

# Personalized Access to Distributed Data Sources

Alessandro Andreadis, Giuliano Benelli, Pasquale Fedele, Giovanni Giambene

Dipartimento di Ingegneria dell'Informazione

Università degli Studi di Siena

Via Roma, 56, 53100 Siena, Italy

Tel.: +390577234608

E-mail: benelli@unisi.it, andreadis@unisi.it, f.pasquale@tin.it, giambene@marconi.ltt.dii.unisi.it

## 1. Abstract

The widespread diffusion of mobile communication systems is paving the way for new information services to be provided to users. This paper deals with the *Personalised Access to Local Information and services for Tourists* (PALIO - IST20656) Project within the fifth research framework of the European Commission. The focus is on tourists that need to access information on the city or on the region they are visiting by means of a cellular phone or at specific locations (kiosks). The PALIO project is committed to produce a prototype showing the potentialities of combining user position information and user preferences in order to provide context-aware information. This paper presents early results of the PALIO Project for what concerns adaptation to the access technologies, terminal characteristics and user preferences.

## Keywords

Content Adaptation, Multimodal Communication, Location-Aware Information Services, Information Retrieval.

## 2. Introduction

The evolution of mobile communications systems is paving the way toward the development of new information services. This work deals with the *Personalised Access to Local Information and services for Tourists* (PALIO-IST 20656) project within the fifth research framework financed by the European Commission [1].

At present, tourist information is static and mainly available through paper guides, Internet sites and information points located in specific positions. PALIO aims at providing tourist services to mobile and fixed users for facilitating their city sightseeing. In the PALIO Project, the user must be able to access services not only by using a common personal computer at home, but also by means of its cellular phone or a *Personal Digital Assistant* (PDA). Accordingly, the provision of personalized services anywhere and anytime requires that the information system be adaptable and scalable, depending on heterogeneous client devices (e.g., different displays, computation capabilities, software releases), user preferences and access

networks (e.g., available bit-rate). Content adaptation to the access device implies to provide the contents with suitable markup languages; whereas, the adaptation to the access technology will require the adoption of scalability of the presented content still allowed by means of markup languages, as explained later in this paper. Within the PALIO system, the *Augmented Virtual City* centre (AVC) adapts the presentation of information to the different access and client technologies (see Figure 1). The PALIO Project prototype will test the following services:

- *Real-time information* on public transportation and traffic situation;
- *Location-specific information* on tourist sites and sites of cultural interest (such as museums, hotels, exhibitions, concerts including navigation);
- *Personalisation of the information* to be provided to the user depending on both the context of use, history of the interaction with the system and user interests;

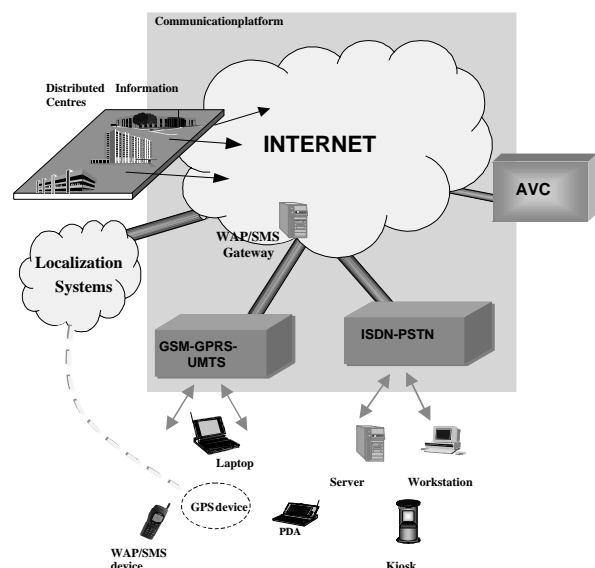


Figure 1: PALIO high-level architecture.

PALIO test sites will be carried out in the island of Crete, in Florence and in Siena. The PALIO partners are: ASSIOMA Group (I), Italian National Research Council (I), Florence

Municipality(I),FoundationResearch&Technology(GR), GMD(D),TelecomItaliaMobile(I),UniversityofSiena (I),SienaMunicipality(I),MA(UK)andFORTH.net (GR).

### 3. System architecture

The high-level system architecture is shown in Figure 1 and encompasses the main building blocks described below.

1. The *communication platform* comprises the Internet, network interfaces and gateways necessary to integrate the distributed components of the PALIO system.
2. The *Augmented Virtual City Centre* (AVC) contains adaptation and control functions.
3. The *Distributed Information Centres* in the territory.
4. The *Localization Systems* comprise the GPS solution and the mobile network-based solution to identify the position of the tourist in outdoor environments (indoor solutions are under investigation).

All these building blocks are detailed in the following subsections.

#### 3.1 The communication platform

The communication platform is based on the HTTP protocol for allowing the interconnection of the different and geographically dispersed elements of the PALIO system [2]. In particular, the elements of the communication platform are the WAP/SMS server [3], the gateway towards the mobile (GSM/GPRS/UMTS) network (including also the location server of the mobile operator) and the gateway towards the fixed (ISDN/PSTN) network.

#### 3.2 The Augmented Virtual City centre

The AVC centre consists of the following elements (see Figure 2):

- a) The *User Communication Layer* (UCL), an 'abstract' entity allowing the communication of the AVC with the other servers (i.e., Web & WAP server, WAP gateway, SMS server, and localization server) in a common format.
- b) The *Service Control Centre* (SCC), the heart of the PALIO system. It provides the runtime platform for the system information services.
- c) The *Generic Information Server* (GIS) for allowing the access to the distributed information space through the Internet. In particular, the GIS permits the retrieval

of information in distributed databases and then provides the obtained contents in a common format to the SCC.

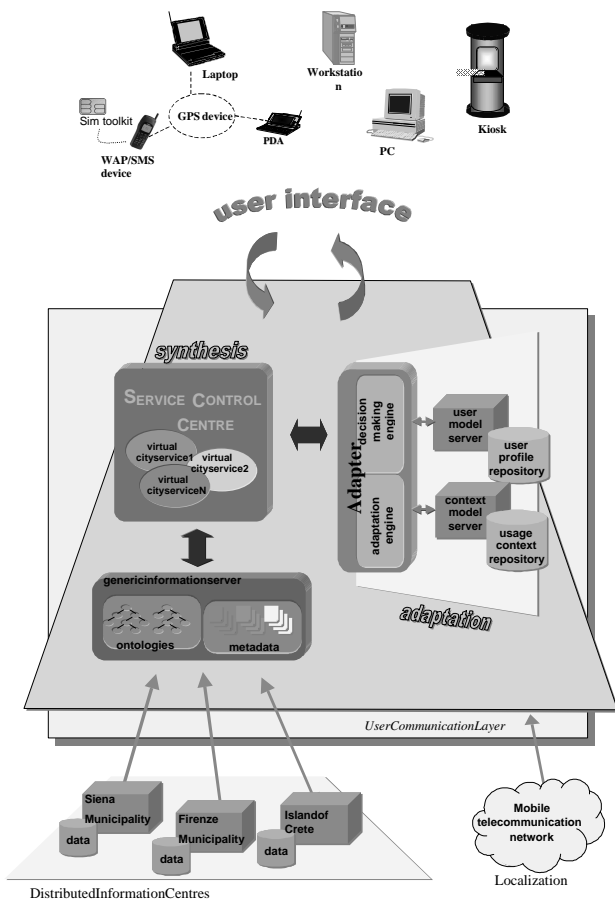
- d) The *adaptation infrastructure*, responsible for content and interface adaptation in the PALIO system. It represents one of the most innovative technological developments of the PALIO Project. This infrastructure is composed of the elements described below.
  - d1) The *Adapter*, the basic adaptation component of the system. It integrates information concerning the user, the context of use, the access environment and the interaction history. The Adapter provides adaptation guidelines to be adopted in the SCC to produce an adapted output content and representation.
  - d2) The *User Model Server* (UMS) that integrates and manages information concerning user characteristics (e.g., interests, interaction style, disabilities) in the form of both individual user profiles constructed during interactive sessions and user stereotypes for groups of users that share a number of characteristics.
  - d3) The *Context Model Server* (CMS) that assembles context profiles by using information retrieved from the *Usage Context Repository* (UCR) and informs the adapter about the current context.

#### 3.3 Distributed information centres

The PALIO system demonstrator will mainly employ the databases made available by Florence and Siena municipalities as well as FORTH.net operator in Crete. As for the Siena municipality database, whose information is partly accessible through the Siena municipality Webpage [4], a lot of tourist information is available encompassing different solutions for accommodations (hotels, residences, villas, etc.), restaurants, special events and news (some of them continuously updated).

#### 3.4 Localization systems

The study for the selection of the best technological option for locating tourists is in progress in order to match the specific characteristics of the information services envisaged by the PALIO Project. The present technological options for outdoor localization (GSM networks) are based on *Cell Identification + Timing Advance* (CellID+TA) and the GPS technique [5]. The GSM Italian operator will make available the use of the *Mobile Positioning System* (MPS) that implements the CellID+TA technique. In the indoor case (e.g., within a museum), Wireless LAN and Infra-Red-based techniques are under investigation.



**Figure 2:** The AVC architecture and interaction with the other PALIO building blocks.

#### 4. The flow control of the AVC

We can consider the SCC as a robot task engine that manages the flow of operations within the AVC, as detailed below with the steps represented in Figure 3.

1. The request from the client (i.e., the tourists) reaches the SCC by means of the HTTP protocol through the UCL. The UCL transforms the request in an XML file, before passing it to the SCC. The UCL is also responsible for requesting the position information, according to the selected localization method.
2. The SCC acquires from the client request all the necessary information (i.e., user identification, user agent type, location and type of request).
3. The Adapter receives the information regarding the request, the user, the client device, the client agent and the location and forwards it to both UMS and CMS. The adapter also forwards some part of the request to the *decision making engine*, which holds adaptation rules and a decision making model.

4. The decision making engine requests and gets from the UMS and the CMS the appropriate user and context based information.
5. UMS and CMS get from the relevant repositories the user profile and the context situation (UCR).
6. UMS and CMS update the user profiles and the context situations, respectively, by using the information coming from step 3.
7. The decision taken by the decision making engine is mapped to a proper adaptation action.
8. The adaptation action is passed to the SCC.
9. The request is adapted by the SCC on the basis of the instructions coming from step 8.
10. The request is analyzed before forwarding it to the GIS.
11. The request is forwarded to GIS.
12. The request is analyzed in order to formulate the necessary query to interrogate the database that contains the requested information.
13. Queries are forwarded to data sources.
14. The response is generated.
15. The responses are forwarded to SCC.
16. All the information (provided by the GIS in XML format) is collected within the SCC.
17. The response is assembled by the SCC with the help of instructions passed from step 8 and with some service domain rules.
18. Final construction of the response.
19. The XML response is generated and sent to the UCL.

The PALIO system will be based on leading edge technologies for what concerns content adaptation, content representation and application development. Java [6] and XML technologies will be the basis for the realization of an integrated environment for information retrieval and representation in the PALIO project. Java, as programming language for serverside applications, supports *servlets*, an efficient approach for creating Web applications, and *Java Server Pages (JSP)* [7], [8], an easy solution for generating documents based on markup languages (e.g., XML).

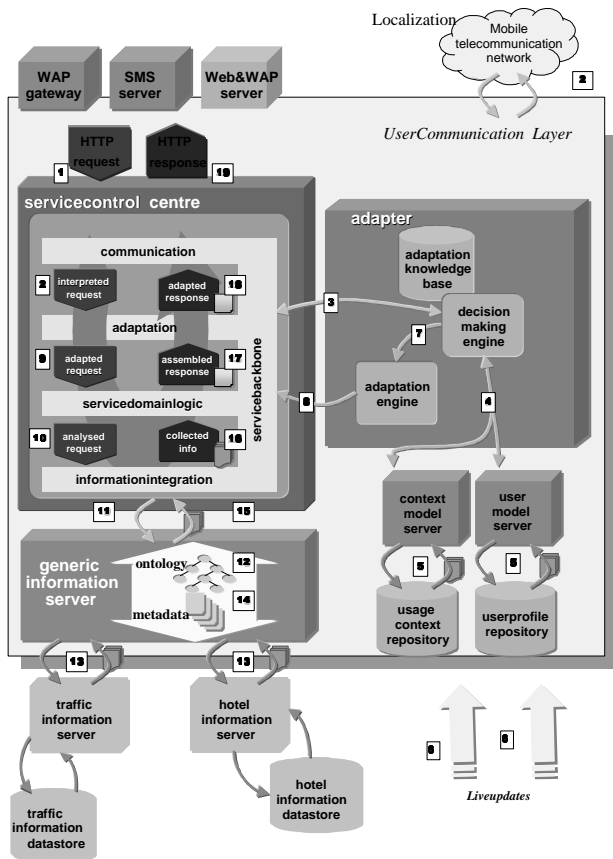


Figure 3: Flow control within the AVC.

XML is the most interesting solution for document and data interchange [9]. XML files are reproduced on the basis of a schema that derives from a specific ontology where each object has a list of attributes. Thus, XML helps to describe the contents of a document.

Several options are available to define a schema language for XML. The most significant are two: *Document Type Definition* (DTD part of XML 1.0) and XML Schema. An XML Schema is used for the description and the validation of both the structure and the content of an XML document. XML Schema also allows the definition of stronger requirements via a grammar: sequence and choice of attributes applied to an instance can be defined. DTD describes the valid XML tags, the order of the tags, whether the tags are optional, the type of data contained within the tags, the tag attributes, the attribute values, processing instructions, entities, and so on.

The use of such technologies enables a platform independence and built-in Internet support.

## 5. Present technological implementations

The PALIO Project is an early design and implementation phase. We describe below the present achievement within the PALIO consortium that are related to the following aspects: (i) identification of the client device accessing the

PALIO system; (ii) retrieval of information in distributed data sources; (iii) generation of multi-language output files depending on the adopted client device.

### 5.1 Identification of the client device

When a user sends a request to the PALIO system, the AVC is able to identify the client device characteristics by extracting information contained in the header of the received HTTP packet. This task can be performed by Java servlets. The packet header fields used for this purpose are:

- HTTP\_ACCEPT: containing the list of the *Multipurpose Internet Mail Extensions* (MIME) types accepted by the client;
- HTTP\_USER\_AGENT: that lists the most significant products information about the client. It is a string like *product/version*.

By means of this information extracted from the user request, the PALIO system (i.e., the AVC) is able to identify the capabilities of a client device, such as the accepted MIME types, display size, browser type and version, etc. This information allows defining terminal classification rules used by the adapter. Such rules can be either very specific (by using name and product version) or generic (based on accepted MIME types; e.g., WML or HTML browser support).

### 5.2 Retrieval of information and generation of the output to SCC

The mapping of the information contained in distributed data sources in the Internet is a task of the GIS within the PALIO Project. This operation is based on the *Mappamond* database contained in the GIS (see Figure 3). By using the *Mappamond* database it is possible to know which data source contains the information requested by the user and the necessary parameters for the access (URL, Driver, User -ID, Password, etc). An XML file in the *Mappamond* database describes the structure of every data source. Data are extracted by means of SQL queries and JDBC connections. Then, data are encapsulated in XML tags and the XML response is sent to the SCC.

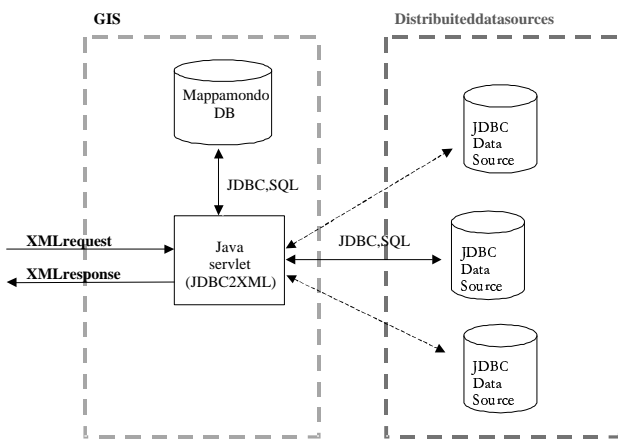


Figure 3: PALIO system query scheme.

### 5.3 Generation of multiple output markup languages for client devices

Once the SCCh has generated the XML response adapted to the user profile, position and context of use, this XML document needs to be adapted to the user device. Hence, there are two approaches to generate multiple markup languages from the same XML code. The first solution is based on *Document Object Model (DOM)* API to build an object tree, representing the elements hierarchy of the XML code. Then, JSP uses this object tree to generate directly the response to be sent to the client. DOM API works by loading the entire hierarchical structure in memory, allowing a faster manipulation of the information, but requiring expensive RAM and CPU resources. In the second solution, it is up to the *XSL Transformation (XSLT)* processor to perform the transformation of an XML document on the basis of a suitable XSL stylesheet, containing the translation rules for the specific client. This technique is described in Figure 4.

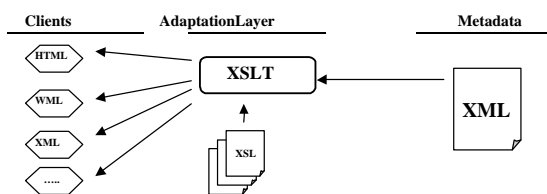


Figure 4: XSLT processor.

## 6. Conclusions

This paper has presented the PALIO Project approach toward the realization of context-aware mobile information services. The PALIO architecture has been detailed by focusing on the adaptation techniques, the retrieval of information on distributed data sources and the user

localization solutions. In this early design and implementation phase of the PALIO Project we have presented initial project results for what concerns the identification of the client device on the basis of the request received by the PALIO system, the retrieval of information on distributed data sources and the generation of multiple markup languages from a given document generated in the XML format.

**Acknowledgements:** The authors wish to thank the PALIO Partners for contributing to the definition of the PALIO system characteristics shown in this paper.

## REFERENCES

1. The PALIO Consortium Web page with URL: <http://palio.dii.unisi.it>.
2. Network Working Group, "Hypertext Transfer Protocol - HTTP/1.1", (WWW page) URL: <http://www.ietf.org/rfc/rfc2616.txt>, June 1999.
3. C. Arehart et al., *Professional WAP*, Wrox Press Ltd., 2000.
4. Siena Municipality Website: [www.comune.siena.it](http://www.comune.siena.it).
5. 3GPPTS 43.059 "Functional Stage 2 Description of Location Services in GERAN", 2001.
6. P. Naughton, *The Java Handbook*, McGraw-Hill, Inc, 1996.
7. H. Maruyama, K. Tamura, N. Uramoto, *XML and Java: developing Web applications*, Addison-Wesley, 1999.
8. M. Hall, *Core Servlets and Java Server Pages*, Prentice Hall, 2000.
9. K. Williams et al., *Professional XML*, Wrox Press Ltd., 2000.