Scalable multimedia delivery with QoS management in pervasive computing environment

Hongguang Zhang, Hang Nguyen, Eduardo Martinez Gracia, Pedro Antonio Tudela Solano, Daqing Zhang, Noel Crespi

Scalable Multimedia Delivery with QoS Management in Pervasive Computing Environment

Hongguang Zhang[1], Hang Nguyen[1], Eduardo Martínez Graciá[2], Pedro Antonio Tudela Solano[2], Daqing Zhang[3], Noël Crespi[1]

[1]*Wireless Networks and Multimedia Services Department, Institut Telecom, Telecom SudParis, 9 rue Charles Fourier, 91000 Evry, France*

[2]*Department of Intelligent Systems and Telematics, Faculty of Informática, University of Murcia, 30100 Campus de Espinardo, Spain*

[3]*Telecommunication Network and Services Department, Institut Telecom, Telecom SudParis, 9 rue Charles Fourier, 91000 Evry, France*

Tel: +33 (0)1 60 76 66 63

Fax: +33 (0)1 60 76 45 78

Email: hongguang.zhang@it-sudparis.eu

Abstract

People are increasingly demanding rich-media and bundled services. However, the diverse terminals, heterogonous networks as well as various user requirements constrain the multimedia access to low quality in the pervasive computing environment. In order to enable rich-media delivery across a wide range of devices and networks, multimedia adaptation with scalable QoS management is an important issue. Our research introduces a Scalable Multimedia Delivery (SMD) framework with QoS management. This framework utilizes the CAM4Home metadata model to aggregate multimodal rich media services into a bundle. MPEG-21 metadata is integrated into the CAM4Home model to enforce interoperable QoS management. The issues in supporting QoS are addressed on both fidelity and modality. We further develop the SMD system in IP Multimedia Subsystem (IMS) architecture, where multimedia adaptation is implemented through application-level QoS negotiation.

Keywords - Multimedia Communication, Multimedia Adaptation, Quality of Service, IP Multimedia Subsystem

1 Introduction

The past decade has witnessed a revolution in mobile multimedia technologies. With the proliferation of mobile devices and deployment of wireless networks, people can access multimedia contents anytime and anywhere, which is the vision of Universal Multimedia Access [1-2]. Under this trend, wireless networks and multimedia services are converging to a ubiquitous multimedia system. Unlike traditional data, multimedia contents have more stringent Quality of Service (QoS) requirements. However, the diverse terminals, heterogonous networks as well as various user requirements constrain the multimedia access to low quality in the pervasive computing environment. Meanwhile, the emerging pervasive services are getting more and more complicated when the rich-media services are aggregated, known as service composition [3-5]. As a result, a uni-model for QoS provisioning is not suitable for the application-driven and mission-critical pervasive services. It is desired to have a scalable and quality-assured multimedia delivery system.

A ubiquitous multimedia system need adapt multimedia contents to various constraints of terminals and networks as well as user preferences, while providing the best possible quality to the end user. T.C. Thang et al. [6-8] have intensively studied the quality in multimedia content delivery. They identified the quality from two aspects: perceptual quality and semantic quality. The former known as fidelity refers to a user's satisfaction, while the latter is the amount of information the user obtains from the content. The semantic quality is sometimes referred as Quality of Experience (QoE) or Information Quality (IQ). In some cases, the perceptual quality of an adapted content is unacceptable or its semantic quality is much poorer compared with that of a substitute modality. A possible solution for this problem is to convert the modality of contents. For example, when the available bandwidth is too low to support the video streaming service for a football match, the text-based statistics service would be more appropriate than the adapted video with poor perceptual quality. This is a typical case of video-to-text modality adaptation. Apparently, the combination of fidelity and modality can enhance user experiences. John R. Smith et al. [9] proposed a scalable model named InfoPyramid, which manages the different variations of media objects with different fidelities and modalities in order to adapt the delivery to different client devices. However, it considered only the diversity of terminal devices, which is far less than enough for heterogeneous networks. And the InfoPyramid model is neither standard-interoperable nor network-applicable. Considering the diverse devices and heterogeneous networks, we argue that a scalable model with end-to-end QoS management is critical for rich-media delivery in the pervasive computing environment. Without such as a model, the quality-assured pervasive services would not be really feasible, especially from the viewpoints of service providers. If the service providers could specify a service model starting service requirements, they would be relieved from implementing complex QoS decision making functionality for each new service being introduced, hence leading to simplified provisioning and quicker time-to-market for new services [10].

In this paper, we present a Scalable Multimedia Delivery (SMD) framework with QoS management in the pervasive computing environment. The advanced multimedia delivery supports the aggregation of multimodal media including video, audio, image, and even text. Most prominently, the framework provides scalable QoS management for multimodal media delivery. The concept of scalability in this paper means that the rich-media contents can be tailored and adapted with the assured quality to diverse terminals and heterogeneous networks. We address the issues in supporting QoS from two levels, namely fidelity and modality. The former supports quality adaptation to guarantee QoS, while the latter provides modality transformation to maximize QoE. To support this notion of scalability, we utilize CAM4Home [11] as the model that combines multimodal media into a bundle on the level of metadata. CAM4Home is an ITEA2 project implementing the concept of Collaborative Aggregated Multimedia (CAM). We further propose to integrate MPEG-21 metadata within the CAM4Home model in order to enable interoperable QoS management. Accordingly, MPEG-21 metadata aims to provide the best user experiences across multimodal services. To verify the proposed model is network applicable, we develop the SMD prototype system based on IP Multimedia Subsystem (IMS) architecture. SMD is implemented during session negotiation by two-step adaptation, pre-session and in-session.

The rest of the paper is organized as follows. Section 2 reviews the background and related works. In Section 3, we describe the requirements on SMD and present the scalable service model. Section 4 discusses the approach to apply the metadata-based SMD model into IMS service architecture. A prototype system and the performance evaluation are described in Section 5. Section 6 concludes the paper and presents some issues for future research.

2 Background and Related Works

Providing QoS support is one of the most challenging problems towards ubiquitous multimedia. The issue of QoS can be addressed at different levels [12], namely user level, application level, transport level and network level, as well as across different stakeholders [13], namely content provider, service provider, network provider, and end user. In terms of application level, dynamic adaptation of multimedia content is seen as an important feature of pervasive systems enabling terminals and applications to adapt to changes in access network, and available QoS due to mobility of users, devices or sessions [14]. Generally, it is a computing intensive process for adaptation decision-taking involved for choosing the right set of parameters that yield an adapted version. The computational efficiency of adaptation can be greatly enhanced if this process could be simplified, in particular by metadata that conveys pre-computed relationships between feasible adaptation parameters and media characteristics obtained by selecting them [15]. Moreover, the development of an interoperable multimedia content adaptation framework has become a key issue, in order to cope with this heterogeneity of multimedia content formats, networks and terminals. To address above problems, MPEG-21 Digital Item Adaptation (DIA) specifying metadata for assisting adaptation has been finalized as part of the MPEG-21 Multimedia Framework [16].

The past researches on multimedia adaptation are more concerned with the perceptual quality from the aspect of end user. However, the intensive studies in [6-8] state that the semantic quality should be considered in some cases. They argue that modality conversion could be a better choice than unrestricted adaptation on fidelity. The Overlapped Content Value (OCV) model is introduced in [6] to represent conceptually both content scaling and modality conversion. The past research works made most efforts to optimize mathematic model for modality adaptation [6, 15, 17], focusing on two questions “when to change the modality?” and “what is the substitute modality?” However, the question on “how to do modality adaptation” is usually ignored, especially from the viewpoints of content providers and service provider. Apparently, the combination of fidelity and modality can bring the best user experiences. John R. Smith et al. [9] presented the InfoPyramid model that support multimedia delivery with different fidelities and modalities in order to adapt the delivery to different client devices. However, it considered only the diversity of terminal devices. Moreover, the proposed model is neither standard interoperable nor network applicable.

Towards all-over-IP network architecture, IP Multimedia Subsystem (IMS) [18] has been recognized as the service architecture for Next Generation Networking (NGN). Many works related to QoS have been studied for IMS-based multimedia delivery. Lea S.K. et al. [19] described the application-level QoS negotiation and signaling for advanced multimedia services in the IMS. T. Ozcelebi et al. [20] presented a solution for multimedia streaming quality adaptation using resource management. J. Arnaud et al. [21] introduced an adaptation of IPTV service according to the perceived QoS degradation at the side of end user. Even though session negotiation mechanisms used above ensure QoS to some extent, it is still not enough for multimedia applications in pervasive computing environment due to the diversity of end devices and highly resource-consuming multimedia contents. Firstly, QoS negotiation occurs at the layer of signaling and control. However, it is the service provider who really knows and manages the quality of multimedia service. Secondly, IMS services are originally designed by service providers for specific end devices or access networks. It is time to move from special-purpose, one-of-a-kind to more widely deployable that can scale to user context. Therefore, mechanisms are needed providing a more advanced decision-making process based on matching restrictive user parameters, service requirements, and network constraints, with the goal of achieving maximum user experiences [19].

In this paper, we concentrate on both the user level by specifying user perceivable service parameter and the application level by adapting contents according to the resource availability of terminal and network. We present a SMD framework enforced by interoperable QoS management. The proposed model considers both fidelity and modality to meet QoS requirements in the diverse terminals, heterogeneous networks as well as dynamic network conditions.

3 Scalable Multimedia Delivery Model

This section firstly describes the concept of SMD in detail. The SMD model is further decomposed into two essential parts: multimodal service aggregation and metadata-based QoS management.

3.1 Concept of Scalable Multimedia Delivery

The scalability in this paper is proposed on two levels: fidelity and modality. To support this notion of scalability, we consider a service as a collection of multimodal service components which are semantic-coherent but independent services offered for certain end devices, access networks or user preferences. The breakdown of a service into service components supports its adaptation for a given device type, network or user by allowing the sorting, filtering and adaptation of its individual components [24]. Let's take "Sports Live Broadcasting" service as an example. The scenario is the last round of the football league where more than one team has the chance to win the champion. All teams start playing at the same time. Fans are watching the live TV broadcasting of their team. Obviously, they may also want to be updated on the information (e.g. goal, penalty, and red card etc.) of other simultaneous matches. The "Sports Live Broadcasting" application contains two sub-services. The first one is an IPTV program delivering a live football game, which is designed for high-end devices with wideband network. The second one is a real-time literal broadcasting service delivering statistics data synchronized to all football matches, which is designed for devices with narrowband network like mobile phones. Furthermore, the IPTV service component can be configured by a set of offered alternative operating parameters (e.g., frame sizes, frame rates and bit rates etc.), which can be adjusted dynamically according to user context. Before multimedia session, the scalable service model firstly selects the service version according to the static capabilities of terminals or networks. During session, this service element of IPTV can be adapted according to dynamic network condition or user preferences. Moreover, if the adapted IPTV service cannot provide the expected user-perceived quality, a cross-modal adaptation from IPTV to Text may occur.

The requirements for SMD exist in three aspects. The first one is related to user requirements. People are demanding rich-media and bundled services that extend well beyond voice to include text, image, audio and video [22]. The second requirement is related to service requirements [23]. For example, it is desirable for the service provider to design an application customized for different user contexts. To maximize user experience or information quality, there is a need for the support of cross-modality adaptation in case a terminal or network cannot support the consumption or transport of a particular modality. The third one is related to network requirements. QoS needs to be adapted to dynamic network conditions in real time. The concept of SMD covers the whole value chain across different stakeholders, namely content provider, service provider, network provider, and end user. In this paper, we are concerned more with the service provider.

To sum up, SMD is a sender-driven service model, which aims to create and deliver advanced multimedia experiences tailored to heterogeneous user devices, varying network conditions and diverse user requirements. This service model requires a service profile aggregating multimodal media types associated with pre-computed QoS parameters. There are two essential parts inside: multimodal service aggregation and metadata-based QoS management.

3.2 Multimodal Service Aggregation

The essential part of SMD is the multimodal service aggregation. In this paper, we use CAM4Home framework as the metadata model for multimodal service aggregation. The CAM4Home is an ITEA2 project implementing the concept of Collaborative Aggregated Multimedia (CAM) [11]. The concept of CAM refers to aggregation and composition of individual multimedia contents into a content bundle that may include references to content-based services and can be delivered as a semantically coherent set of content and related services over various communication channels. This project creates a metadata-enabled content delivery framework by bundling semantically coherent contents and services on the level of metadata. The CAM4Home metadata model supports the representation of a wide variety of multimedia content and service in CAM Element as well as its descriptive metadata. CAM Object is the integrated representation of CAM Element and CAM Element Metadata on the association rule "*isMetadataOf*". CAM Bundles are the aggregation of two or more CAM Objects on the association rule "*containsCAMObjectReference*". CAM Object and CAM Bundle can be uniquely identified by "*camElementMetadataID*" and "*camBundleMetadataID*". Fig. 1 illustrates a conceptual view of CAM Bundle and CAM Object.

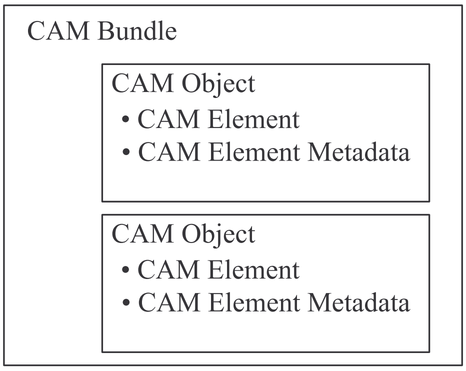

Fig. 1 Conceptual view of CAM Bundle and CAM Object

The taxonomy of CAM Element has two subclasses, Multimedia Element and Service Element. The Multimedia Element is the container of a specific multimedia content, which is further divided into four types, document, image, audio and video. The Service Element is the container of a specific service. The physical content in CAM Element is referred by the attribute "*EssenceFileIdentifier*" which is a Universal Resource Locator (URL). Actually, the physical content can be referred by the URL. The Service Element includes the other attribute "*ServiceAccessMethod*" indicating the methods used to access the service. With the instinctive of CAM, we use the metadata-based approach for the content and service delivery. For example, the attribute "*EssenceFileIdentifier*" can be used to indicate the Public Service Identity (PSI) of an application server in IMS. And the other attribute "*ServiceAccessMethod*" indicates the Session Initiation Protocol (SIP) methods (e.g. INVITE) accessing the service. However, the services are not limited to SIP-based. This model can be used to encapsulate any types of services.

In this paper, the CAM4Home metadata model is adopted as the rich-media aggregation model for SMD. Fig. 2 shows an example for the aforementioned "Sports Live Broadcasting" service**.** To be noted, CAM4Home provides an open metadata model to integrate with external metadata schemas. MPEG-7 metadata can be used to enrich content descriptions and facilitate content discovery. MPEG-21 can be introduced to enforce QoS management.

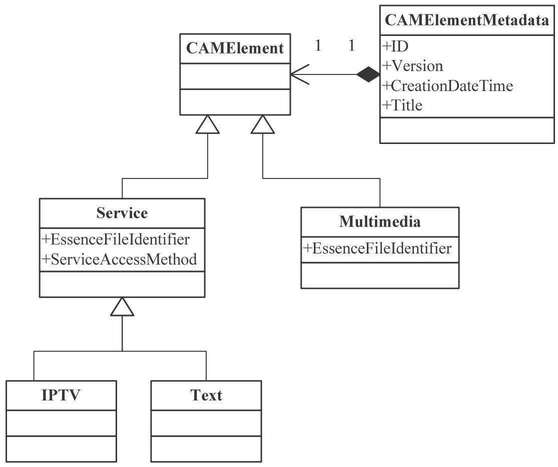

Fig. 2 CAM4Home metadata example

3.3 Metadata-based QoS Management

It is necessary to provide a quality-guaranteed and interoperable content delivery across converged complex devices and distribution networks as well as an optimized use of underlying delivery network bandwidth and QoS characteristics. Toward this purpose, we propose to integrate MPEG-21 DIA into CAM4Home model enabling QoS management. MPEG-21 DIA aims to standardize various adaptation related metadata including those supporting decision-taking and the constraint specifications [15]. MPEG-21 DIA specifies normative description tools in syntax and semantic to assist with the adaptation. The central tool is the Adaptation QoS (AQoS) representing the metadata supporting decision-taking. The aim of AQoS is to select optimal parameter settings that satisfy constraints imposed by a given external context while maximizing QoS. The adaptation constraints may be specified implicitly by a variety of Usage Environment Description (UED) tool describing user characteristics (e.g. user information, user preferences, and location), terminal capabilities, network characteristics, and natural environment characteristics (e.g., location, time). The constraints can also be specified explicitly by Universal Constraints Description (UCD).

Syntactically, the AQoS description consists of two main components: Module and Input Output Pin (IOPin) [25]. Module provides a means to select an output value given one or several input values. There are tree types of modules, namely, Look-Up Table (LUT), Utility Function (UF), and Stack Function (SF). IOPin provides an identifier to these input and output values. Originally in

MPEG-21 DIA, the output values are utilized by Bitstream Syntax Description (BSD) [26] for content-independent adaptation. However, in the proposed SMD model the adapted target is altered to CAM Bundle. As mentioned in 3.1, the multimedia adaptation is proposed on two levels: quality and modality. The quality adaptation is to adapt one of the aggregated service component adjusting QoS parameters. The modality adaptation is to select the most appropriate modality among aggregated multimodal services components. We propose to embed AQoS in each CAM Object for quality adaptation as well as to associate AQoS with CAM Bundle for modality adaptation. In this regard, for quality adaptation the output values (e.g. bit rate, frame rate, resolution) are utilized to yield an adapted version on a single service component**.** For modality adaptation we set "IOPin" as the "camElementMetadataID" that identifies a specific service component within the CAM Bundle. The model of multimedia adaptation is illustrated in Fig. 3. In the value chain of SMD, the service providers take the responsibility on specifying these QoS management parameters.

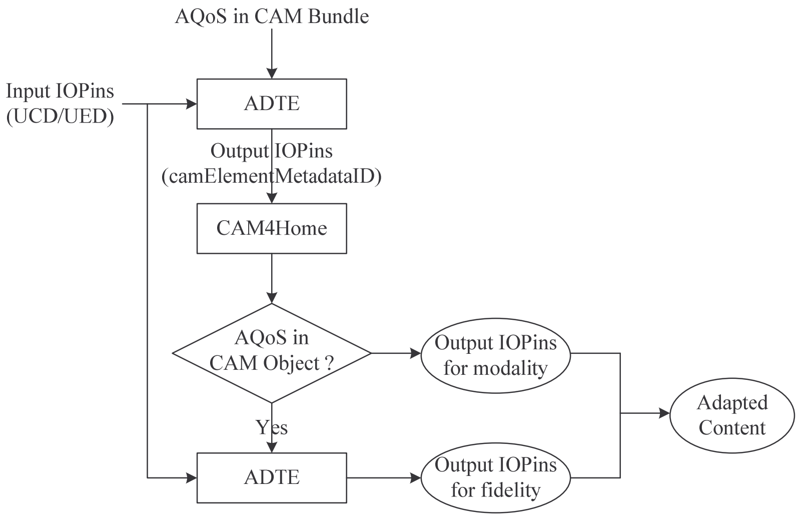

Fig. 3 Metadata-based multimedia adaptation

To further explain the use of MPEG-21 DIA in CAM4Home model, we take the aforementioned "Sports Live Broadcasting" service as an example, along with a part of AQoS description given in Fig. 4. We will not go into detailed explanation of the elements and attributes of the above example, as the names of these are self-explanatory and the complete specification of all these elements and attributes and their associated measurement units are defined in detail in the MPEG-21 DIA standard. The constraint for modality adaptation is the average available bandwidth of network condition as defined in [27]. The utility function is utilized to reflect the semantic quality. The utility can be pre-computed off-line as presented by [15, 28], which is beyond the scope of this paper. The constraints for quality adaptation include the width and the height of display capabilities.

```
<!-------------------------------------------------------------------------------
            MPEG-21 DIA AQoS Utility Function for Modality Adaptation
-------------------------------------------------------------------------------->
<DIA>
  <Description xsi:type="AdaptationQoSType">
    <Module xsi:type="UtilityFunctionType">
      <Constraint iOPinRef= "BANDWIDTH">
        <Values xsi:type="FloatVectorType">
          <Vector>25 50 125</Vector>
        </Values>
      </Constraint>
      <AdaptationOperator iOPinRef="camElementMetadataID">
        <Values xsi:type="IntegerVectorType">
           <Vector>1208327, 1208328, 1208328</Vector>
         </Values>
      </AdaptationOperator>
      <Utility iOPinRef="">
        <Values xsi:type="FloatVectorType">
           <Vector>2 1 3</Vector>
         </Values>
      </Utility>
    </Module>
    <IOPin id="BANDWIDTH">
      <!--Describes the average AvailableBandwidth as defined in this part of
ISO/IEC 21000-->
      <GetValue xsi:type="SemanticalDataRefType"
Semantics="urn:mpeg:mpeg21:2003:01-DIA-AdaptationQoSCS-NS:6.6.5.3"/>
    </IOPin>
    <IOPin id= "camElementMetadataID"/>
  </Description>
</DIA>

<!-------------------------------------------------------------------------------
              MPEG-21 DIA AQoS Look-Up Table for Quality Adaptation
-------------------------------------------------------------------------------->
<DIA>
  <Description xsi:type="AdaptationQoSType">
    <Module xsi:type="LookUpTableType">
      <Axis iOPinRef= "WIDTH">
        <AxisValues xsi:type="IntegerVectorType">
          <Vector>180 480 640 720</Vector>
        </AxisValues>
      </Axis>
      <Axis iOPinRef= "HEIGHT">
        <AxisValues xsi:type="IntegerVectorType">
          <Vector>144 320 480 576</Vector>
        </AxisValues>
      </Axis>
      <Content iOPinRef="SCALE">
        <ContentValues xsi:type="IntegerMatrixType" mpeg7:dim="4">
          <Matrix>1 2 3 4</Matrix>
        </ContentValues>
      </Content>
    </Module>
    <IOPin id="WIDTH">
      <!-- Describes the width as defined in this part of ISO/IEC 21000 -->
      <GetValue                                    xsi:type="SemanticalDataRefType"
Semantics="urn:mpeg:mpeg21:2003:01-DIA-AdaptationQoSCS-NS:6.5.9.1"/>
    </IOPin>
    <IOPin id="HEIGHT">
      <!-- Describes the height as defined in this part of ISO/IEC 21000 -->
     <GetValue xsi:type="SemanticalDataRefType" Semantics="urn:mpeg:mpeg21:2003:01-
DIA-AdaptationQoSCS-NS:6.5.9.2"/>
    </IOPin>
    <IOPin id="SCALE"/>
  </Description>
</DIA>
```

Fig. 4 MPEG-21 AQoS example

4 Scalable Multimedia Delivery Framework

In this section, we firstly overview the IMS architecture. Then we propose the approach to implement SMD based on IMS from two aspects: multimedia adaptation and session negotiation.

4.1 IMS Service Architecture

IMS has been recognized as the service architecture for NGN, offering multimedia services and enabling service convergence based on diverse network access technologies. With the advantage of an All-over-IP network, the opportunity for integration and convergence is amplified in IMS. The IMS service framework can be used to implement new services, but its main strength might eventually be in its ability to seamlessly integrate existing services and deliver them to end-users in an optimized manner. The IMS architecture is made up of two layers: the service layer and the control layer. The service layer comprises a set of Application Servers (ASs) that host and execute multimedia services. Session signaling and media handling are performed in the control layer. The key IMS entity in this layer is the Call Session Control Function (CSCF) which is a SIP server responsible for session control. There are three kinds of CSCF, among which Serving CSCF (S-CSCF) is the core for session controlling and service invocation. Home Subscriber Server (HSS) is the central database storing the subscriber's profile. Regarding the media delivery, the key component is Media Resource Function (MRF) that can be seen as media server for content delivery. In IMS, SIP is used as the signaling protocol and Real-time Transport Protocol (RTP) is used as the media transport protocol.

In order to support advanced media-rich applications across a wide range of user devices and access networks, IMS supports negotiable QoS for multimedia sessions. The goal of QoS negotiation is to determine "the best" service configuration and network resources allocation that would maximize user perceived service quality [19]. Technically, SIP is used for initiating and managing communication sessions. The application level signaling messages carry session descriptions which allow participants to exchange end-system capabilities and agree on a set of compatible media types. The session description is commonly formatted in Session Description Protocol (SDP) [29], while the session negotiation procedure is based on Offer/Answer model [30]. The syntax of

SDP is extensible and new attributes are added to the standard occasionally. It is possible to negotiate QoS parameters at the beginning of session by SIP signaling messages. Considering resource-consuming and real-time multimedia services, the pre-session QoS negotiation is not enough. It is also necessary to renegotiate or modify QoS parameters to satisfy dynamic constraints. As a result, the session renegotiation depending on network conditions or application requirements is also specified.

In the IMS service architecture, AS hosting a specific service can be seen as a provider of media modality. Hereafter, an IMS service in this paper is referred to a content-based service. As a result, the SMD framework should firstly support the combined delivery of multimodal services based on CAM4Home model. Further, the QoS management enforced by MPEG-21 DIA metadata should be applied into IMS service architecture. Especially, the cross-modal adaptation is implemented as service switching among aggregated services. AS also interacts with MRF in order to ensure the adaptive delivery of media. Fig. 5 illustrates the conceptual SMD framework in IMS.

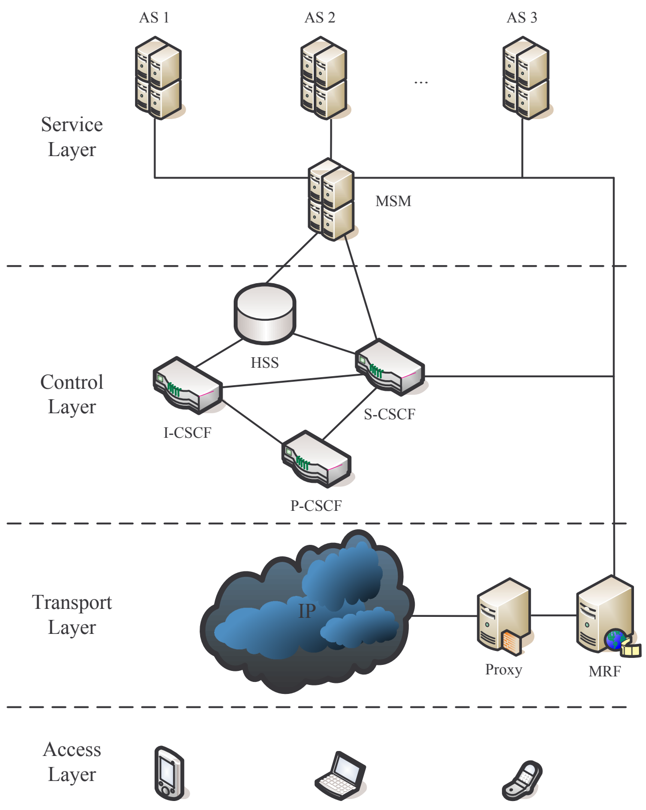

Fig. 5 IMS-based SMD framework

4.2 IMS-based SMD Framework

4.2.1 Adaptation Framework

The essential component of SMD is Multimedia Service Management (MSM) shown in Fig. 5. This component is evolved from our previous research work [31] where MSM implements the metadata-based multimodal content delivery in IMS. As mentioned in Section 3.1, SMD is a receiver-driven service model. Therefore, MSM is firstly proposed as a generic component of Service Deliver Platform (SDP), responsible for service-related functionalities, such as service registration and service discovery etc. Services represented as CAM metadata entities (e.g. object or bundle) are registered in MSM. To service providers, the rich semantic information may facilitate service composition and service discovery. To end users, they may also query a specific service registered in MSM, which helps them select or mash up services. The service repository holds both service objects and service bundles. To be noted that the service repository can be in MSM or in an external database alternatively. For instance, the CAM4Home project provides a web service platform for metadata generating, storing and searching. In this case, MSM needs to access the external platform through web service interface.

Besides above functionalities, the vital role of MSM is service routing. MSM provides address resolution decision-making on ASs. As shown in Fig. 5, MSM is located between S-CSCF and AS. For the consideration of scalability and extensibility, we collocate MSM in a SIP AS behaving as B2BUA. One hand, MSM is configured to connect with IMS. On the other hand, MSM interfaces with SIP ASs which host those aggregated service elements. In order to enable SMD, we extend MSM mainly from two aspects: Adaptation-Decision Taking Engine (ADTE) and UED collecting. ADTE either selects appropriate content modalities among the aggregated service components or to choose adaptation parameters for a specific media service. Additionally, MSM needs to collect UED as inputs of ADTE. For modality adaptation, MSM can act on the incoming requests and route them to AS according to the outputs of ADTE. Thanks to MPEG-21 QoS management, it is more intelligent compared with the routing criterion in [31] where it is based on the user requested service element.

For quality adaptation, we hereafter take video as the target considering video is the most challenging media type. We introduce a Real-time Streaming Protocol (RTSP) proxy, as shown in Fig. 5. RTSP is an application-level protocol for control over the delivery of media data with real-time prosperities [32]. Behaving as a Media Aware Network Element (MANE) [33] between media client and media server, the RTSP proxy intercepts RTSP messages from media client and forwards back RTP traffic. Fig. 6 describes the message redirection flow. The proxy is able to adapt video in real-time according to the configuring parameters that in fact come as the outputs of ADTE. The target configuration of video that can be generated, including bit rate, frame size and frame rate (the frame rate must be a fraction of the original frame rate, 1/2, 1/4, 1/8, etc). Currently the RTSP proxy supports the re-encoding of H.264 video, but it could be modified to support formats, such as MPEG-2 or MPEG-4 video. The RTSP proxy is a light-weight solution for video adaptation, because it introduces few impacts on the current media plane of IMS. To be noted, the configuring parameters generated by ADTE are firstly sent to AS hosting the service by MSM and then set to the proxy by the AS.

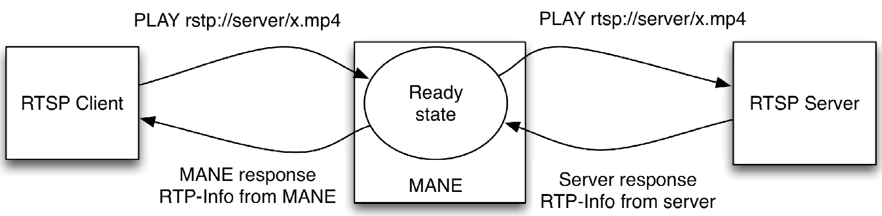

Fig. 6 RTSP/RTP message redirection

4.2.2 Session Negotiation Extensions

The scalability we describe in this paper relies on the information exchange between client and server, which includes both static capabilities (e.g. terminal or network) and dynamic conditions (e.g. network or user preference). It allows participants to inform each other and negotiate about the QoS characteristics of the media components prior to session establishment. SIP together with SDP is used in IMS as the multimedia session negotiation protocol. However, the ability is very limited for SDP to indicate user environment information such as terminal capabilities and network characteristics. The User Agent Profile (UAProf) [34] is commonly used to specify user terminal and access network constraints. It is also not enough, because UAProf contains only static capabilities. Although RFC 3840 [35] specifies mechanisms by which a SIP user agent can convey its capabilities and characteristics to other user agents, it is not compatible with MPEG-21 based ADTE. It is important to reach interoperability between IETF approaches for multimedia session management and the MPEG-21 efforts for metadata-driven adaptation, in order to enable personalized multimedia delivery [36]. In our model, UCD and UED serve as the input of ADTE, as shown in Fig. 5. These input values are in the format of XML document with a known schema. UCD includes the constraints imposed by service providers. We can assume that UCD is available for ADTE. However, UED should be collected for dynamic multimedia session in real time since it is the constraint imposed by external user environment. Therefore, there should be a way to query and monitor UED, particularly terminal capabilities and network characteristics.

In order to collect UED, we propose to extend the Offer/Answer mechanism. According to [30], SDP negotiation may occur in two ways, which are referred to as "Offer/Answer" and "Offer/Counter-Offer/Answer". In the first way, the offerer offers an SDP, the answerer is only allowed to reject or restrict the offer. In the latter way, the answer makes a "Counter-Offer" with additional elements or capabilities not listed in the original SDP offer. We slightly modify the latter way to put querying information in the "Counter-Offer". DIA defines a list of normative semantic references by means of a classification scheme [27], which includes normative terms for the network bandwidth, the horizontal and vertical resolution of a display, and so on. For instance, the termID "6.6.5.3" describes the average available bandwidth in Network Condition. Table I show some examples of the semantic references. To indicate these normative terms in SDP, we define a new attribute/value pair as shown in Table II. "Offer" and "Answer" are distinguished by "recvonly" and "sendonly" respectively. The value in "Offer" means the threshold set by offerer, which is optional. The value in "Answer" is mandatory as return. In the adaptation framework, MSM extracts the semantic inputs of AQoS and format them into SDP formats. During the Offer/Answer session negotiation procedure, the requested parameters are sent to UE in SDP. We assume that there is a module in User Equipment (UE) responsible for providing answers and monitoring dynamic conditions if necessary (e.g. presented by [20]). Accordingly, the answering values are also conveyed in SDP sending back to MAM activating adaptation.

Table 1 Examples of Semantic termID in DIA

termID	Semantic References
6.5.9.1	The horizontal resolution of Display Capability
6.5.9.2	The vertical resolution of Display Capability
6.6.4.1	The max capacity of Network Capability
6.6.4.2	The minimum guaranteed bandwidth of Network Capability
6.6.5.3	The average available bandwidth in Network Condition

Table 2 SDP Extension

Method	Syntax
Offer (for query)	q= * (termID) a= * (recvonly:<value>)
Answer (as reply)	q= * (termID) a= * (sendonly:<value>)

The proposed adaptation process is divided into three phrases: session initiation, session monitoring and session adaptation. In the session initiation phrase, the party who invokes the service offers the default parameters in SDP by a SIP signaling message, normally SIP INVITE. Besides those well known parameters as answer, MSM extracts input parameters in AQoS and offers them again as request. Some input parameters can be answered immediately such as terminal capabilities and network capabilities, which is enough for modality selection. However, some of them need to be monitored in real-time, for example network conditions. In case that any parameter varies out of the threshold set by AQoS, a SIP UPDATE with the specific SDP is feedback to MSM. Once ADTE in MSM receives the inputs and makes a decision, the adaptation starts with session renegotiation. In case of quality adaptation, MSM commands the RTSP proxy with the new parameters.

5 Prototype System and Performance Evaluation

5.1 Prototype System

This section takes the “Sports Live Broadcasting” service as an example to show the proposed approach in detail. Fig. 6 shows the two-phrase adaptation of “Sports Live Broadcasting” service. For the reason of simplicity, we omit some session messages and media plane data. The successfully registered UE initiates the service by sending a SIP INVITE request with a normal SDP. MSM parses the offered SDP in order to check whether the requested inputs exist. If not, MSM replies SIP 183 Session Progress with the SDP extension for querying capabilities. If the querying values can be retrieved immediately such as those static parameters (e.g. terminal capabilities or network capabilities), UE replies SIP PRACK with SDP as response. During pre-session adaptation, MSM makes an adaptation decision-taking based on the static device capabilities (e.g. width and height of display screen) and network capabilities (e.g. maxCapacity and minGuaranteed). The outputs include the "camElementMetadataID" indicating an appropriate service element for modality adaptation as well as the QoS parameters (e.g. bit rate and frame size) for quality adaptation. For example, in case of wide band connection MSM routes the request to the IPTV AS. Then the normal IPTV service is invoked. It is possible that the querying parameters are based on dynamic information in real-time, such as available bandwidth. In that case, UE sends back the response by SIP UPDATE with the payload of SDP. The in-session adaptation starts when MSM receives SIP UPDATE containing the requested available bandwidth. In case of quality adaptation, MSM sends SIP UPDATE with configuration parameters to IPTV AS that set then the RTSP proxy. If it has to make a modality adaptation, MSM responds SIP 503 Service Unavailable together with the header "Retry-After". The retry request will be routed to Text AS and invoke the text-based statistic service as the modality adaptation of IPTV.

To verify the proposed approach, we develop a prototype system that is the integration of several open source projects. The prototype system is illustrated in Fig. 7. On the server side, Open IMS Core [37] is deployed as IMS testbed. We make use of UCT Advanced IPTV [38] to provide IPTV service. MSM and Text AS is set up by Mobicents SIP Servlet [39] and configured to connect with Open IMS Core. In order to simplify the development, the client is simulated in the signaling plane and in the media plane separately. We emulate IMS signaling client by SIPp [40]. In the media plane, we have three types of terminal: Apple iPhone, HP iPAQ and Toshiba laptop.

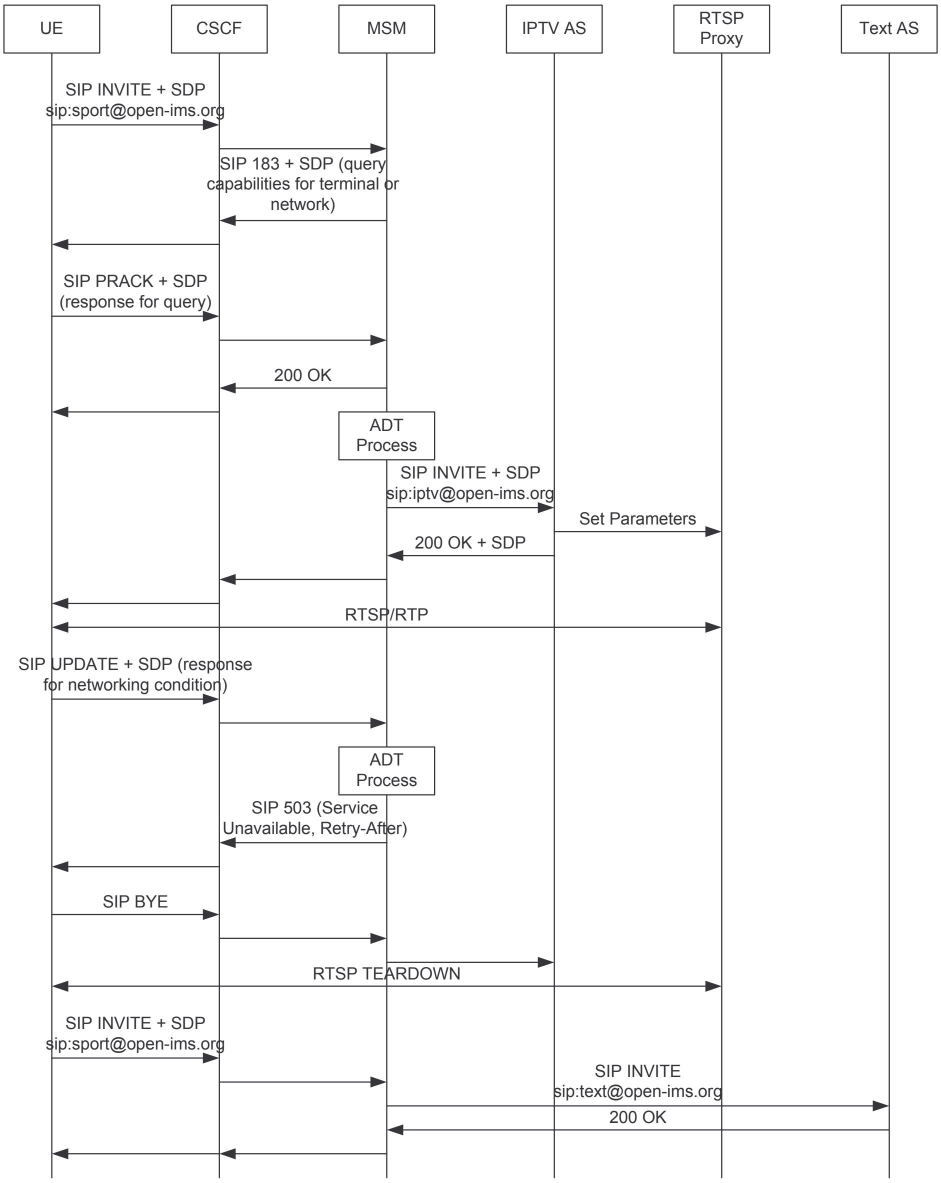

Fig. 6 Multimedia adaptation working flow

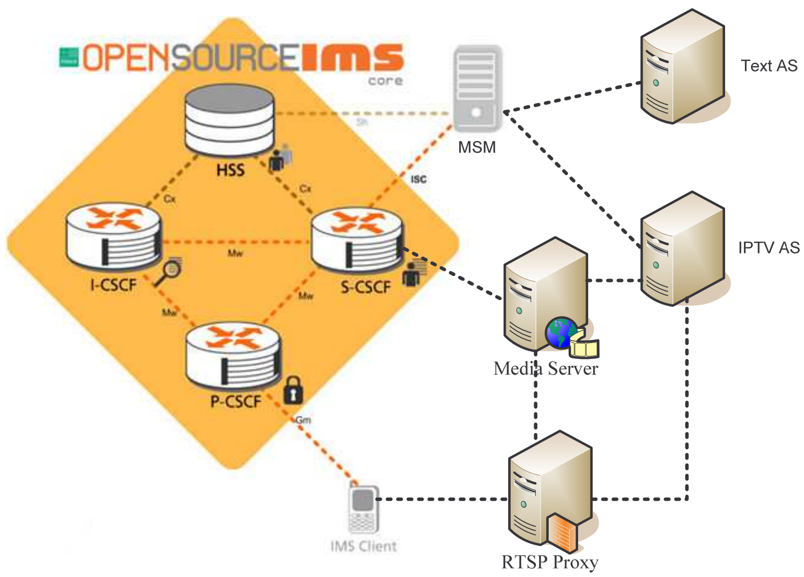

Fig. 7 Prototype system

5.2 Performance Evaluation

The prototype system demonstrates that the proposed SIP/SDP extension works compatibly with the standardized IMS platform. Regarding the performance, we analyze it in the signaling plane and in the media plane respectively. In the signaling plane, we observe that there are notably two kinds of latency: UED collecting and ADTE. The first one is more related to the characteristics of UED themselves. For instance, if the screen size is considered in UED, it could be retrieved immediately by UE. But in terms of available bandwidth, it depends on the time for sampling. Without considering UED, we further observe that ADTE-incurred delay is 0.1s in average. To some extent, this result confirms that the metadata-based adaptation is efficient, because the pre-computation saves significant time over parameter selection. The media plane is related to quality adaptation. We run tests with three terminals in two types of access network, as listed by Table 3. The test sample is a promotion video file for CAM4Home project with the bitrate of 3.9Mbps and the frame size of 720x576. As shown in Fig. 8, the output bitrates of adapted videos are basically consistent with the settings. Fig. 9 presents the output Peak Signal to Noise Ratio (PSNR) curves of adapted videos within 180 seconds. It can be seen that the adapted videos have different qualities, measured by means of PSNR. And the fluctuation of bitrate affects the PSNR in real-time.

Table 3 Terminal, access network and settings

Terminal	Access Network	Bitrate Setting	Frame Size Setting
Apple iPhone	EDGE	150Kbits	320x480
Apple iPhone	UMTS	1.5Mbits	320x480
HP iPAQ	UMTS	1.5Mbits	489x640
Toshiba Laptop (CPU: Intel Core 2 Duo P8700 @ 2.53GHz, RAM: 4GBytes)	Ethernet	4Mbits	720x576

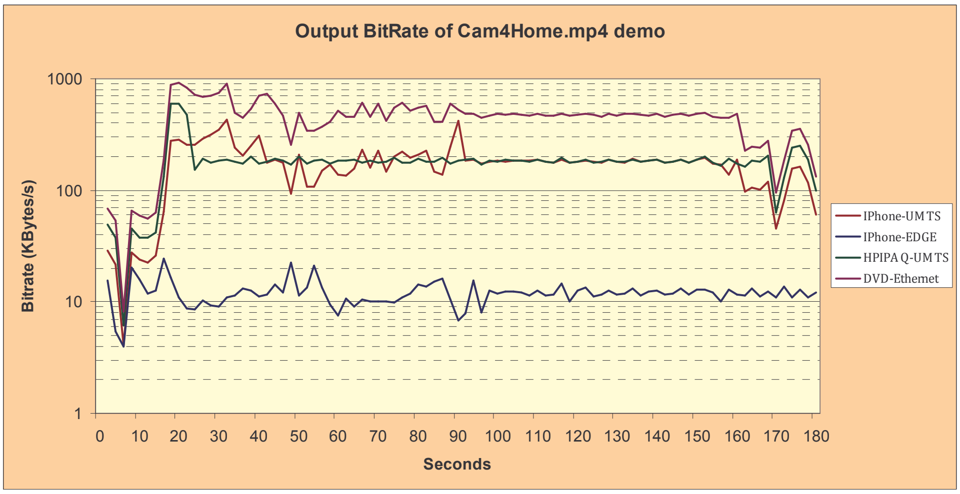

Fig. 8 Output bitrate of adapted video

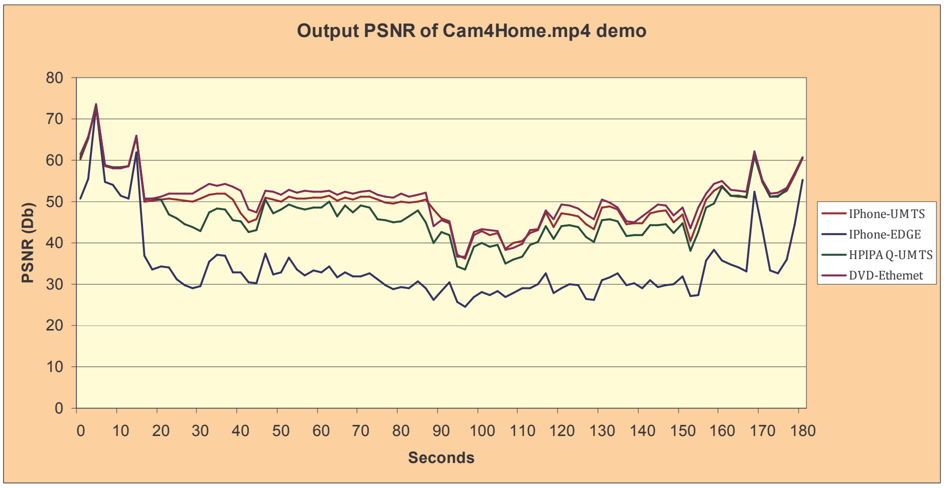

Fig. 9 Output PSNR of adapted video

6 Conclusions

This paper presents the SMD system in pervasive computing environment. It not only provides a novel way of multimedia provisioning but also assures the quality of multimedia delivery. To achieve that, we propose a flexible framework that introduces the CAM4Home metadata model as a bundle of multimodal media. MPEG-21 DIA is further integrated into CAM4Home model to meet end-to-end QoS requirements. We address the issues in supporting QoS from two aspects, namely fidelity and modality, in order to tailor and adapt multimedia to the diverse terminals and the heterogeneous networks, as well as dynamic network conditions. A prototype system is developed on IMS architecture to validate the proposed model. With the use of rich metadata, context awareness and personalization could be challenging topics in the future.

Acknowledgements

This work was supported by the ITEA-2 CAM4Home project. We would like to thank all partners for their contributions. And thank Hui Wang and Mengke Hu for their simulation work.

References

[1] C. Timmerer and H. Hellwagner, "Interoperable Adaptive Multimedia Communication", IEEE Multimedia Magazine, vol. 12, no. 1, 2005.

[2] Vertro A., Timmerer C., "Digital Item Adaptation: Overview of Standardization and Research Activities", IEEE Transactions on Multimedia, vol. 7, no. 3, pp. 418-426, Jun. 2005.

[3] Williams M.H., Yuping Yang, Taylor N., McBurney S., Papadopoulou E., Mahon F. Crotty M., "Personalized Dynamic Composition of Services and Resources in a Wireless Pervasive Computing Environment", The First International Symposium on Wireless Pervasive Computing, Jan. 2006.

[4] Xiaohui Gu, Nahrstedt K., "Distributed Multimedia Service Composition with Statistical QoS Assurances", IEEE Transactions on Multimedia, vol. 8, no. 1, pp. 141-151, Feb. 2006.

[5] Kalaspur S., Kumar M. Shirazi B., "Personalized Service Composition for Ubiquitous Multimedia Delivery", The Sixth IEEE International Symposium on World of Wireless Mobile and Multimedia Networks, Jun. 2005.

[6] T.C. Thang, Y.J. Jung, Y.M. Ro, "Modality Conversion for QoS Management in Universal Multimedia Access", IEE Proceeding Vision, Image Signal Process, vol. 152, no. 3, pp. 374-384, 2005.

[7] T.C. Thang, Y.J. Jung, Y.M. Ro, Semantic Quality for Content-aware Video Adaptation, Proceedings of the IEEE MMSP2005, Shanghai, 2005.

[8] T.C. Thang, J.G. Kim, J.W. Kang, J.J. Yoo, "SVC Adaptation: Standard Tools and Supporting Methods", EURASIP Signal Process: Image Communication, vol. 24, pp. 214-228, 2009.

[9] John R. Smith, Rakesh Mohan, Chung-Sheng Li, "Scalable Multimedia Delivery for Pervasive Computing", The Seventh ACM International Conference on Multimedia, 1999.

[10] T. Grgic, V. Huskic, M. Matija, (2008) Resource authorization in IMS with known multimedia service adaptation capabilities, Studies in Computational Intelligence, Springer Berlin, 10.1007/978-3-540-68127-4.

[11] CAM4Home Official Website, http://www.cam4home-itea.org/. Accessed 26th Nov. 2009.

[12] Shaha N., Desai A., Parashar M., "Multimedia Content Adaptation for QoS Management over Heterogeneous Networks", International Conference on Internet Computing, June 2001.

[13] T. Ahmed, A. Asgari, A. Mehaoua, E. Borcoci, L. B. Equille, and K. Georgios, "End-to-end Quality of Service Provisioning through an Integrated Management System for Multimedia Content Delivery", Computer Communications, vol. 30 (3), 2007, 638-651.

[14] A. Szwabe, A. Schorr, F.J. Hauck, A.J. Kassler, "Dynamic Multimedia Stream Adaptation and Rate Control for Heterogeneous Networks", Journal of Zhejiang University ACIENCE A, vol. 7, pp. 63-69, January 2006.

[15] D. Mukherjee, E. Delfosse, J.G. Kim, Y. Wang, "Optimal Adaptation Decision-Taking for Terminal and Network Quality-of-Service", IEEE Transactions on Multimedia, vol. 7, no. 3, pp. 454-462, June 2005.

[16] Information Technology – Multimedia Framework (MPEG-21) – Part 1: Vision, Technologies, and Strategy, 2002.

[17] M.W. Tong, Z.K. Yang, Q.T. Liu, "A Novel Model of Adaptation Decision-taking Engine in Multimedia Adaptation", Journal of Network and Computer Applications, vol. 33, no. 1 pp. 43-49, January 2010.

[18] IP Multimedia Subsystem (IMS), Stage 2, v7.6.0, TS 23.228, 3GPP, December 2006.

[19] L. Skorin-Kapov, M. Mosmondor, O. Dobrijevic, M. Matijasevic, "Application-Level QoS Negotiation and Signaling for Advanced Multimedia Services in the IMS", IEEE Communication Magazine, vol. 45, no. 7, pp. 108-116, July 2007.

[20] T. Ozcelebi, I. Radovanovic, M. Chaudron, "Enhancing End-to-End QoS for Multimedia Streaming in IMS-Based Networks", The Second International conference on System and Networks Communications, Augest 2007.

[21] J. Arnaud, D. Negru, M. Sidibe, J. Pauty, H.G. Koumaras, "Adapted IPTV Service within Novel IMS Architecture", The Fifth International ICST Mobile Multimedia Communications Conference, 2009.

[22] B. El Saghir, N. Crespi, "A New Framework for Indicating Terminal Capabilities in the IP Multimedia Subsystem", IEEE Global Telecommunications Conference, 2006.

[23] Z. Sahinoglu, A. Vetro, "Mobility Characteristics for Multimedia Service Adaptation", Journal of Signal Processing: Image Communication, vol. 18, no. 8, pp. 699-719, September 2009.

[24] G. Durand, G. Kazai, M. Lalmas, U. Rauschenbach, P. Wolf, "A Metadata Model Supporting Scalable Interactive TV Services", The Eleventh International Multimedia Modeling Conference, January 2005.

[25] I. S. Burnett, F. Pereira, R. V. de Walle, R. Koenen, "The MPEG-21 Book", pp.308, John Wiley & Sons, Ltd. 2005.

[26] S. Devillers, C. Timmerer, J. Heuer, H. Hellwagner, "Bitstream Syntax Description-based Adaptation in Streaming and Constrained Environments", IEEE Transaction on Multimedia, vol. 7, no. 3, pp. 463-470, 2005.

[27] A. Vetro, Ch. Timmerer, S. Devillers, "Information Technology – Multimedia Framework – Part 7: Digital Item Adaptation", ISO/IEC JTC 1/SC 29/WG11/N5933, October 2003.

[28] M. Prangl, H. Hellwagner, T. Szkaliczki, "Fast Adaptation Decision Taking for Cross-modal Multimedia Content Adaptation", IEEE International Conference on Multimedia & Expo, 2006.

[29] IETF RFC 2327, "SDP: Session Description Protocol", April 1998.

[30] IETF RFC 3264, "An Offer/Answer Model with Session Description," June 2002.

[31] Hongguang Zhang, Hang Nguyen, Noël Crespi, Shanmugalingam Sivasothy, Tien Anh Le, Hui Wang, "A Novel Metadata-based Approach for Content and Service Combined Delivery over IMS", submitted.

[32] IETF RFC 2326, "Real Time Streaming Protocol (RTSP)", April 1998.

[33] IETF RFC 3984, "RTP Payload Format for H.264 Video", February 2005.

[34] OMA-UAPROF, "User Agent Profiling Specification (UAPROF) 1.1", Open Mobile Alliance, 12 December 2002.

[35] IETF RFC 3840, "Indicating User Agent Capabilities in the Session Initiation Protocol (SIP)", August 2004.

[36] A. Kassler, T. Guenkova-Luy, A. Schorr, H. Schmidt, F. Hauck, I. Wolf, "Network-Based Content Adaptation of Streaming Media using MPEG-21 DIA and SDPng", The Seventh International Workshop on Image Analysis for Multimedia Interactive Services, 2006.

[37] T. Magedanz, D. Witaszek, K. Knuettel, "The IMS Playground@FOKUS an Open Testbed for Next Generation Network Multimedia Services", First International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, 2005.

[38] UCT Advanced IPTV, http://uctimsclient.berlios.de/. Accessed 12 September 2009.

[39] Mobicents, http://www.mobicents.org/. Accessed 18 September 2009.

[40] SIPp, http://sipp.sourceforge.net. Accessed 18 September 2009.