networks. Amongst other requirements, the ITU ordained that 3G compliant networks should support data rates of:

- 144 kb/s for user moving quickly e.g. vehicles;
- 84 kb/s for pedestrians;
- 2 Mb/s for users in a low mobility or office environment.

If we consider the issue of multimedia dissemination in such an environment, the situation appears quite promising, at least initially. On closer examination, a number of problems arise. First of all, the data rates that a user can expect will vary according to the user's context. Obviously, the difference between 144 kb/s and 2 Mb/s will have a significant impact on the end user experience and must be considered when planning services for mobile users. This issue is further complicated when normal operating conditions and network operator policies are factored into the equation. The spectrum allocated to operators is usually divided into channels, which are in turn shared by multiple subscribers. Some channels may be reserved for voice calls as voice traffic still makes a major contribution to an operator's revenue. All of which can reduce the data rates available to subscribers considerably. Thus services that seemed quite attainable, for example, video-conferencing and online games with significant streaming media content, may not be feasible except in limited circumstances: a factor that has contributed immensely to the current turbulence in the cellular telecommunications sectors.

Most networks are currently in the process of being migrated to 3G via a series of interim technologies which are frequently termed 2.5G. Though data-rates continue to improve, they may not be adequate in meeting users' expectations. In the longer term, the issue may well be addressed through 3.5G technologies, such as High Speed Downlink Packet Access (HSDPA) [2], or possibly the proposed 4G networks that are already on the drawing board[3]. In the meantime, operators continue migrating their networks toward 3G, modestly improving the data rates available to customers as they do so. However, the effective delivery of services with a rich multimedia content is, and is likely to remain in the medium term at least, a significant obstacle to prospective application developers and services providers.

References

1. ITU-R Recommendation M.687-2, "International Mobile Telecommunications 2000 (IMT2000)", International Telecommunications Union – Radiocommunications (ITU-R). 1997.3GPP TS 25.308 Version 5.4.0 Release 5.
2. Universal Mobile Telecommunications System (UMTS); UTRA High Speed Downlink Packet Access (HSDPA); Overall Description, 2003.
3. T. Zahariadis and D. Kazakos, "(R)evolution toward 4G Mobile Communication Systems", IEEE Wireless Communications, vol. 10, no. 4, August 2003.

The Wireless Data Bottleneck

Traditionally, users have been used to high specification hardware and high speed networks so it can be something of a culture shock when they first embark on computing in a mobile context. The first obvious difference is the inferior computational power and facilities offered by state-of-the-art PDAs when contrasted with conventional workstations. A second noticeable difference is the poor data rates supported by wireless networks, in comparison with their fixed network cousins; something that is starkly illuminated when doing something as mundane as browsing the WWW. Remedying this problem continues to preoccupy the telecommunications sector. Solutions based on satellite and microwave technologies have had some limited success. However, solutions based either WiFi (IEEE 802.11) or cellular telecommunications (any 3G solution ratified by the International Telecommunications Union) offer the best hope for a solution in the medium term. Though frequently presented as competing technologies, both WiFi and 3G actually complement each other in certain areas. WiFi offers higher data rates than 3G but is only deployed in very select localities. In contrast, cellular telecommunications networks are deployed over vast regions of the world although full 3G compliant networks are only operational in Japan at present. However, most network operators are in the process of migrating to 3G albeit incrementally, so data rates are likewise increasing. Similarly, deployment of WiFi is accelerating thus increasing its geographical availability. However, 3G networks will never offer comparable data rates to WiFi (Table 1). Likewise, WiFi will never cover as wide a geographic area as cellular networks. All of which gives those aspiring to provide services to mobile users ample food for thought.

Family	Technology	Generation	Maximum Data Rate	Expected Data Rate [a]	Expected Download Time [b]
GSM	GPRS	2.5G	115 kb/s	30 kb/s	4.4 min.
	EDGE	2.5G	384 kb/s	60 kb/s	2.2 min.
	UMTS	3G	2Mb/s	300 kb/s	26 sec.
CDMA	CDMA(1xRTT)	2.5G	153 kb/s	60 kb/s	2.2 min.
	CDMA(3xRTT)	3G	2 Mb/s	300 kb/s	26 sec.
WLAN	WiFi	-	11 Mb/s	1 Mb/s [c]	8 sec.

[a] Approximate data rate, as the obtained data rate is completely dependant on the prevailing network operating conditions.

[b] Assume 1 MB file. Note that the average MP3 file would be 3 MB.

[c] data rates supported by WLAN can vary substantially according to the distance from the access point.

Table 1. Data rates that users can expect using either cellular networks or WiFi.

Intelligent Precaching: A Possible Solution

As part of our research into the practicalities of realising applications and services for mobile users, we quickly realised that the delivery of dynamic information services would prove problematic under the constraints within which the average mobile computing user would normally operate. We focused on the tourism domain and sought to investigate how information might be best disseminated to tourists as they explored some city – an endeavour a number of other research groups in differing disciplines have also engaged in. While a number of systems have been documented in the literature[1,2,] , CyberGuide[3] and GUIDE[4] are the two pioneering examples. Other notable efforts include Archeoguide[5], which is oriented towards historical reconstructions, and the hypermedia guide developed for the Costa Aquarium in Genoa, Italy[6]. In our case, we envisaged the tourist being equipped with a PDA or similar device. As navigation support would be an essential service, the use of GPS was assumed. The delivery of personalised multimedia-rich presentations on attractions that a tourist might encounter offered significant opportunities to enhance the individual tourist experience. Delivering such presentations in an appropriate and timely manner would prove challenging, particularly when the well documented computational restraints of PDAs, for example poor computational power, limited memory and so on, are considered. One obvious solution was to host the entire multimedia repository on the PDA, using some combination of Smartcard or Memory Stick technologies, for example. Though feasible, such an approach would severely curtail the possibilities for both personalisation of content and the inclusion of inherently dynamic content. In the latter case, the use of such technologies could potentially lead to circumstances where content is not revised and maintained in a systematic and diligent manner, resulting in the decay of content accuracy and, ultimately, a frustrated tourist. The other alternative was to host the content on a centralised server and access it via a wireless data network. But which wireless technology should be used? Though WiFi was successfully used by GUIDE, this would severely curtail where the tourist could operate. CRUMPET[7], a project concerned with the provision of location-aware services to tourists, had successfully used the 2.5G General Packet Radio Service (GPRS), which yields a data-rate of about 30 kb/s. However, this project concentrated on the provision of location-aware services and was not overtly concerned with services that incorporated a significant multimedia component. Nevertheless, GPRS covers a large geographic area and was therefore identified as the most appropriate technology.

Distributing information with a significant multimedia component over a GPRS network is problematic even under good network operating conditions[8,9]. To counteract this, we developed the following two step strategy:

1. To intelligently and dynamically precache information on the tourist's PDA.
   Implementing such a strategy requires the availability of a model of the tourist's environment. Using this model in conjunction with the tourist's known position and cultural interests, it is possible to derive a ranked list of potential attractions that the tourist may visit. On identifying an attraction that the tourist is likely to visit, a presentation can be requested and downloaded in a *just-in-time* basis. The idea being that when the tourist finally encounters the attraction, a presentation is available on the PDA for instant consultation. As the tourist continues to explore, the cycle must be repeated, thus requiring the process itself to be extremely dynamic and proactive. In this way, network latency is reduced thus giving the illusion of instantaneous information downloading and accessibility. However, there is a trade-off between immediate information access and the downloading of material that may be discarded later due to incorrect assumptions about the user's anticipated behaviour.

2. To minimise the amount of information sent to the tourist.
   To achieve this, we only dispatch information that is relevant to the tourist's position. Secondly, all information sent is consistent with the tourist's cultural interests thus improving the probability that they will be satisfied with the presented information while further minimising the amount of information that must be dispatched.

Precaching (sometimes called prefetching) of data is a well-known and established technique in mobile computing for handling network disconnections[10,11]. Traditionally, files might have been precached on a user's laptop using some criteria, possibly on a "most-recently-used" basis. As most laptops can currently be augmented with a wireless modem, the issue of disconnection is no longer the problem it once was. However, the recent surge of interest in location-aware services has again focused attention back on the core issue of cache management, and what invalidation and replacement strategies should be adopted when working with location-dependent data in a mobile computing environments remain the subject of much research[12,13].

While precaching has been used to great effect as a means of disguising network bandwidth limitations from users, it can also aid in obtaining maximum benefit from the limited memory resources on PDAs, thus minimising two inherent constraints of modern mobile computing systems. Indeed these two constraints are most acute when working with multimedia data, primarily due to the large size of the files involved. If a model of the tourist's immediate environment is available, along with a profile of the tourist's interests and preferences, these can provide a basis for the system to reason about what the contents of the cache should be. Furthermore, by making the precaching mechanism itself extremely dynamic, the contents of the cache on the PDA can be maintained in such a way that it is always relevant to the tourist's position. In this way, information can be made available in an almost *on demand* fashion to mobile users.

A number of factors are essential if such a precaching strategy is to be implemented successfully. Firstly, the entire process must be transparent to the tourist. This suggests a preference for an autonomous solution. Secondly, the contents of the cache must be maintained in such a way as to be both relevant to the tourist's position, and ideally, their immediate future position. In practice, the system must react to the tourist's behaviour as well as proactively seeking to "second guess" the tourist's likely future behaviour, and update the cache contents accordingly. In anticipating the tourist's possible future behaviour, a number of factors need to be reconciled. These include the features of the tourist's immediate environment, the tourist's own interests and preferences as well as the memory constraints of their device.

To achieve this, a reasoning capability is essential. Autonomy, reactivity, proactivity and strong reasoning capabilities are synonymous with intelligent agents. In addition, the inherent social ability of agents facilitates the intuitive assignment of appropriate tasks to individual agents - an attractive facility from a system engineering perspective. In this way, the agents may then collaborate to deliver the necessary service. Therefore intelligent agents form the basis of the architecture proposed in the next section. Some of the key steps in realising what we have termed intelligent precaching may be found in figure 1.

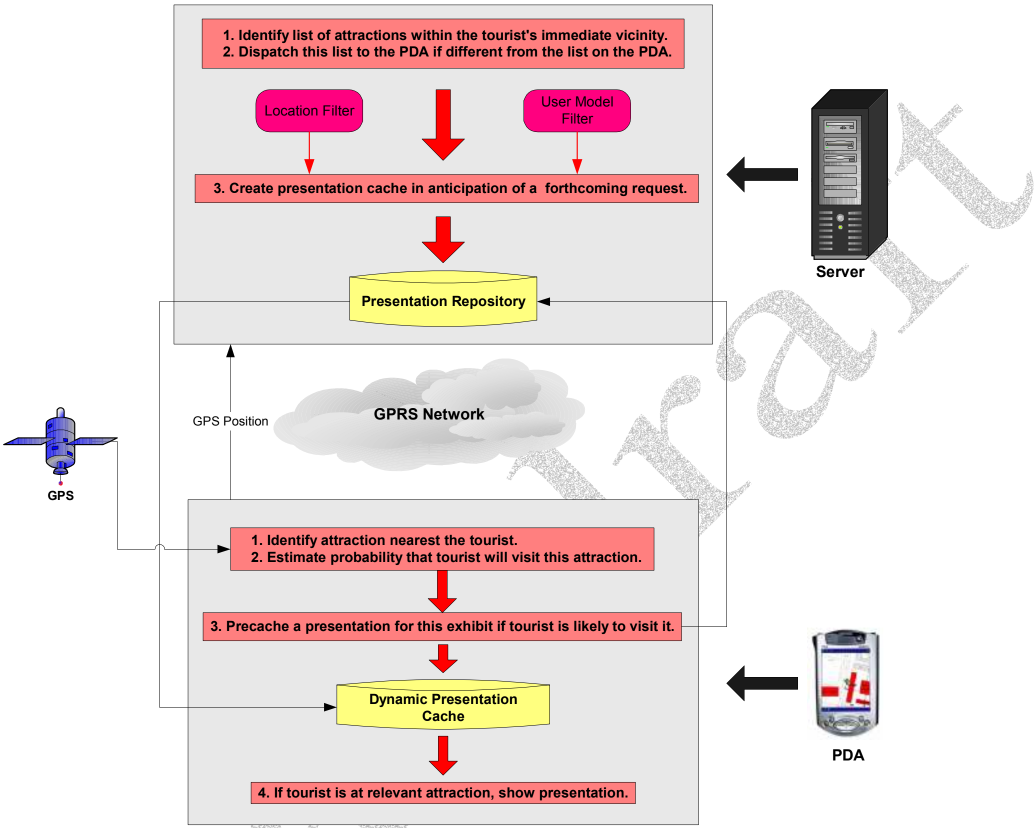

Figure 1. Essential stages of the proposed precaching strategy.

Architecture

To verify the validity of the intelligent precaching approach, we incorporated it into a system that we have developed in our research laboratory. This system, which we have dubbed Gulliver's Genie[14], is the result of research that has been ongoing for some time now into various facets of mobile computing. Though providing significant navigation support, it is the Genie's strategy for the intelligent delivery of information to tourists that is of particular interest. Tourists almost invariably seek out information of a cultural nature. The Genie's *raison d'être* is to deliver multimedia-enriched presentations to tourists on the various attractions that they encounter while exploring some city, archaeological site and so on. Such presentations consist of some combination of images, sound and video. To maximise the tourist's interest, all presentations conform to the tourist's own personal interest model. As well as being delivered in the native language of the tourist, those tourists who express an explicit interest in a certain topic, for example architecture, are supplied with additional information on that topic, assuming it is relevant to the attraction in question of course.

Gulliver's Genie conforms to a relatively straightforward client/server architecture with the tourist hosting the client component on their PDA, which in turn is connected to the server component via a GPRS connection. Intelligent Agents form the fundamental building blocks upon which the Genie is built and it may be regarded as a Multi-Agent System (MAS) encompassing agents on both the client and on the server, all collaborating to deliver the required services to the tourist in a responsive, efficient and economic basis. The architecture of Gulliver's Genie is outlined in figure 2 and its principal agents are now described:

The Intelligent Agent

Intelligent Agents offer an alternative paradigm for software development. In particular, domains that are complex and dynamic, and are hence difficult to model using traditional techniques, are potentially suitable for agents. An example of such a domain might well be the environment that the average mobile computer user operates within.

Agents come in many varieties and a number of taxonomies have been proposed. One popular taxonomy classifies agents into weak and strong varieties[1]. Weak agents may be reactive, proactive, social and autonomous. Strong agents may be rational, mobile, benevolent and veracious (truthful) as well as inheriting some of the characteristics of weak agents. A classic example of strong agents has been realised through the Belief-Desire-Intention (BDI) paradigm[2]. Very briefly, beliefs represent a snapshot of the state of the agent's environment at a given time. Desires represent those things that agents would like to do or, in other words, its reason for existing. Intentions represent those desires that the agent is in a position to realise at a given time. Intentions are formulated as a result of the agent reconciling its existing model of its environment, represented through its beliefs, with its desires. The programmer will usually specify what beliefs or conditions are necessary before an agent can commit to realising its individual desires.

Agents differ fundamentally from traditional software in that they are goal-oriented and will only perform some task if it contributes to the realisation of one of their goals. In contrast, most modern software systems are task-oriented and comprise programs performing tasks without any inkling as to what they are trying to achieve. Interestingly, the differences between objects, upon which the OOP paradigm is defined, and agents, are not correspondingly clear-cut. Some software practitioners consider agents as being nothing more than objects with some addition minor characteristics. Others regard both agents and objects as being intrinsically different, while sharing some characteristics. A detailed discussion of this viewpoint is beyond the scope of this article and may be found elsewhere[3]. However, agents do bring a new way of thinking to the traditional software development process and offer an alternative mechanism for modelling and realising solutions to problems in various domains.

References

1. M. Wooldridge and N. Jennings, "Intelligent Agents: Theory and Practice", Knowledge Engineering Review, vol. 10, no. 2, 1995, pp. 115-152.

2. M. Kinny, M. Georgeff and A. Rao, "A Methodology and Modelling Technique for Systems of BDI Agents", Proc. 7th European Workshop on Modelling Autonomous Agents in a Multi-Agent World (MAAMAW'96), Springer Verlag Lecture Notes in Artificial Intelligence, vol. 1038, 1996, pp. 56-71.

3. J. Odell, "Objects and Agents Compared", Journal of Object Technology, vol. 1, no. 1, 2002, pp. 41-53.

Spatial Agent

The Spatial Agent is assigned the task of monitoring the tourist's spatial context, that is, their position and orientation. From these it can make some deductions about their behaviour, for example, are they moving or stationary? All of this information is, of course, shared with other interested agents, particularly the Cache Agent and the Tourist Agent. To carry out its task successfully, the agent continuously monitors the GPS device and extracts both position and orientation parameters. It also verifies that these parameters are consistent with previous values and that the quality of the signal used during the calculation was adequate before it briefs other agents about the tourist's status.

Cache Agent

The Cache Agent is responsible for implementing the intelligent precaching strategy on the client. To do this, it is dependent on information from the both the Tourist Agent and Spatial Agent. From the Tourist Agent, it obtains a model of its immediate environment, which contains, amongst other things, a list of attractions within the tourist's immediate vicinity. From the Spatial Agent, it continuously receives position updates. Using both, it maintains a snapshot of the tourist's position with respect to the attractions within its environment. On deciding that the tourist is likely to visit a certain attraction, it requests a presentation on that attraction from the Tourist Agent so that it will be downloaded by the time the tourist reaches the attraction. After downloading the presentation, the Cache Agent confirms that the user's position is still relevant and, if so, proceeds to play the presentation. If the tourist should have not yet reached the attraction, it waits until such time as they within the vicinity of the attraction before playing it. If the downloading process takes longer than planned, the tourist must wait. Alternatively, if they have passed the attraction, then the presentation will not be played and the Cache Agent will dispose of it when it is satisfied it is no longer required and the tourist appears intent on visiting another attraction.

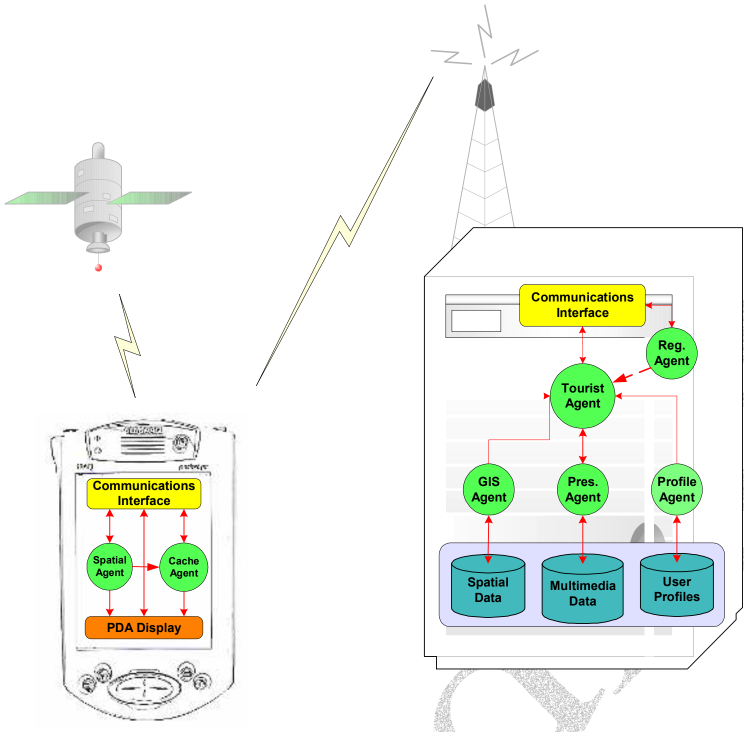

Figure 2: Architecture of Gulliver's Genie.

Tourist Agent
All tourists are assigned their own individual agents on the commencing a session with the Genie. In essence, this agent acts as the tourist's primary focal point for interfacing and accessing those services provided by the Genie. In conjunction with the other agents, it ensures that the environmental model is up-to-date and informs the Cache Agent of any changes. In conjunction with the Presentation Agent, it maintains a repository of pre-built presentations in anticipation of a request from the Cache Agent. This repository is continuously updated to reflect any changes in the tourist's environmental model.

Presentation Agent
The Presentation Agent is responsible for building presentations that are consistent with the tourist's individual profile. To do this, the Tourist Agent must provide it with details of both the tourist in question and their environmental model. It then proceeds to search the database for all relevant information on those attractions contained within the model. These are then filtered using the tourist's profile (obtained from the Profile Agent) after which presentations are constructed.

Profile Agent
The Profile Agent maintains the profiles of all users registered with the Genie and provides the relevant information to the Presentation Agent so that presentations can be customised accordingly. After a tourist has listened to a presentation, a record of their interaction is returned to the Profile Agent for further analysis and refinement of their profile. In this way, presentations are always consistent with the latest version of the tourist's profile.

GIS Agent
The GIS Agent advises on what a tourist's environmental model should be comprised of, given a particular geographical position. At present, it focuses on providing information about attractions in the tourist's vicinity.

Registration Agent

The Registration Agent administers the creation and termination of Tourist Agents as tourists join and leave the Genie community. It provides the initial contact point for tourists wishing to avail of the Genie's services.

Database

Underlying Gulliver's Genie is a multi-faceted database. Data stored within it may be classified as follows:

- Geospatial: Information concerning the position of all tourist attractions known to Gulliver's Genie is stored in this section. Naturally the storing and management of electronic maps is also of critical importance.
- Multimedia: All Genie presentations consist of some combination of sound, images, text or video. These media elements including some relevant meta-data are stored in the multimedia section of the database.
- Profile: All prospective users of Gulliver's Genie must first register. Details required include language, nationality etc. An initial snapshot of the user's interests is also stored.

Both the geospatial and the multimedia data comprise what we term a region's *information space*. Before the Genie can be used, an information space must first be defined for the area in question. Based on this, the GIS Agent can construct models of the tourist's immediate environment for closer observation.

Some Implementation Details

An IPAQ H3850 has been chosen to host the Genie. As it can be augmented with a dual-slot expansion sleeve, incorporating GPS and GPRS PCMCIA cards is not a problem. These components are illustrated in figure 3. The PDA connects to a server over GPRS using the standard Internet protocol HTTP. The runtime environment on the client and server is based on Java with the Jeode JVM being used on the IPAQ. All agents have been implemented using Agent Factory[15]. This provides a complete environment for designing and realising BDI agents. In particular, a lightweight implementation of its reasoning engine has been developed for PDAs. DB2 has been used to host the database.

Deploying multimedia on PDAs has proved somewhat problematic aside from the wireless communications issue. Displaying images is straightforward as most of the important formats are supported. The situation becomes more complicated when audio and video are involved. The Java Media Framework (JMF) provides extensive and sophisticated features when working with multimedia in a multi-platform environment but is limited on PDAs. While we have successfully deployed the JMF on the IPAQ, it only supports a fraction of the formats that it does on a normal workstation environment. Furthermore, while it can successfully render a number of formats, in practice, some are too slow on a PDA to be considered practical. Eventually, we settled on the reliable but somewhat low quality *au* format for rendering audio. Unfortunately, in the case of video, it proved impossible to render a video at an acceptable rate so we have not included video in any of the presentations at this time.

Figure 3: Genie Components showing an IPAQ, expansion sleeve, GPS and GPRS cards.

What the Tourist Experiences

When the tourist encounters an attraction, a presentation is automatically displayed. All presentations follow a similar template and consist of an image, image title and a number of *Links* (hyper-links) (figure 4.a). Each Link constitutes a concise information snippet about some aspect of the tourist attraction in question and consists of:

1. A name to indicate the subject of the Link, for example architecture. By convention, the first Link is usually used to identify the attraction in question thus allowing the tourist time to orientate themselves.
2. A sound recording on the subject of the Link;
3. An optional image to reflect the subject of the Link, should it merit its own image. If not, the default image of the relevant exhibit is automatically displayed.

All Links are activated using a stylus, the default interaction mechanism for interacting with a PDA, although some shortcuts have been programmed using the IPAQ's navigation pad.

While such an interface structure would prove adequate for most tourists, the Genie must also attempt to cater for those tourists who might have more experience in certain topics and thus have higher expectations. For example, a general description of the architecture of a building may be adequate for most people. However, for those with experience or additional expertise in the field, this might not be adequate and an opportunity could be lost for taking advantage of this fact and customising the presentation accordingly. To cater for such tourists, the Genie offers the user the possibility of accessing further information through what we term *Followups*. If a Followup is available for a Link, a plus or expansion icon is placed beside it otherwise the default negative symbol is used. By selecting this, the Link expands into a sub-list of highly specialised sub-Links that the tourist can access in the normal way. A simple example is shown in figure 4.b.

(a) (b)

Figure 4. Screen snapshots showing presentations.

How Agents deliver Intelligent Precaching

In light of the previous discussion, it is instructive to review how Gulliver's Genie uses intelligent precaching to make multimedia presentations available to users in an *as needed* fashion. As a tourist explores an area, the Spatial Agent is continuously monitoring their movements and informing the Tourist Agent of the latest position readings. Based on these, the GIS Agent advises on what the contents of the Cache Agent's environmental model should be, and the Tourist Agent informs the Cache Agent accordingly. Simultaneously, it arranges with the Presentation Agent to build and store presentations consistent with the attractions in the environmental model, in anticipation of a future request from the Cache Agent. Meanwhile on the PDA, the Cache Agent, acting on position updates from the Spatial Agent, attempts to anticipate where the tourist is likely to visit. To do this, it continuously monitors the position of the tourist relative to the attractions within their vicinity (obtained from the environmental model). When it deduces that the tourist is obviously converging on a certain attraction, it requests the corresponding presentation from the Tourist Agent and stores it in its cache on the PDA. When the tourist encounters the attraction, it automatically displays the presentation. Should the tourist have had a late change of heart and moved somewhere else, the presentation will remain in the cache until such time as another one is requested whereupon it is discarded. This entire process takes place in a transparent fashion to the user and is continually repeated as long as the tourist is availing of the Genie services.

It is instructive to reflect on why intelligent agents are particularly suited to implementing the precaching mechanism as just described. Their autonomous nature enables the entire process to take place in a transparent manner to the user. In continuously monitoring the tourist's behaviour, agents can react to events as they happen, reason about the implications of any course of action before proactively seeking to anticipate and address possible future needs and requests. From a modelling and implementation perspective, the social ability of agents enabled an intuitive and modular approach to system design. An example of the pertinent interactions is illustrated in figure 5.

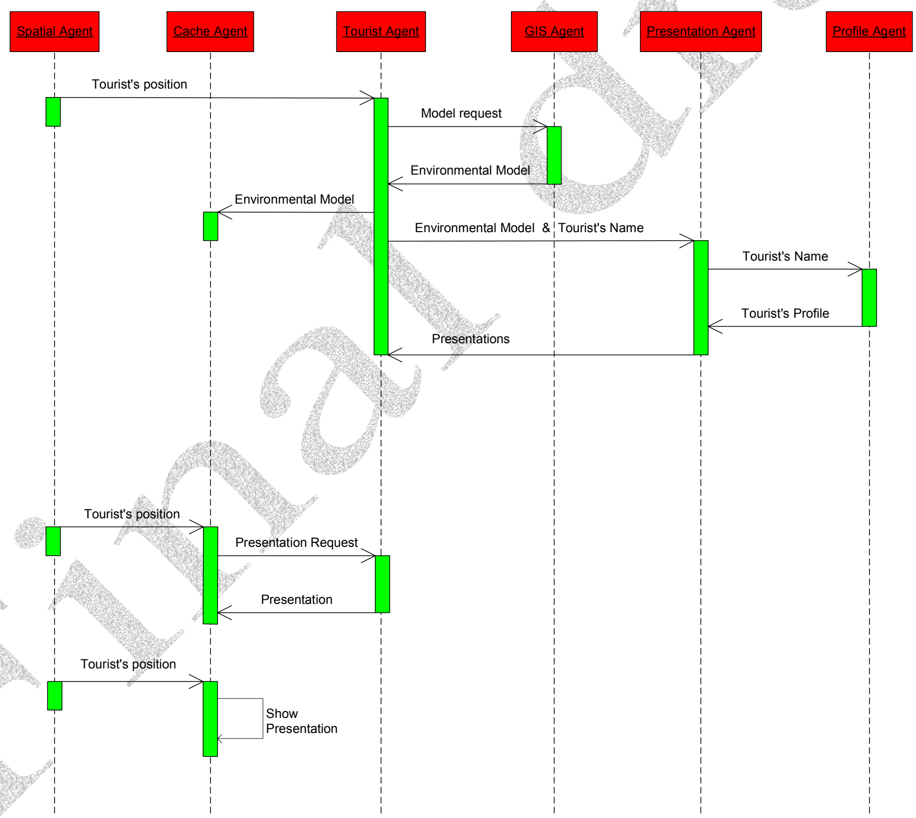

Figure 5. UML Sequence Diagram illustrating the essential interactions between agents as they collaborate to deliver a precaching solution.

Some Evaluation Feedback

To verify the approach taken with the Genie, it was necessary to conduct user evaluations and trials. Before doing this, we defined an information space by identifying the most interesting buildings and art works on the campus of University College Dublin (UCD). Fortunately, the campus is well spread out, comprises a number of historical buildings and is well endowed with art works of various kinds. Therefore we were able to include attractions that would appeal to a wide spectrum of users.

A sample group of 40 users took part in the evaluation. This group, 56% male, 44% female, tended towards a younger age group with 65% being less than 30 years of age. Most had some computing experience as expected but only 44% had some prior experience with PDAs. From an occupational perspective, the group was mixed comprising students, visiting students from other universities, people working on campus and some day visitors to the university. All evaluations took place during daylight hours and under varied though not wet weather conditions. Each evaluation took about 45 minutes after which volunteers were asked to complete a questionnaire, which amongst other things asked them to rate various aspects of the Genie on a Lickert scale of 1 to 7 (7 meaning being completely satisfied) as well as make some recommendations for improvements. Overall, 85% of participants were quite satisfied with the system, scoring it within the 5-6 range. 56% were very satisfied with the navigational aspects of the Genie, scoring it within the 6-7 range. The level of satisfaction regarding the presentations was excellent with 84% of participants scoring it within the 5-7 range. Finally over 70% were satisfied with the system functionality, scoring it with the 5-7 range.

Open feedback was used to tease out what users considered to be the Genie's strengths and weakness as well as to identify how the system could be improved. Users saw navigation support and the Genie's intuitive interface as being its principal strong points. A third characteristic identified by users was the potential of the Genie to deliver information at the right place and time. A critical limitation of the Genie, as perceived by users, was its responsiveness when using the stylus. More interestingly, it was noted that the timing of the presentations was occasionally out of phase as they were displayed after the user had passed the attraction in question. This was an implementation decision as we decided that even if the user had moved some small way beyond the attraction, the presentation would still be shown. Though only observed by 10% of users, it does suggest a need to review and refine the precaching process. In particular, the ability of the agents to cater for the inherent inaccuracy of GPS and the variable data-rates delivered by GPRS needs to be strengthened further, possibly through the incorporation of 3G and SBAS technologies as outlined in the next section. Having said that, it must be stated that, contrary to our initial expectations, GPS and GPRS performed reliably during the evaluations. Due to the nature of the campus, GPS coverage was not an issue and no discernible loss of signal was experienced. In the case of GPRS, signal loss was only experienced by 4% of users, but as almost instant reconnection was available, it was not perceived as a major problem. A third issue that arose was the perceived control of the Genie. A number of users had reservations about the autonomous nature of the Genie and expressed a desire to actively decide when the Genie should download data. It may well be that such users do not quite understand the implications of this. However, it seems reasonable that they should be given the choice, particularly as they may have to pay for such a service in real life and as such, a facility to let them specify this feature of the Genie's behaviour will be included in the next version.

Figure 6. In this case a tourist is consulting the Genie about a sculpture that he has encountered.

Observations & Future Developments

Overall, the results of the evaluations have proved encouraging, not withstanding the fact that a number of areas have been identified for improvements. Therefore, we intend to continue augmenting and refining the agents' behaviour thus improving the robustness and reliability of the Genie. Though satisfied that intelligent precaching offers a viable approach to multimedia dissemination; nevertheless, some improvement is necessary before it could form the basis of a commercial application. While reliability is not an issue, users would expect the Genie to have completed the precaching of the appropriate content at the appropriate time; something that is very difficult given the unpredictable nature of the data rates supported by wireless networks as well as the inherent position errors in GPS. Fortunately, two on-going developments will address these problems. Firstly, the somewhat protracted migration to 3G networks will increase the available data-rates thus reducing download time. And secondly, the development of Satellite-Based Augmentation Systems (SBASs), for example, the Wide Area Augmentation System (WAAS)[19] and the European Ground Navigation Overlay Service (EGNOS)[20]. Such technologies are currently being integrated into GPS receivers and will lead to position estimates in the five meters range; a significant improvement on the current 20 meters norm. To conclude this discussion on precaching, it is essential to consider when such an approach should be contemplated. Two situations immediately spring to mind. If the mobile user is equipped with a device of limited storage capacity, then intelligent precaching offers one approach for distributing multimedia content to such users. Secondly, should the content have a dynamic component, for example, new exhibitions in a museum etc, a precaching solution might well be considered. Finally, the nature of the application domain needs consideration. The tourism domain is ideal as attractions are usually liberally dispersed and sufficient time is available in order to deduce the tourist's destination, before dynamically assembling the presentation and proceeding to download it to the device. Alternatively, some domains may be particularly suitable to a hybrid solution. Static content could be stored on the device using a Smart Card and the dynamic content may be precached on the device as the occasion demands.

Our work with Gulliver's Genie will continue in a number of areas. Initially, we plan on incorporating more agents to provide additional services. The navigation component of the user interface requires additional enhancements to cater for users of differing abilities and it is our intention to expand the support of the Genie for further multimedia data types. For example, we plan on augmenting the 2D map with some 3D constructs. The incorporation of panoramic images is also envisaged, as is the use of VRML or Java3D for historical reconstructions. The use of scrolling text commentaries is also planned to cater for those people who might have hearing difficulties.

Given the experience and expectations people currently have for technology, the Genie's lack of support for video would be generally viewed as a significant deficiency. However, a number of promising solutions are already on the horizon. The most prominent of these, namely MPEG-4 [16], is a comprehensive open standard aimed at delivering audio and video content to heterogeneous devices across networks of varying bandwidth. If we regard the Genie as a potential 3G service, as distinct from an application in its own right, the use of a standard mechanism for defining and delivering presentations would prove highly desirably. SMIL 2.0 (Synchronised Multimedia Integration Language)[17] offers an intriguing possibility here, particularly when it is considered that SMIL forms an intrinsic component of the 3GPP transparent end-to-end packet-switching streaming service (PSS) specification [18]. Should either or both these possibilities come to fruition, the possibility would then exist for the identification of heuristics that could be exploited by the agents for determining those circumstances under which either a streaming or precaching solution would be preferable.

Conclusion

Intelligent precaching provides an alternative mechanism by which multimedia data may be made accessible to mobile computer users. In particular, it can help minimise and disguise two of the prominent limitations facing prospective suppliers of services to mobile users: poor data rates supported by the current generation of wireless data networks and the limited computational resources of popular mobile devices. Given that mobile computing is likely to be the prevailing computer usage paradigm in the coming years, the necessity to deliver a complex range of multimedia data in a timely manner will become increasingly urgent if mobile computing in its various guises is to fulfil its potential.

Acknowledgements

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

References

1. T. Pfeifer, T. Magedanz and S. Hübener, “Mobile Guide – Location-aware Applications from the Lab to the Market”, Proc. 5th International Workshop on Interactive Distributed Multimedia and Telecommunications Services (IDMS’98), Oslo, Norway, September 1998, pp. 15-28.

2. M. J. O’Grady and G. M. P. O’Hare, "Accessing Cultural Tourist Information via a Context Sensitive Tourist Guide," Information Technology and Tourism Journal, vol. 5, no. 1, 2002, pp. 35-47.

3. G. Abowd et al, “Cyberguide: A Mobile Context-aware Tour Guide”, Baltzer/ACM Wireless Networks, vol. 3, 1997, pp. 421-433.

4. N. Davies et al., “Using and Determining Location in a Context-Sensitive Tour Guide,” IEEE Computer, vol. 34, no. 8, August 2001, pp. 35-41.

5. V. Vlahakis et al, “Archeoguide: An Augmented Reality Guide for Archaeological Sites”, IEEE Computer Graphics and Applications, vol. 22, no. 5, September/October 2002, pp. 52-60.

6. F. Bellotti et al, “User Testing a Hypermedia Tour Guide”, IEEE Pervasive Computing, vol. 1, no. 2, April-June 2002, pp. 33-41.

7. S. Poslad et al., "CRUMPET: Creation of User-friendly Mobile Services Personalised For Tourism," *Proc*. 2nd International Conference on 3G Mobile Communication Technologies, London, UK, March 2001.

8. C. Hoymann and P. Stuckmann, “On the Feasibility of Video Streaming Applications over GPRS/EGPRS”, IEEE Global Telecommunications Conference, Taiwan, 2002.

9. R. Prasad and M. Ruggieri, “Future Trends in the Wireless Multimedia Environment”, International Conference on Communication Technology (ICCT’2003), Beijng, China, 2003, pp. 11-20.

10. D. Barbará and T. Imielinski, “Sleepers and Workaholics: Caching Strategies in Mobile Environments”, Proceedings of ACM SIGMOD International Conference on Management of Data, Minneapolis, Minnesota, USA, 1994, pp. 1-12.

11. Z. Jiang and L. Kleinrock, “Web Prefetching in a Mobile Environment”, IEEE Personal Communications, vol. 5, no. 5, October 1998, pp. 25-34.

12. J. Xu, X. Tang and D. Lee, “ Performance Analysis of Location-Dependent Cache Invalidation Schemes for Mobile Environments”, IEEE Transactions on Knowledge and Data Engineering, vol. 15, no. 2, March/April 2003, pp. 474-488.

13. S-Y. Wu and K-T. Wu, “Dynamic Data Management for Location Based Services in Mobile Environments”, Proc. 7th International Database Engineering and Applications Symposium (IDEAS’03), Hong Kong, July 2003.

14. G. O'Hare and M. O'Grady, "Gulliver's Genie: A Multi-Agent System for Ubiquitous and Intelligent Content Delivery," Computer Communications, vol. 26, no. 11, 2003, pp. 1175-1185. Elsevier Press, Amsterdam, NL.

15. R. W. Collier et al., "Beyond Prototyping in the Factory of Agents," Proc. 3rd International Central & Eastern European Conference on Multi-Agent Systems (CEEMAS’03), Prague, Czech Republic, June, 2003, Lecture Notes in Computer Science (LNCS), Springer Verlag Publishers.

16. F. Periera and T. Ebrahimi, “The MPEG-4 Book”, Prentice Hall, ISBN 0-13-061621-4, 2002.

17. W3C Recommendation: “Synchronised Multimedia Integration Language (SMIL 2.0)”, http://www.w3.org/TR/2001/REC-smil20-20010807/

18. 3GPP TS 26.246 V1.0.0, "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Transparent end-to-end packet switched streaming service (PSS); 3GPP SMIL Language Profile (Release 6), 2003.

19. Wide Area Augmentation System (WAAS) - http://gps.faa.gov/Programs/WAAS/waas.htm

20. European Ground Navigation Overlay Service (EGNOS) - http://www.esa.int/export/esaSA/navigation.html