A Semantics-aware Platform for Interactive TV Services

Antonis Papadimitriou, Christos Anagnostopoulos,
Vassileios Tsetsos, Sarantis Paskalis, Stathes Hadjiefthymiades
Pervasive Computing Research Group, Department of Informatics & Telecommunications,
Athens, Greece,
{a.papadimitrioy, bleu, b.tsetsos, paskalis, shadj}@di.uoa.gr

Abstract: Interactive digital TV is becoming a reality throughout the globe. The most important part of the picture is new services for the user, in terms of audio-video quality, but mostly in terms of entirely new content and interacting experience. To this end, we explore the potential of introduction of semantics in the distribution, processing and usage of the media content. We propose a smart iTV receiver framework capable of collecting, extending and processing (reasoning) semantic metadata related to the broadcast multimedia content. System architecture is presented along with an example service to illustrate the combination of semantic metadata, user preferences and external data sources.

Keywords: multimedia, metadata, semantic processing, interactive TV, MPEG-7, MHP, OSGi
Categories: H.5.1 Multimedia Information Systems

1 Introduction

Over the past years, there has been a coordinated effort from multiple organizations towards the establishment of standard transmission technologies and application execution environments for digital interactive TV (iTV). This has led to the development of several standards such as Digital Video Broadcasting (DVB) [1] transmission specifications (for satellite, cable and terrestrial TV) and the Multimedia Home Platform (MHP) [2] as middleware for interoperable interactive applications. There is already an extensive network of digital TV infrastructure based on these specifications, numbering a multitude of homes subscribed to iTV services from different broadcasters. Hence, an open market is evolving with great prospects of benefit both in terms of provider profit and user satisfaction.

The next step to the future of digital TV is the enhancement of interactive applications provided to the iTV subscribers. The platform described in this paper is named POLYSEMA, which is a prototype system which focuses on the development of a “smart iTV receiver” (smart residential gateway) capable of collecting, extending and processing semantic metadata related to the broadcast multimedia content, in order to offer pioneer services to users. In this context, content providers have to supply metadata descriptions along with multimedia content, in a certain metadata standard format (MPEG-7). Broadcasters multiplex metadata information into their transport streams so that smart receivers can take advantage of it. Receivers download metadata from the transport stream and/or the web sites of content providers, and process the information, taking into account user-defined rules and preferences.

POLYSEMA is based on several standards widely adopted by international research and industry communities. Particularly, the DVB-T standard [3] is chosen as the multimedia transmission specification, the MHP standard is chosen as an application execution environment, the OSGi platform [4] as a service-oriented integration platform, the MPEG-7 standard [5] as a metadata description standard and Semantic Web [6] as the base for semantics representation and logic-based inference.

The rest of the paper is organized as follows: Section 2 describes the overall architecture of the POLYSEMA system. In Section 3, we present how the proposed system can be used to support interactive services based on content provider information and user preferences. The following sections consider prior work to that research area and conclude the article by giving some insight for future work.

2 The POLYSEMA Architecture

The core of the POLYSEMA system is the receiver. Yet, in order to exhibit application and video synchronization at the subscriber's receiver, there is a need to add synchronization information at the broadcaster side. In this section, we describe both the novel architecture of the receiver and the additional operations running at the server. It is worth mentioning that POLYSEMA focuses on compatibility with the standards introduced in Section 1. This enables seamless interaction of the system with existing products through the addition of a single component.

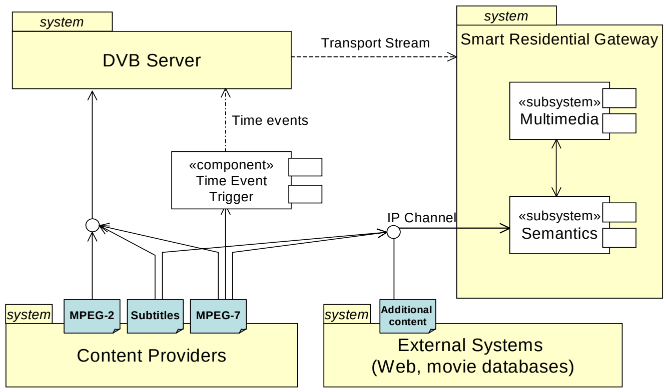


Figure 1: Overall POLYSEMA architecture

Moreover, the media player displaying the multimedia content at the receiver is a common MHP-DVB player. The system interacts with it by using its "default" interfaces (e.g. return channel) and by loading specially tailored MHP applications. Fig.1 depicts the generic components of the POLYSEMA system and their relationship with the broadcasting server, content provider and other external systems.

2.1 Extending the Broadcasting Server

MPEG-7 files can describe multimedia content on a scene-by-scene basis, by providing distinct descriptions for different sections of the video separated by clearly defined time markers. This model allows for a wide range of interactive applications, which adjust their behavior as the content presentation advances. A remarkable problem inhibiting such applications is that, in broadcast environments, the receiver cannot maintain a value depicting the concept of "default media time". That happens because iTV subscribers may tune to a specific program at any time during the broadcast. In such cases, applications are unaware of the “absolute media time” of a program event. The only way to synchronize video with an application is to transmit *stream events* within the stream, as described in the DSM-CC specification [7].

To overcome this problem, POLYSEMA utilizes a software module to convert MPEG-7 timestamps to stream events sent by the broadcasting server. The module parses the MPEG-7 document, which describes the content to be broadcast, and performs a mapping of MPEG-7 time elements to stream events, at the granularity of video segments. A time event signals the transition to a new scene of the video, which is associated to an MPEG-7 segment. A stream event consists of an *identifier*, *name* and *data*. Data can contain the MPEG-7 segment identifier. Any time the receiver tunes to a channel, the application can determine which part of the video is being displayed by listening for stream events and consulting the MPEG-7 document.

Summing up, the transport stream produced at the server side and sent to the receiver contains: AV content, MHP applications, and the broadcast file system mounted on the transport stream (the broadcast stream may include MPEG-7 files), stream event objects that convey synchronization information, and files in the DSM-CC object carousel, containing a list of URLs, so that the receiver can locate and obtain the corresponding metadata files.

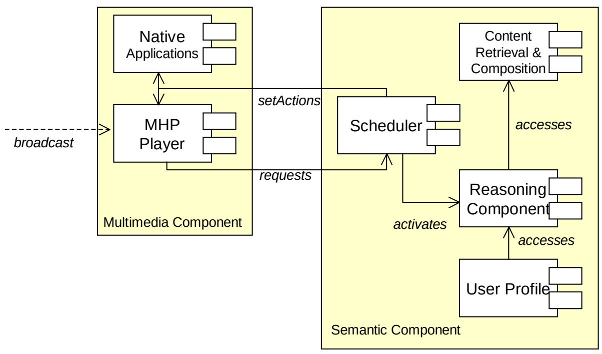


Figure 2: POLYSEMA Software Components

2.2 The POLYSEMA Smart Receiver Architecture

Each interactive service consists of a presentation part and a logic part. The former part refers to actions that the media player ought to perform when it comes to display the results of the latter part. Such actions may be implemented either in a platform-independent way over MHP and Java Virtual Machine, or using native applications. The software module responsible for the media presentation is the Multimedia Component, illustrated in Figure 2. The Semantic Component of POLYSEMA determines the actions which comprise each service. The Semantic Component is also in charge of retrieving and integrating any external resources in the POLYSEMA system. All components are built in an OSGi-based service-oriented fashion, in order to provide maximum flexibility in composing interactive TV services.

2.2.1 The Multimedia Component

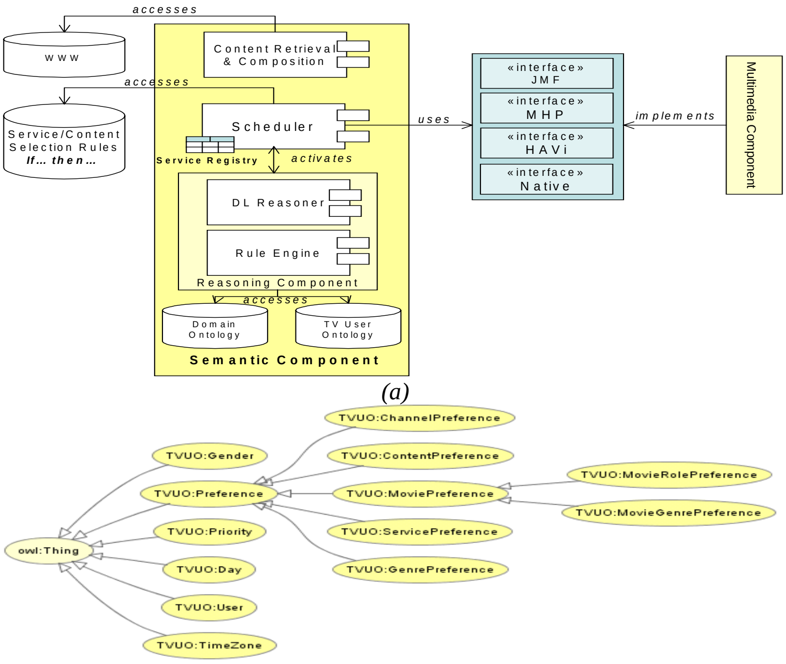
The Multimedia component is the basic constituent of the proposed system that produces multimedia content, interacts with the users and adapts to their preferences accordingly. Specifically, this component refers to the MHP application environment, in which, the MHP application instances run (e.g., on a Set-Top-Box). A user interface is supported through which the user is able to define and place several rules or actions related to their preferences (e.g., on demand multimedia content presentation and retrieval). Moreover, the Multimedia component communicates with the Scheduler of the Semantic component (Figure 2) when further contextual information is required to be fetched and processed. The Semantic component is responsible for a set of actions denoting the specific type of activity that the MHP player has to perform, e.g., change content presentation style and format, tune sound volume, display retrieved info or record part of a program. The Multimedia component can be envisaged as the ingredient that realizes certain actions derived from the reasoning tasks in the Semantic component, as discussed below.

2.2.2 The Semantic Component

The Semantic component refers to the semantic processing of relevant metadata. A more detailed view of its internal architecture is presented in Figure 3. A basic assumption for our system is that every AV content item is described by an MPEG-7 document. In order to reason over this document we transform it to a corresponding ontology, which was based on that proposed in [13]. In fact, we have developed a stripped down version of the ontology in order to eliminate any unused elements. Moreover, the user “defines” their service preferences by combining templates of possible actions and declaring rules about when such services should launch and how they should be presented. The user input is based on the TV User Ontology (Figure 3b). The Reasoning component of the system uses the MPEG-7 document of the TV program and the user ontology (along with other domain ontologies such as the TV-Anytime classification schemes of MPEG-7) to infer which services should be activated during the broadcast. The Scheduler tells the Content Retrieval & Composition and the Multimedia components what actions they should perform.

The Reasoning component wraps the functionality of a reasoning and a rule engine. The first engine is required mainly for classifying the multimedia content and the user preferences to predefined categories, while the second one is used for deciding which services should be executed given the TV program metadata and the user profile. Bossam [14] is used as both a reasoning and a rule engine. Once the appropriate services have been selected for execution by the Reasoning component, it is the responsibility of the Scheduler component to coordinate the execution of the respective application logic. Such logic is registered in the Service Registry module through procedures specified by OSGi (Figure 3).

The Content Retrieval & Composition component is a framework for registering and managing interfaces with external information sources. For each new source that is registered (e.g., Web site, multimedia database, RSS feed), the available content is described along with its type (e.g., text, video) and the invocation details (e.g., URL, parameters).



(a)

(b)

Figure 3: (a) The Semantics Component, and (b) the TV User Ontology

2.2.3 System Integration with OSGi

POLYSEMA uses both OSGi services and MHP applications. The two platforms exhibit different properties because they follow different design principles, see [8] for a discussion. In order to retain full compatibility with industry standards such as MHP and DVB, and still harvest the service management flexibility of OSGi, we decided to base the component interaction on the IP return channel of the receiver. We incorporated a delegate bundle in the OSGi platform that conceals the nature of MHP applications from the rest of the system. If any other OSGi-based component of POLYSEMA wishes to interact with an MHP application, the MHP-delegate OSGi service should be accessed.

3 Provision of Interactive Services

This section gives an example of an interactive application that could be provided by POLYSEMA, and outlines its implementation. A content provider is assumed to supply the metadata describing the content, while the user enters their profile. POLYSEMA undertakes the responsibility of semantically matching descriptions with profiles, and activate appropriate services. Consider a user who includes in their profile their interest in cars. It could be requested that, in case of appearance of a car in a TV program the system should collect information about it from the Web (e.g.

Wikipedia). Figure 4 depicts an indicative screenshot of such an application. This is just one of the various possible alternative applications, but it illustrates how complex services can be supported by the proposed system.

Figure 4: Screenshot of the interactive service which collects web information.

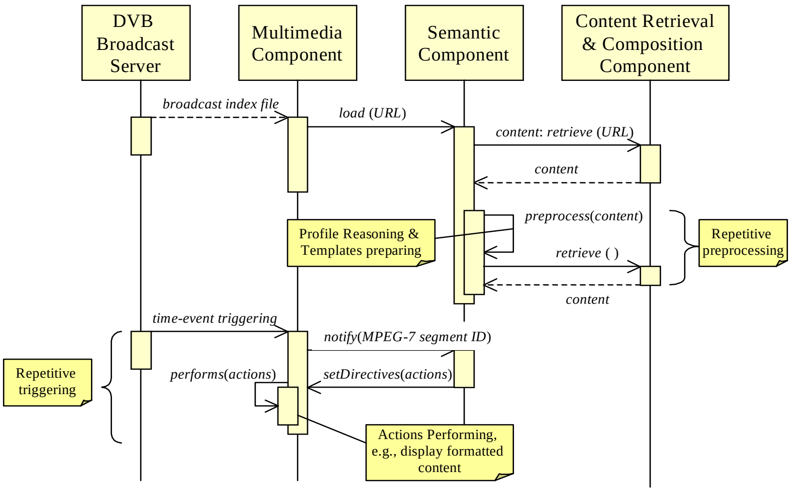


Figure 5: Sequence Diagram of POLYSEMA components functionality

Figure 5 depicts a sequence diagram illustrating the interaction of the Semantic and Multimedia components in order to outline the implementation of the aforementioned application. Specifically, the MPEG-7 document is assumed to contain detailed metadata that annotate the scene in which a car chases a small aircraft (Figure 4). A URL link to the MPEG-7 document is transmitted through the transport stream. At the beginning of the film, the MHP application running at the receiver requests that the Semantic component downloads the MPEG-7 from the Internet. Afterwards, the Semantic component creates an MPEG-7 ontology from the document, processes the descriptions of the film's scenes to match them with the user rules and produces the respective actions. As the broadcast advances, each film scene annotated by the MPEG-7 metadata (e.g., a description of a car make), is eventually displayed on screen. The corresponding time event is triggered at the receiver and the

MHP application requests from the Scheduler component to return the desired actions, i.e. the outcome of the preprocessing of the Reasoning component for that media scene. Subsequently, the Semantic component, supplies the desired content through certain layout templates gathered and formatted by the Content Retrieval & Composition component.

4 Prior Work

The AVATAR project discussed in [9] utilizes most of the techniques and technologies that the POLYSEMA project uses, such as video annotation, ontology-based modeling, multimedia metadata, and user profiling/personalization through semantics reasoning. The main objective of this project is to create a personalized digital TV program "recommender", based on the use of TV-Anytime formats and on techniques that have been widely used in the Semantic Web. The project has partial overlap with POLYSEMA, which is focused on a more broad range of applications.

MediaNET [10] is one of the major research efforts concerning multimedia networking. This project is divided in several sub-projects, each of which covers a significant area of the multimedia content creation, service providers and network operators. MediaNET emphasises on the broadcasting issues of the multimedia content and, as far as interactive services are concerned, it delivers the AmigoTV [12] service and a Personal Video Recorder (PVR). The project does not investigate the benefits from using metadata during the various phases of the multimedia content lifecycle. Although POLYSEMA and MediaNET share some infrastructure design issues their focus is on different aspects of the provision of iTV services.

The SAMBITS project [11] is implemented by some of the major telecasting organizations across Europe. The two main objectives of the project is the development of a set of tools for the creation of new interactive multimedia services that use MPEG-4 and MPEG-7 technologies, as well as the development of the technologies that would enable user terminals to access those services. On the other hand, POLYSEMA develops tools for creating multimedia semantics and focuses on the manipulation of such metadata *inside* the residential gateway.

There have been previous attempts to design systems combining both MHP and OSGi platforms. The work in [8] introduces a low-level implementation of a system, which is both OSGi and MHP compliant. The problem with this approach, however, is that, the reference implementation of the MHP platform had to be modified, and, thus, the system does not retain full compatibility with industry standards.

Recent work (see [15] for a comprehensive presentation of the respective ISO ammendment) defines different methods to carry metadata over MPEG-2. Metadata can be sent either by using private sections of MPEG-2, PES packets or the Broadcast File System. The latter approach was preferred, as it allows for in advance loading of the complete metadata, so that it can be timely preprocessed by the semantic component, before respective video scenes arrive. Moreover, in our design, the server can only send links to the metadata and not the entire resource. We believe that this model of metadata transmission is more appropriate, because it saves bandwidth for AV information in the TS, while the metadata can be fetched concurrently from an Internet connection available at the receiver.

5 Conclusions and Future Work

The research work carried out in the context of the POLYSEMA project is driven by the great importance of metadata in providing future iTV services and the need to manage them efficiently. Moreover, the POLYSEMA platform supports applications which adapt their behavior as the content presentation advances, allowing thus for innovative iTV services. Additionally, we believe that more effective personalization can only occur if the preferences of each user are known. This can only be achieved if semantic reasoning process takes place in the end-user premises. Future research may include an even more generic framework for designing services and integrating a variety of external web resources into the TV watching experience.

Acknowledgements

This work was partially funded by the Greek General Secretariat for Research and Technology (GSRT) under the Operational Program "Information Society", co funded by the EU.

References

[1] The Digital Video Broadcasting Project (DVB), http://www.dvb.org/.

[2] ETSI TS 102 812 V1.2.1, "DVB - Multimedia Home Platform (MHP) Specification 1.1.1".

[3] ETSI TR 101 190, "Implementation guidelines for DVB terrestrial services; Transmission aspects", November, 2004.

[4] "About the OSGi Service Platform - Technical Whitepaper Revision 4.0",www.osgi.org/

[5] Martinez, J, "Standards: Overview of MPEG-7 description tools", IEEE Multimedia, 2002.

[6] Antoniou, G., and van Harmelen, F. "A Semantic Web Primer", MIT Press, 2004.

[7] ETSI TR 101 202, v1.2.1, "DVB - Implementation guidelines for Data Broadcasting".

[8] Vilas, A. F. et al.,"MHP-OSGi convergence: a new model for open residential gateways", Software Practice and Experience, Vol. 36(13), November, 2006.

[9] Y. Blanco Fernández et al., "AVATAR: An Improved Solution for Personalized TV based on Semantic Inference", IEEE Trans. on Consumer Electronics. Vol. 52(1), February, 2006.

[10] Serge Travert, Michel Lemonier, "THE MEDIANET PROJECT", Proc. Image Analysis for Multimedia Interactive Services, Portugal, April, 2004.

[11] SAMBITS Project, "Deliverable 1: Project Description and Plan", IST1999-12605/Brunel/WP0/PU/P/001, 2000.

[12] AmigoTV: a social TV experience through triple-play convergence, Alcatel/Lucent, 2005.

[13] C. Tsinaraki, P. Polydoros and S. Christodoulakis, "Interoperability support for Ontology-based Video Retrieval Applications" C. In Proc. of Image and Video Retrieval, 2004.

[14] M. Jang, J Sohn, Bossam: an extended rule engine for the web, Proc. of RuleML 2004.

[15] A. Lopez et al. "Synchronized MPEG-7 Metadata Broadcasting over DVB networks in an MHP Application Framework". Proc. of Inrernational Broadcasting Convention, 2003..