The AVATON System: Location-Based Multimedia Services for Tourists

A. D. Sotiriou, P. Kalliaras, P. Papageorgiou, S. Zoi, John N. Karigiannis, N. Mitrou
National Technical University of Athens,
School of Electrical and Computer Engineering, Computer Network Laboratory,
9 Heroon Polytechneiou Street, Zographou 15773, Athens, GREECE
Tel: +30 2107721495, Fax: +30 2107722534
E-mail: dsot@telecom.ntua.gr

Abstract- In this paper, the AVATON system is presented, which allows the provisioning of ubiquitous, personalized, location-based multimedia services for tourists within the area of the Aegean Volcanic Arch. A uniform architecture is adopted in order to dynamically release the geographic and multimedia content to heterogeneous user devices, fixed and mobile, over the Internet, via Television channels and through wireless networks.

Keywords: Location-based, personalized content, multimedia services.

I. INTRODUCTION

The evolution of mobile technologies and their convergence with the Internet enable the development of interactive services targeting users with heterogeneous devices and network infrastructures [1]. Specifically, as far as cultural heritage and tourism are concerned, several systems offering location-based multimedia services through mobile computing and multimodal interaction have already appeared in the European research community [2],[3].

Although such services introduce new business opportunities for both the mobile market and the tourism sector, they are not still widely deployed, as several research issues haven’t been resolved yet and also available technologies and tools are not mature enough to meet end-user requirements. Furthermore, user heterogeneity both from different device and network technologies is another open issue as different versions of the multimedia content are often required.

This paper presents the AVATON system. AVATON aims at providing citizens with ubiquitous user-friendly services, offering personalized, location-aware [4], tourism-oriented multimedia information related to the area of the Aegean Volcanic Arch. Towards this end, a uniform architecture is adopted in order to dynamically release the geographic and multimedia content to the end users through enhanced application and network interfaces, targeting different device technologies (mobile phones, PDAs, PCs and TV sets). Advanced positioning techniques are applied for those mobile user terminals that support them.

The rest of the paper is organized as follows: Section 2 describes the AVATON services. Sections 3 presents the AVATON architecture, while Section 4 covers deployment and usage details. Finally, Section 5 outlines directions for further work.

II. SERVICES

The “AVATON” is an ambient information system that offers an interactive tour to the user (visitor) in the area of the Aegean Volcanic Arch [5]. The system can serve both as a remote and as on-site assistant for the visitor, by providing multimedia rich content, through various devices and channels:

- Over the Internet, via web browsers with the use of new technologies such as rich-clients and multi-tier architecture in order to dynamically provide the content.
- With portable devices (palmtops, PDAs) and 2.5G or 3G mobile phones, which are capable of processing and presenting real time information relevant to the user’s physical position or areas of interest.
- Via television channels

AVATON allows users to directly correlate geographic with informative space and conceivably pass from one space to the other, in the context of Worldboard [6]. With the use of portable devices equipped with positioning capabilities, the system provides:

- Dynamic search for geographical content, information related to users’ location or objects of interest that are in their proximity.
- Tours in areas of interest with the aid of interactive maps and 3-D representations of the embossed geography.
- Search for hypermedia information relative to various geographic objects of the map.
- User registration and management of personal notes during the tour that can be recalled and reused during later times.
- Interrelation of personal information with natural areas or objects for personal use or even as a collective memory relative to visited areas or objects.

The services described above are provided across all access platforms with varying degrees of functionality. More specifically:

- Workstation users using the Internet have full functionality of the system. These services are developed as web based applications, so they can be accessed via rich clients. They offer dynamic representations such as 3-D animations, interactive maps, streaming video and sound.
- Handheld devices, such as mobile phones and PDAs, support the same spectrum of operations as workstations in limited scale (e.g. a subset of the content), as determined by their processors and displaying capabilities.
• The access via television channel is set up so as to assimilate the interactivity of the other applications and also provide easy distribution of content through high speed channels. The streamed video is structured in such a way that it provides the illusion of an interactive application. This is achieved by splitting the television screen in four different regions, each one presenting a different information stream from different channels. One of these regions supplies content, simulating a user using the AVATON application, a virtual user of the system, while the other regions provide additional information in video and text format.

III. THE AVATON ARCHITECTURE

A. Overview

The Avaton System is based on a client-server architecture composed of three main server components, the Application Server, the Content Server and the Location Server. The Application Server combines information and content from the Content and Location Servers and replies to client requests through different network technologies. The content is retrieved from two kinds of databases, the Geographical and Multimedia Content DBs. The above architecture is shown in Fig. 1.

![Figure 1. The Avaton Architecture](image)

In more detail:

• **Multimedia Content Database**: This database contains the multimedia content such as images, video, audio and animation.

• **Geographical Content Database**: A repository of geographical content such as aerial photos, high-resolution maps and relevant metadata.

• **Content Server**: The Content Server supplies the Application Server with multimedia content. It retrieves needed data from the Multimedia Content Database according to user criteria and device capabilities and responds to the Application Server.

• **Location Server**: Serves requests for geographical content from the Application Server by querying the Geographical Content Database. The content retrieved is transformed into the appropriate format according to user device display capabilities and network bandwidth available.

• **Application Server**: The Application Server receives requests from different devices through GPRS, UMTS (3rd generation mobile phone), W-LAN, Internet (PDA, laptop, PC) and RF (television). The server identifies each device and transmits data in an appropriate format. More precisely, the Application Server incorporates a Web Server and a TV Server in order to communicate with PCs and televisions respectively.

B. Client

This section focuses on the mobile-phone and PDA applications. The scope of the AVATON system includes Java enabled phones with color displays and PDAs with WLAN or GPRS/UMTS connectivity. While all the available data for the application can be downloaded and streamed over the network, data caching is exploited for better performance and more modest network usage.

![Figure 2. The XML-based technologies in the client side](image)

When the users complete their registration in the system, they have in their disposal an interactive map that initially portrays the entire region as well as areas or individual points of interest. For acquiring user position, the system is using GPS. The client also supports multi-lingual implementation, as far as operational content is concerned, for example menus, messages and help. These files are maintained as XML documents. XML is extensively used in order to ease the load of parsing different data syntaxes. A single process, the XML Parser is used for decoding all kinds of data and an XSL Transformer for transcoding them in new formats. The different XML formats are XHTML, SVG, SMIL, SOAP and GML, as shown in Fig. 2.

1) **Geographical Info Presentation**

In order to render the geographical data, the Client receives raster images for the drawing of the background map, combined with metadata concerning areas of interest, links to additional textual or multimedia information. The raster data is aerial high resolution photographs of the region on two or three scales. Because of the high resolution of the original images, the client is receiving small portions, in the form of tiles from the Raster Data Processing Engine in the server side, which are used to regenerate the photorealistic Image Layer in a resolution that is suitable for the device used. The attributes of the geographical data, are generated in vector ShapeFile[7] Format, which is quite satisfactory for the server side but not for lightweight client devices. So, a
SHP TO SVG Converter at the server side is regenerating the metadata in SVG[11] format that can be viewed properly from a handheld device. As soon as the metadata is downloaded to the client device, a final filtering (XSLT Transformation) is done and the additional layer is opposed to the image layer in the SVG Viewer. On the SVG Data Layer, the user can interact with points of interest and receive additional information in the form of text or multimedia objects. The above are shown in Fig. 3.

![Client – side map rendering](image)

2) Multimedia Info Presentation
The presentation of multimedia information mainly depends on user position. The system is designed to provide audio and video clips, 3-D representations and also textual information concerning each place of interest. Not all devices, though, receive the same content, since they differ in display, processor or network speed. For that purpose, for each registered user, the system decides what kind of content is more suitable for them to receive and the Multimedia Content Server generates the appropriate script. Depending on the available memory of the client’s device, media objects stay resident in the cache memory, so that frequently requested content is accessed without delays that occur due to network latency.

![Client – side map rendering](image)

In Fig. 4 the components that are involved in the multimedia presentation are shown. The TourScript Data contains the script which describes the multimedia presentation. It is transcoded inside the SMIL Generator to a SMIL message that follows the XML syntax, so that it can be incorporated seamlessly to the messages that are exchanged in the AVATON system. At the client side, the SMIL message is received by the SMIL Processor which coordinates the process of fetching the Multimedia Objects from the Client Cache Memory to the suitable renderer, so that the Multimedia Presentation can be completed.

C. Location Server
The Location Server is the component that handles the geographical content of the AVATON system. It provides a storage system for all geographical data and allows querying of its contents through location criteria, such as global position and areas of interest. Content management is based on a PostgreSQL [8] relational data base. A JDBCInterface uses the JDBC APIs in order to provide support for data operations. A GISExtension is also present, based on PostGIS [9], in order to enable the PostgreSQL server to allow spatial queries. This feature is utilized through a GISJDBCInterface, a PostGIS layer on top of PostgreSQL.

1) Cartographic data
As long as the photorealistic information is concerned, the user can choose from several distinct zoom levels. The mobiles phones and PDAs in the market that support GPRS or WLAN have displays of different resolutions that, in most cases, are multiples of 16 pixels. Hence, the Location Server can generate tiles with a multiple of 16p x 16p, which can be presented in the user’s mobile device. The server always holds multiple resolutions for every level of cartographic (photographic) information. The levels of cartographic information define the degree of focus.

D. Multimedia Content Server
The Multimedia Content Server component comprises the major unit that controls the mixing and presentation of different multimedia objects. Its purpose is to upload all the objects necessary and present them in a well-defined controlled order that in general depends on the user position, interactions, and tracking information available. The multi-lingual audio-visual information scheduled for presentation is coordinated so that several objects may be presented simultaneously. The Multimedia Content Server component is also responsible for choosing “relevant” objects for the user to select among in the case the user requires more information on a topic. The Multimedia Content Server interfaces with the Multimedia Content Database, a relational database storing the multimedia content. The database is organized thematically and allows the creation of hierarchical structures. It also contains a complete list of multimedia material, covering all content of the physical site, such as 3D reconstructed plants, audio narration, virtual 3D models, avatar animations and 2D images.

1) Media Objects
As mentioned already, the Multimedia Content Server is responsible for mixing the basic units of multimedia information. These elements are hierarchically ordered. At the finest level of granularity, there are atomic objects called MediaObjects with specializations such as
AudioMediaObject, ImageMediaObject, 3DMediaObject and CompositeMediaObject. These objects contain the actual data to be rendered along with additional profile metadata characterizing them. At a higher level of complexity, a TourScript represents an ordered sequence of MediaObjects, all of which are to be presented if the script is chosen.

According to user requirements, the user will be able to navigate through the site in a geographically-based tree. This is made possible through the use of Points of Interest (Pol) and Areas of Interest (AoI). A Pol can only contain TourScripts and can be viewed as the end node of the site tree. In contrast, a AoI may contain either another Pol, an AoI or TourScripts. This allows the system to map the actual site into a hierarchy model containing Pol at the top and MediaObject components at the leaf level.

The Multimedia Content Server is also responsible for managing this site-tree for the entire site. Moreover it is responsible for traversing the tree. The use of the site tree is quite interesting: when a media object, for instance audio object is presented to the user, it belongs to a node in the site hierarchy. Fig. 5 shows the structure of the site in a tree view as described previously. The Multimedia Content Server is responsible for coordinating the rendering components in order to provide a synchronized presentation to the user, according to user preferences, position and commands.

![Figure 5. Hierarchy formulation of media objects at the Multimedia Content Server](image)

2) Inter-communication

The system makes extensive and ubiquitous use of XML technologies in all of its components (client side, location server, content server) as well as in intra-communication between them. In order to have a light and fast communication language that all components can share, the SOAP[13] technology has been adopted. The SVG format used for the geographical information, the SMIL format used for multimedia presentations and the XHTML format used for rich-text presentations are also XML-compliant. In addition, another XML-based language, GML has been used for describing GPS position information. The XML/SOAP parsing code is generated by the Sun JWSDP toolkit along with the use of a SAX-based parser and Java SAX API calls. In more detail, messages are XML documents which contain AVATON message data, as well as, depending on the type of the message, XHTML code and/or SVG data. Message structure is thus well defined structure and is expressed by means of XML Schema. There is a basic schema which captures the generic message structure, plus additional sub-schemas which further refine message types. AVATON supports the following messages types:

- **Session preamble**
  The Client application initiates a session with the AVATON server. Purpose of this message is to establish the session’s characteristics, namely:
  1. Client device type
  2. Device position
  3. Caching capability
  4. User data
     - Language selection
     - Email
     - Nickname

- **Area content request**
  The client device requests a summary of available information regarding points of interest that are contained within the area that is determined by client’s type, the client’s position and a zoom indicator. Basically the server can calculate a tile having as center the client’s current location and dimensions the device’s display dimensions as a multiple of 16px16p. The actual geographical area that can fit within that window is determined using the zoom indicator.

- **Media Request**
  The client device requests a specific multimedia resource, an HTML page or a video file that can be, under certain circumstances, streamed and/or adapted to meet requirements of clients with limited resources or other special requirements. This request is supported for the benefit of clients who do not display the full content related to an area or point of interest. Having this extra request type they, can choose for example to present part of the content to the user or when applicable, choose whether to present a video or not. The server receives the request and appropriately redirects to the appropriate multimedia resource URL.

- **Point of Interest Request**
  The client device requests a summary of available information related to a specific point of interest. The server can either create a summary of available information adding relevant links to full information or alternatively provide the main content for the particular content in XHTML format. The choice depends on the client type.

- **Conclusion request**
  The session can be terminated either by the visitor or by the server. The visitors can indicate that sightseeing is over and in case they have created a list of interesting pictures, can request that the pictures be sent at the given e-address. The server will generate an e-mail message which will contain any requested content as attachments. Alternatively, the session timeout mechanism can automatically terminate a session when there is inactivity for a certain amount of time.

IV. DEPLOYMENT AND USAGE

Based on the proposed architecture, the AVATON services are being deployed to physical sites within the Aegean Volcano Arc (such as Santorini and Lesvos islands) and evaluated by real end users, under different scenarios. The main air interfaces that will be used by the
system (along with standard wired access through common LANs or the internet) are:

- **WLAN**
  A standard 802.11b wireless access network provides connectivity for users equipped with portable devices (such as PDA with built in WLAN cards or laptops).

- **GPRS/UMTS**
  For mobile users and smart phones, access will be provided through the GSM network, using GPRS [10]. This restricts the system from providing video or 3-D animations to such users, and the services offered are focused on text, images (including map information) and short audio. As GPRS is already packet-oriented, our implementation can be easily transferred to UMTS, if available.

Figure 6. Implementation plan for the Lesvos island site

In Fig. 6 the actual implementation plan is given for the location of the Sigri Natural History Museum, in Lesvos island [12]. The area consists of an open geological site, the Petrified Forest, where the ash from a volcanic eruption some 15 to 20 million years ago covered the stand of sequoia trees, causing their petrification. Wireless access is provided by the use of 3 Netgear 54Mbps access point equipped with additional Netgear antennas in order to overcome the physical limitations of the area (hills, trunks and hollows).

Figure 7. SVG map and interface

Visitors to the site are equipped with PDAs or smart phones (provided at the entrance kiosk) and stroll around the area. A typical scenario consists of the following: The users enter the archaeological site and activate their devices. They then perform a login and provide personal details to the server, such as username, language selection and device settings. The client then requests information from the server and loads the map of the area in SVG format, as seen in Fig. 7. The circles on the map present distinct points of interest (yellow indicating trees and blue indicating leaves). The application monitors the users’ location and updates the SVG map in real time, informing the users of their position. The SVG map is interactive and when the users enter the vicinity of a point of interest, the application automatically fetches and displays (via their browser or media player) the corresponding multimedia content (in the form of HTML pages, audio or video) at the requested language. The users are also able to navigate manually through the available content and receive additional information on topics of their interest. During the tour, the users’ path is being tracked and displayed (the red line on Fig. 7), in order to guide them through the site. They are also able to keep notes or mark favourite content (such as images), which can be later sent to them when they complete the tour.

V. FURTHER WORK

The system is currently being deployed and tested in two archaeological sites. Users are expected to provide useful feedback on system’s capabilities and assist in further enhancements of its functionality. Also, as 3G infrastructure is being expanded, incorporation of the UMTS network in the system’s access mechanisms will provide further capabilities for smart phone devices and also use of the system in areas where wireless access cannot be provided.

Towards commercial exploitation, billing and accounting functionalities will be incorporated into the proposed architecture. Finally, possible extensions of the system are considered in order to include other cultural or archaeological areas.

REFERENCES


