Dynamic Personalisation of Media Content

Benedita Malheiro
Electrical Engineering Department
Instituto Superior de Engenharia do Porto
Porto, Portugal
e-mail: mbm@isep.ipp.pt

Jeremy Foss
School of Digital Media Technology
Birmingham City University
Birmingham, UK
e-mail: Jeremy.Foss@bcu.ac.uk

Juan C. Burguillo, Ana Peleteiro and Fernando A. Mikic
Telematics Engineering Department
Universidade de Vigo
36310-Vigo, Spain
e-mail: J.C.Burguillo@det.uvigo.es

Abstract—Dynamic personalisation of media content is the latest challenge for media content producers and distributors. The idea is to adapt in near real time the content of a video stream to the viewer’s profile. This concept encompasses any type of context-awareness customisation, expressed preferences and viewer profiling. To achieve this goal we propose a multi-tier framework composed of a content production tier, a content distribution tier and a content consumption tier, representing producers, distributors and viewers, plus an artefact brokerage tier, implemented as an agent-based e-brokerage platform, to support the dynamic selection of the content to be inserted in the video stream of each viewer.

Keywords: component; video on demand; personalisation; context-awareness; agent-based brokerage; B2B; metadata.

I. INTRODUCTION

The goal of this paper is to propose a framework for the dynamic personalisation of media content. The players in this challenging scenario are the media content producers, the media service distributors and the viewers. The transacted objects are video objects (streams). From the point of view of the system there are, however, two sorts of video objects: the user-selected stream and the external streams that will be dynamically inserted according to the user profile.

Personalisation relies heavily on the adoption of appropriate video metadata descriptions as well as on the construction and maintenance by the content producers of large libraries containing compliant video objects. On one hand, the user-selected source video object must include placeholders indicating the properties and the location of the insertion of the personalised external objects. On the other hand, all video objects must be appropriately annotated in order to support the proposed approach.

The profiling of the viewers is continuously performed by the system and is based on all available user interactions, namely the user selected media streams, navigation patterns and social networking.

The identification of the potential candidate objects to be inserted in a given play-out stream and the selection of the actual object to be inserted according to the user profile is performed by an agent-based brokerage platform. It is through this platform that content producers and distributors meet, negotiate and transact the personalised video objects.

Finally, the acquired objects are automatically inserted into the user play-out streams, substituting the placeholders and providing the viewer with a personalised program. This personalisation may contemplate cultural, political, age, gender or religious expressed preferences as well as context-aware and web-based inferences.

II. BACKGROUND

Near real time personalisation of media content is a complex problem involving knowledge representation, dynamic profiling, matching, negotiation and video insertion.

A. Related Projects

The main research efforts in this field focus on system design and knowledge representation ([1], [2], [3], [4], [5] and [6]) and on personalised search and recommender systems ([7], [8], [9] and [10]) projects:

COAST [1] intends to develop a content-centric network overlay architecture linking content sources to content consumers, including content-aware retrieval, delivery and streaming. CAM4HOME [2] proposes a metadata enabled content delivery framework to support end users and commercial content providers to create and deliver rich multimedia experiences. ACDC [3] goal is to develop user-aware content-delivery architecture for content navigation and personalisation based on semantics, online media resources and relying on the usage of new infrastructures based on cloud. Boemie [4] addresses the automatic extraction of semantics and annotation of multimedia content to allow multimedia search both to content providers and content consumers. aceMEdia [5] is a framework combining advances in knowledge, semantics and multimedia processing technologies to support self-analysing, self-annotating and self-adapting content.
OMWeb [6] intends to provide personalised, just in time media objects, constructed on the fly from distributed components and according to user preferences.

MyMedia [7] provides dynamic personalization of multimedia search via recommendations that are integrated from many sources based on user preferences. SAME [8] uses the context of the sonic environment (street, restaurant, car, meeting) to provide the end-user with a personalised playlist. GLOCAL [9] manages media according to context and events, catching media semantics in a natural way and sharing consistent event schemas built upon local and global interaction, supporting user access to the requested media. SAMBITS [10] integrates MPEG-2, MPEG-4 and MPEG-7 technologies and complementary Internet services to enable a personalized and easier access to multimedia applications.

B. Representation of Media Objects

Producers and distributors should use a common ontology for the representation of media content (metadata). MPEG-7, formally named Multimedia Content Description Interface, is a standard developed by the Moving Picture Experts Group (MPEG) for the structural and semantic description of multimedia content [11]. However, since MPEG-7 is an XML Schema, it lacks formal semantics. To overcome this problem, several MPEG-7 based ontologies have been proposed during the last decade: the Harmony ontology [12], SmartWeb project ontology [13], the Boemie project ontologies (Multimedia Content Ontology and Multimedia Descriptors Ontology) [4], DS-MIRF ontology [14], Rhizomik ontology [15], aceMedia project RDFS ontologies (Multimedia Structure Ontology and Visual Descriptor Ontology) [5] and Core Ontology for Multimedia (COMM) [16] based on both the MPEG-7 standard and on the DOLCE foundational ontology [17]. The different modelling approaches followed by these ontologies illustrate the significance of adopting formally founded standardised description models for practical semantic metadata management services [18]. As a result, we intend to adopt an MPEG-7 based ontology for the representation of all media content.

The metadata of the source content describes the restrictions and requirements and characteristics of objects which will be sought for integration into the content. Imported object properties are described by their own metadata. Along with the viewer profile, a selection is made for a suitable object to insert into the source stream. The scene graph is then updated to include references to the objects which have now populated the placeholders in the source content. The user-end equipment will utilise the scene graph to assemble the personalised stream.

C. Personalisation and Filtering

Nowadays, there are several techniques for filtering and personalization of contents:

- Content-based Filtering (CBF) uses the description of the resource and the user’s interests to provide recommendations [19]. These recommendations do not take care about the information provided by other users; therefore the description of the resources is really relevant to provide precise recommendations.
- Collaborative Filtering (CF) techniques usually consider the comparison of ratings, provided by the users to the resources, with other similar users (concerning their profile) in order to produce recommendations. User-based valuations or resource-based valuations are the two main alternatives in this approach. Recently, some hybrid approximations have been proposed, that usually are a combination of content recommendation and collaborative filtering [20], [21].
- Collaborative Tagging (CT) systems allow users to describe contents by means of tags and to share such description. These systems generate two different types of structures: tag clouds and folksonomies [22]. These structures can be used for content recommendation, and we can distinguish between two different approaches: (i) systems that use tagging information to improve recommendation algorithms [20] where we can find the ones that consider the number of resources tagged and the numbers of tags used [23] and (ii) works where only tagging information is used for recommendation [24].
- The Case-based Reasoning (CBR) paradigm covers a range of different methods for organizing, retrieving, utilizing and indexing the knowledge retained in past cases [25]. The CBR approach focuses on inexact reasoning by a similarity measurement among cases, and integrates four phases: Retrieve, Reuse, Revise and Retain. Basically, the system retrieve previous cases similar to the new one presented, reuse the closer ones (adapting them to the new context), revise the proposed solution(s), and finally, if it has become successful, the system retains the new solution for future use. Cases may be kept as concrete experiences, or a set of similar cases may form a generalized case. CBR techniques provide a good solution to suggest recommendations that worked well in the past, taking into account user context, user profile and resource descriptions.

D. Electronic Brokerage

To negotiate the external video objects to be inserted, we propose an agent-based e-brokerage platform. E-brokers are electronic intermediaries or cybermediaries that facilitate exchanges between buyers (consumers) and sellers (producers) by meeting the needs of both parties [26]. In August 2009, a survey conducted by a not-for-profit project funded by the trade promotion organizations of various countries, produced a non exhaustive directory of over 628 active electronic marketplaces around the world [27]. Agent-based e-marketplaces and, specifically, agent-based brokerage systems have been around for more than one decade. Whereas Kasbah [28], AuctionBot [29], Multi Attribute Resource Intermediary (MARI) [30], Global Electronic Market Stands (GEMS) [31] and the Multi AGent
NEgotiation Testbed (MAGNET) [32] are examples of agent-based e-commerce platforms – see [33] and [34]. Agent-based BROkerage Services in Electronic commerce (ABROSE) [35], Object Framework for Electronic Requisitioning (OFFER) [36] and the Multimedia Publishing Brokerage System [37] are instances of e-brokerage systems supported by agent technology – see [38]. However, as a significant sector of e-commerce, e-brokerages are still a relative new and poorly understood type of business [26]. This is the case of media brokerage advertising services that are still overwhelmingly conducted by humans that bring together advertisers and publishers and facilitate transactions for a fee. The existing so-called media e-brokerage systems are mainly for resource allocation and not for media content delivery. This means that there are currently no commercial platforms performing the proposed electronic media content brokerage services.

E. Insertion of Video Objects

There are two specific video technologies which can be utilised for near real time video insertion: (i) frame-based overlays in which objects are overlay images, e.g., Adobe Flash; and (ii) object-based video supported by MPEG-4 extensions. Our preferred approach is object-based video. Digital Media Broadcasting (DMB) is a service which currently employs these technologies, although not for the type of personalised services we envisage here. In a typical DMB implementation transmitted video objects are multiplexed via the Delivery Multimedia Integration Framework (DMIF), defined in MPEG-4 part 6 [39]. Objects are related to each other in their video scene and in their temporal relationships by the Binary Interchange Format (BIFS), specified in MPEG-4 part 11 [40], or through the Lightweight Application Scene Representation (LASeR), defined in MPEG-4 Part 20 [41]. The user-end equipment (i.e. computer or set-top box) will reassemble the objects into the desired scene for playback to the user interface. Objects may be given specific properties, for example for user interaction.

In this proposal personalisation is achieved by selecting different objects and integrating them into the "source video" play-out stream. This requires the source video to be produced with suitable spaces in the content where imported objects may be placed; the spaces are "placeholders" defined in the source video. The properties of the placeholder define the type of object which may be inserted, and would be in accordance with the editorial integrity of the content as specified by the producer of the source video.

III. NETWORK ARCHITECTURE

All current video distribution formats would be feasible, including Satellite TV, Cable TV, Telco IPTV, WebTV (Over-The-Top TV) and Terrestrial TV. Figure 1 shows a potential network implementation. The video head-end will be unaffected by the requirement to host the source content stream. The architecture is augmented by an "integration server" which (i) receives the source video file from the head-end, and (ii) will request and retrieve the externally acquired video objects from third party video object servers (i.e., commercial libraries).

![Figure 1. Network Architecture](image1)

Video objects are selected on the basis of (i) the source metadata requirements; (ii) the external video object metadata description and (iii) the profile of the viewer. Objects are arbitrary shaped video objects [42] which are video files to be multiplexed with other objects and the source streamed from the head-end.

The integration server then re-multiplexes the new objects and recalculates the new scene graph and then streams the modified content and scene graph to the end user. For any source video, this process will be repeated for each end user, but with different objects selected on the basis of the user’s personal profile. In this proposed architecture there are no proprietary coding standards, and no development to existing head-ends other than the additional integration server. Implementation of the personalised video distribution is therefore easily achievable in existing network deployments.

When several viewers are receiving a video stream of a broadcast or on-demand program, each viewer will receive the requested program. The viewer may interact with the objects, if allowed by the objects or the source content producer. Viewers may also participate as user groups, where inserted objects are relevant to a group of users.

IV. SYSTEM OVERVIEW

Our approach to dynamic personalisation of media content is supported by the generic architecture presented in Figure 2.

![Figure 2. Architecture](image2)

This multi-tier architecture is constituted of four main tiers: the content production tier, the content distribution tier, the content consumption tier and the artefact brokerage tier. The key players are the producers, the distributors and the viewers.
A. Representation of Video Objects

In this framework video objects are MPEG-4 instances annotated in an MPEG-7 based OWL ontology. This applies both to the source video objects (the user-selected video streams) and to the external video objects (automatically selected and inserted by this framework).

B. Representation of Producers and Distributors

The producers and distributors of media content are modelled by autonomous intelligent agents. These so-called enterprise agents must, on one hand, be entirely controlled by their real world counterparts to ensure the privacy of the company strategic knowledge and, on the other hand, be fully compatible and interoperable with the remaining components of the framework. The latter is achieved by the adoption of a Web Service interface guaranteeing interoperability and allowing the creation of loosely coupled enterprise agents that can enter and leave freely the proposed transaction environment. The resulting Service-Oriented Architecture (SOA) relies on Universal Description, Discovery and Integration (UDDI) service registries to hold the descriptions of existing agent services. On one hand, producers and distributors can publish, update and remove their service descriptions – metadata descriptions of the objects they hold or seek to insert in the viewer stream. On the other hand, any entity can discover, download and interact with any service (agent) automatically.

C. Viewer Profiling

In order to achieve a good viewer profiling, and to provide him the best possible object insertion experience, we will use a combination of some techniques that we have applied successfully but isolatedly to other contexts in the past ([21], [22], [43] and [44]). Of course, we will adapt them to consider the particularities of the media scenario.

The approach we propose considers the creation and maintenance of users’ profiles as CBR cases containing information about personal, social, professional, cultural, political, or religious data. On the one hand, the use of CBR techniques can be very useful when managing ontologies and semantic data. On the other hand, with Web 2.0 profiling has evolved from an individual to a social activity, and the profile of any user can be enhanced by social networking, especially by means of folksonomies, that are the structures that emerge thanks to the use of collaborative tagging [22].

Therefore, in order to find the best possible objects to insert within the main stream, we model every user as a CBR case that contains all his/her personal and semantic data, together with his tag cloud resulting from his social interactions in the Web or from the previous selection of video streams he/she has done. From all this data, a CBR engine will use classical filtering techniques [21] enhanced with collaborative tagging [22] to search for recommended objects to be inserted. In case that the objects are distributed over a set of databases, we can follow a P2P scheme, as the one that we describe in [43] to perform the distributed search, selecting the most appropriate objects concerning their characteristics and the user profile.

The viewer profiling activity is carried out by profiler agents launched by the distributors to monitor all relevant user interactions – see Figure 3.

![Figure 3. Viewer profiling](image)

By relevant interactions we mean all types of user interaction that can be filtered by the distributor: selected video streams, Internet browsing activity, social networking, etc. The profiler agents rely on the proposed approaches to create and maintain the individual viewer profiles.

D. Agent-based Brokerage

The brokerage platform is a competitive multi-agent system where enterprise agents (producers and distributors referred to as AgProd and AgDist, respectively) and market regulator meet in order to trade media components according to the market profiles of the agents and the rules of the market. The platform, presented in Figure 4, is structured in two layers.

![Figure 4. Brokerage platform](image)

The top layer is composed by coarse grain agents representing the producers and distributors – the enterprise agents and market profiler agent – and, the bottom layer, is made of a set of finer grain agents constituting the marketplace – the delegate transaction agents and the market agent. Figure 5 presents the top layer in further detail.

To ensure interoperability and platform independence, all enterprise agents have a Web Service interface. Whenever a producer or a distributor adopts our framework, a dedicated top layer agent is created. The enterprise interacts via the Web Service interface and retains full control of its representative. Additionally, the enterprise agent registers in the UDDI Market Service Registry the description of the services it provides using the adopted ontology. Producers
and distributors use this ontology to classify the video objects they hold or the video objects they seek, respectively.

![Multimedia Annotation and Retrieval enabled by Shared Ontologies](image)

Figure 5. Top layer

The bottom layer constitutes the marketplace that is governed by the market agent and populated by delegate agents.

The market agent periodically checks the Market Service Registry for producer (external videos) objects matching distributor (placeholders) objects and profiles of connected viewers. The selection of the object to be inserted relies on a market-based negotiation. The identified distributors and producers are then invited to send delegate agents to the marketplace to negotiate these specific objects. A delegate agent negotiates, according to the specified user profile, a unique object – either a placeholder or a video – and has an ephemeral life – once the negotiation ends, reports back the outcome to its master and terminates.

V. CURRENT STATUS

This framework is under development. Our current stage is focussed on the development of the brokerage platform. It is being developed in Java Agent DEvelopment Framework (JADE), which is a software framework to develop agent-based applications in compliance with the Foundation for Intelligent Physical Agents (FIPA) specifications for interoperable intelligent multi-agent systems [45]. The multi-agent system itself, i.e., the different type of agents, their market profiles, actions and relationships also share a common OWL ontology. The media components are annotated in a MPEG-7 based OWL ontology and the agents rely on an automated negotiation approach that suits the market type [46].

Producers and distributors update frequently the state of their representative agents by specifying new strategic knowledge to improve market behaviour (market profile) and updating the lists of external objects available for transaction (producers) or the lists of needs identified by integration servers (distributors).

Currently experiments are being conducted with the Iterated Contract Net negotiation protocol. Other negotiation protocols (types of markets) such as auction-based protocols will follow. Although media objects can be negotiated under several dimensions such price, resolution, delivery time, etc., they are currently based only on price.

So far, video objects and viewers have been classified based on standard socio-economic classifications A, B, C and D [47]. Although this is far from our intended level of personalisation, it allows us to illustrate the concept.

VI. CONCLUSION

The proposed framework constitutes an ambitious solution to the dynamic media content personalisation challenge. There are a number of issues which need to be approached in the realisation of the interactive personalised services proposed in this paper, including codecs to support object-based video and video head-end equipment to support the object insertion.

Content producers will also require intuitive authoring and editing tools to manipulate programme material and stipulate the conditions for which external objects may be imported into their creative output. The proposal given in this paper allows a scalable market model to cover the global media industry across the full range of TV distribution methods including cable, satellite and Telco IPTV.

The proposal relies on existing standards such as Web Services, MPEG-4 for media encoding, MPEG-7 based OWL ontology for content description and standard negotiation protocols. The design concept is modular, open and the technologies selected are open source and Java based. Furthermore, the adoption of a Semantic Web (ontology-based) approach for the knowledge representation intends to contribute to promote the interoperability with other systems and to allow future expansions.

ACKNOWLEDGMENT

This work was partially supported by the Xunta de Galicia under INCITE Project Number 10 PXIB 322 039 PR.

REFERENCES


