



SEVENTH FRAMEWORK PROGRAMME
ICT-1-1.5
Networked Media

Specific Targeted Research Project

My eDirector 2012
(FP7-215248)

My eDirector 2012 - Real-Time Context-Aware and Personalized Media Streaming Environments for Large Scale Broadcasting Applications

[D4.1 User and Network Context Acquisition and Tailored Streaming Creation]

Due date of deliverable: [31-08-2010]

Actual submission date: [23-07-2010]

Start date of project: 04-02-2008

Duration: 36 months

Summary of the document

Code:	D4.1 User and Network Context Acquisition and Tailored Streaming Creation – v1.1
Last modification:	05/08/2010
State:	Proof read
Participant Partner(s):	QMUL, AIT
Editor & Authors (alphabetically):	Editor: Stefan Poslad(QMUL) Authors: Uko Asangansi, Alan Pearmain, Stefan Poslad(QMUL), Aristodemos Pnevmatikakis (AIT)
Fragment:	Yes/ No ,
Audience:	<input checked="" type="checkbox"/> public <input type="checkbox"/> restricted <input type="checkbox"/> internal
Abstract:	<i>This document is deliverable D4.1 “User and Network Context Acquisition and Tailored Streaming Creation” and presents present techniques for taking into consideration user, content and network context information which can assist the formulation of the final stream for transmission.</i>
Keywords:	<ul style="list-style-type: none"> • Network Context • User Context • Context Awareness
References:	Refer to the corresponding section at the end of the deliverable

Document Control Page

Version number	V1.1
Date	05/08/2010
Modified by	Alan Pearmain
Comments	Proof read
Status	<input type="checkbox"/> draft <input checked="" type="checkbox"/> WP leader accepted <input type="checkbox"/> Technical coordinator accepted <input type="checkbox"/> Project coordinator accepted
Action requested	<input type="checkbox"/> to be revised by partners involved in the preparation of the deliverable <input type="checkbox"/> for approval of the WP leader <input checked="" type="checkbox"/> for approval of the technical coordinator <input type="checkbox"/> for approval of the project coordinator Deadline for action: 13/08/2010

Change history

Version number	Date	Changed by	Changes made
0.1	28/05/2010	Asangansi Uko	1 st version of TOC
0.2	14/06/2010	Asangansi Uko	Updates on Section 1 and 3
0.3	16/06/2010	Stefan Poslad	Updated scope to include user context, added background, outline and restructured document and added sections overall.
0.4	02/07/2010	Asangansi Uko	Updated background, motivation, and supplied previously omitted references.
0.5	02/07/2010	Stefan Poslad	Updated context-awareness and most sections.
0.6	13/07/2010	Asangansi Uko	Updated Use-cases
0.9	02/08/2010	Stefan Poslad	Major rewrite of most of document to make the final draft version for consortium comments
1.0	04/08/2010	Aristodemos Pnevmatikakis	Incorporated comments from partners
1.1	05/08/2010	Alan Pearmain	Conclusions amendment by Stefan Poslad added. Proof read.

Table of Contents

1	Executive Summary	6
1.1	Scope	6
1.2	Audience	6
1.3	Summary.....	6
1.4	Structure.....	6
2	Background and Motivation.....	7
2.1	The concept of Context in the My eDirector 2012 system	7
2.2	Personalised Video and Text Streaming.....	8
2.3	Network-aware Video and Text Streaming	8
2.3.1	Streaming Adapting to changing Access Network Conditions User Scenario	10
2.3.2	High-resolution Streaming On-demand User Scenario	10
2.3.3	Higher resolution of Smaller Views User Scenario	10
2.4	Terminal-aware Video and Text Streaming	11
2.4.1	Terminal Specific Streaming User Scenario.....	11
2.4.2	Seamless Roaming User Scenario.....	11
2.5	Location-aware Streaming	11
2.5.1	Location-Aware Streaming Use case	11
2.6	Light-level aware Streaming	11
3	Context Representation and Acquisition	13
3.1	Context Representation	13
3.2	ICT Context Environment Acquisition	14
3.2.1	Network Context Acquisition – Network QoS Conditions	15
3.2.2	Network Context Acquisition – Wireless Networks.....	15
3.2.3	Network Context Acquisition from User Terminal.....	15
3.2.4	Terminal Context Acquisition	15
3.3	Physical Environment Context Acquisition.....	15
3.3.1	Location Context Acquisition	15
3.3.2	Light level Context Acquisition.....	16
4	Context Adaptation Design.....	17
4.1	User Context Adaptation	17
4.2	Network-Context Adaptation	17
5	Conclusions	19
6	References	20

1 Executive Summary

1.1 Scope

This document constitutes the My eDirector 2012 project Deliverable D4.1. It presents the techniques for acquiring and utilizing the user, and network context, and for adapting the video streaming to the user. In this deliverable, concepts and processes specified in detail in other deliverables are not repeated but just referred to. Furthermore, it is worth noting that the underlying network topology upon which a user accesses the My eDirector 2012 platform is abstracted by the HTTP Adaptive Streaming scheme applied in the My eDirector 2012 project. Hence, the need for network topology information is nullified. The issue of acquiring the end-user context is dealt with in WP5 deliverables D5.1 and D5.2 and is not described in detail in this report.

1.2 Audience

This deliverable is public.

1.3 Summary

Factoring in network context information in a networked environment is important in ensuring that end users enjoy the anticipated Quality of Experience envisaged by the content provider. This deliverable presents the approaches taken to acquire and subsequently utilize the user and network context to adapt video streaming by the My eDirector 2012 system.

1.4 Structure

The overall structure of this document is as follows: Section 2 presents the concept of 'context' as applied in the My eDirector 2012 and the motivation for applying contextual information in the My eDirector 2012 platform. Context acquisition and representation approaches are discussed in Section 3. Section 4 presents the My eDirector 2012 context adaptation design and discusses further how the network context is being applied. Finally, in Section 5 the conclusions are drawn.

2 Background and Motivation

2.1 The concept of Context in the My eDirector 2012 system

There have been numerous attempts at the definition of the concept of context in literature. In a seminal publication, Schilit and Theimer [10] define context as the location, identities of nearby people and objects, and changes to those objects. In a similar definition, Brown et al. [2] define context as location, identities of the people around the user, the time of day, season and temperature, while Ryan et al. [9] referred to context as the user's location, environment, identity and time. In the same vein, Dey [4] defines context as the user's emotional state, focus of attention, location and orientation, date and time, as well as objects and people in the user's environment. Hull et al. [7] describe context as the aspects of the current situation. These sorts of definitions are often too broad, but a more specific definition is presented by Brown [1]. Brown defines context to be the elements of the user's environment which the computer knows about. However, a more accurate definition is provided by Dey and Abowd [5]. These authors refer to context as "any information that can be used to characterize the situation of entities (i.e., whether a person, place or object) that are considered relevant to the interaction between a user and an application, including the user and the application themselves."

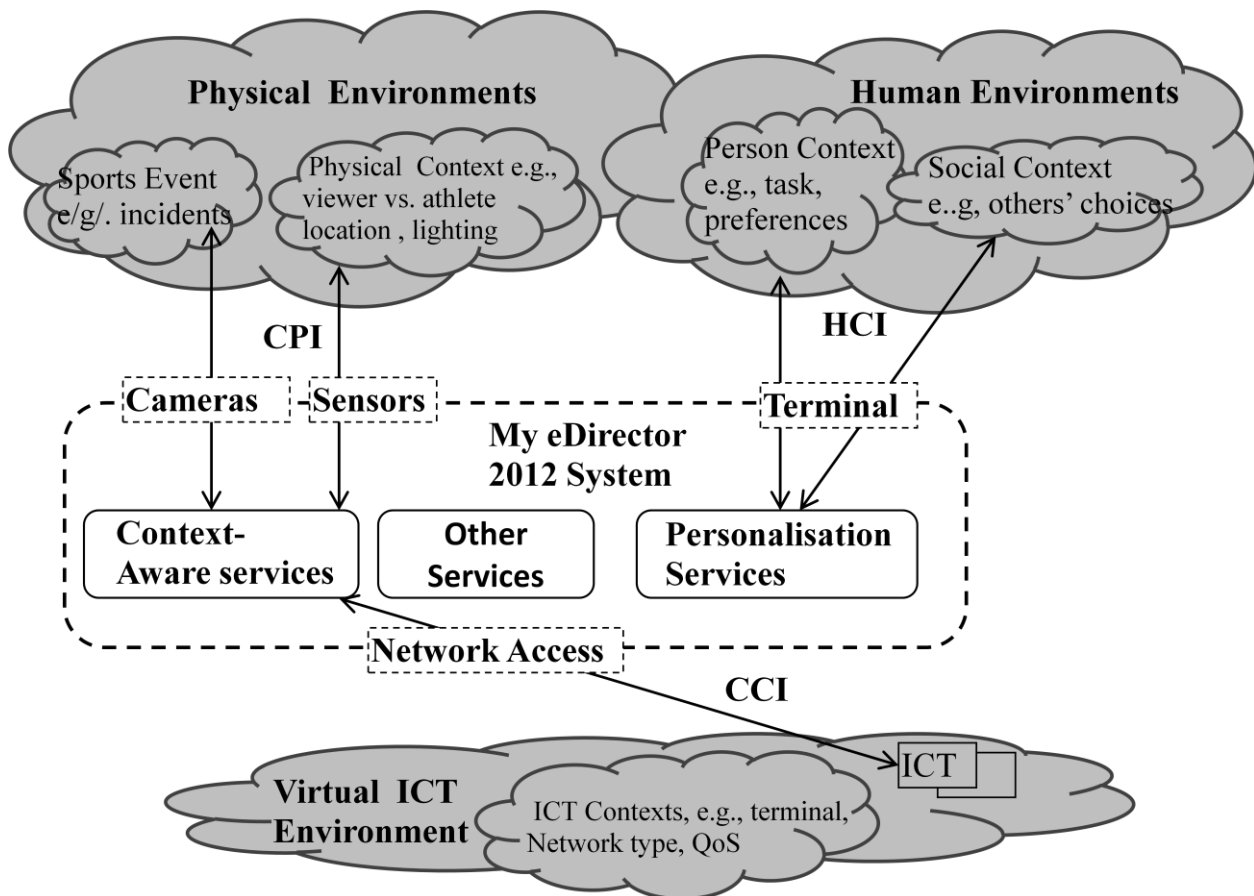


Figure 1: Overview of Context-awareness in the My eDirector 2012 system

Poslad [6] has reviewed in detail and discussed existing context definitions, representations and management life-cycle processes. He argues that there are two types of context model definitions, either quite abstract ones that act as a 'catch all' for much general system behaviour, e.g., Schilit and Theimer [10], Dey and Abowd [5], etc, or contexts that are defined via membership of a closed finite set of contexts, e.g., location, time, person etc. Instead, Poslad [6] considers context first of all in terms of one of three types of external environment context: physical, ICT and human environments (See Figure 1).

The most common choices for context-awareness are personalisation, network awareness and location-

awareness for mobile end-users. Personalisation involves designing a video streaming system that can be aware of a person's characteristics and preferences and then showing only choices of video streams that match these preferences that can be streamed. Network context-awareness involves being aware of the network characteristics such as throughput and delivering content so that it can be transmitted.

The motivation for context acquisition and awareness is considered with respect to the following types of context:

Human Environment Contexts

- End user context

Virtual ICT Environment Contexts

- Network Context
- Terminal Context

Physical Environment Contexts

- Location Context
- Light-level Context

Each of these is discussed in turn in more detail below.

2.2 Personalised Video and Text Streaming

End user context awareness most often refers to adapting services to the user preferences, this is also called personalising services. The motivation for this is to provide more content and service interaction that is of greater interest to the user thus enriching the user experience and to reduce the content less of interest to the user thus reducing the information and potential cognitive overload on the user. The user environment context is multi-dimensional. It covers users' demographic information (gender, nationality age etc). It covers their service specific preferences that can be explicitly defined by users at the terminal interfaces (preferences for sports etc). Many personalised viewing systems have in the past assumed that users' preferences are static over a user session and between sessions. Current state of the art systems are tending to observe users' usage history and to use this as a dynamic user profile to orientate services. The information gleaned from the user's context also makes it possible for the system to automate or simplify the execution of more complex tasks [8]. A key concern in personalising live video delivery is the issue of scaling video delivery to a large number of end users at a major event. The extreme would be multi-cast streaming to many millions of users with unique individual preferences. Rather than develop servers with the capacity to do this which would be prohibitively expensive to develop and operate, video is streamed with respect to groups of users and through creating user groups, matching individual users to those groups and making recommendations for stream choices to users based upon individual rather than group preference. End user context awareness including user group awareness as a form of personalisation is discussed in more detail in Work Package 5 (User Access and Personalisation) deliverables, see Deliverable D5.2 [27].

In addition to users' personal profiles, the end-user context overlaps with the user access device or terminal profile, to the network infrastructure and to the physical environment in which the user is situated in terms of the relative location of the user at a sports event (i.e., users' locations with respect to the location of the athletes and the different live camera locations and subsequent feeds) and the lighting.

2.3 Network-aware Video and Text Streaming

In conventional broadcast multimedia (video and text) streaming, the network context is not considered, as it may be assumed to be static while delivering multimedia content to the users, however this may lead to a variable network performance in terms of a variable Quality of Service (QoS) from the provider perspective and a variable Quality of Experience (QoE) as seen from an end-user perspective. Time-based video streams contain individual elements or frames that need to be generated and accessed at some fixed rate. If timing is not maintained or cannot be masked, e.g., because of congestion etc, distortions such as jitter occur. A video streamed at too high a rate for the bandwidth available results in choppy (normal speed inter-

persed with intermittent low speed sections) video and audio, while an unstable Internet connection can result in video stopping suddenly. The end user perceives a low quality streamed video.

One main factor that affects network performance is congestion in the core network and network routing in the face of this. Routing algorithms can be classified as deterministic or adaptive. In deterministic routing the path from the source to the destination is completely determined by the source and the destination addresses. Thus this cannot change routes because of congestion. In adaptive routing, multiple paths from the source to the destination are possible, thus providing packets with a better chance to avoid hot spots or faulty regions in the network as compared to deterministic routing. Hence, if network services are aware of the network topology changes, in terms of their throughput they could provide a better QoS. This was the original intention of the project in aiming to support network topology awareness.

A second major factor which affects the QoS is congestion in the access networks such as WLAN, WiMax, 3G and DVB-H by mobile users. In particular wireless access networks are very prone to this problem. This problem is made more complex when users have the flexibility to roam across different access points and to seamlessly switch between different types of wireless access network. One simplification is that if a network path consists of a mix of high-bandwidth wired core network and a lower bandwidth shared access network, the performance bottle neck is taken to be at the wireless access network caused by congestion and lower throughput. Simple measures of network context for overall network performance that have been used to determine the QoS across an access network for wireless services are to monitor the round trip time for network messages and error rates. 'Out of band' PING type messages are more suitable for non-streamed intermittent content access [29]. For streamed, e.g., video content, QoS signalling mechanisms, tend to be more tightly coupled with the corresponding transport protocol because separating them out would result in a very complex design. Typically, an existing implementation of the real-time streaming protocol (RTSP) transported over the real-time transport protocol (RTP) is used for video streaming. Either RTP can be used directly for flow control which is limited or it can be extended with a TCP friendly flow control algorithm to adjust the transmission rate to the network data rate by measuring round trip times and packet loss using pre-coded videos in MPEG-4 format [30]. The new generation of streaming services, based on adaptive HTTP streaming (i.e. IIS Smooth Streaming and Move Networks Adaptive Stream) rely on HTTP as the transport protocol and perform the media download as a long series of very small progressive downloads, allowing the media to get easily cached along the network, and much closer to the end users. Adaptive HTTP streaming can be considered as a new cross layer extension of TCP and HTTP protocols. The most important innovation of the adaptive HTTP streaming technology is the capability to provide multiple encoded bitrates of the same media content, therefore allowing the client media playback applications to seamlessly and dynamically switch between these bitrates depending on network condition (like current throughput, delay, packet loss and jitter). RTSP can identify congestion problems and request actions from the server in order to avoid such problems, but cannot respond as directly and fast as HTTP that uses the underlying TCP mechanisms for this, before passing to the second level of reaction that is the application layer response (reducing the media stream bitrate over the network). Hence adaptive HTTP rather than RTSP is used to support network context awareness in this project [31].

QoE is a subjective measure from the user's perspective of the value of services provided that is more than the effective quality of the service (QoS) seen from the provider perspective. Although, these are related and overlapped when researchers sometimes refer to user's QoS requirements [30], a lower QoS may still result in a high QoE and vice versa. The QoE considers contributions to this user experience not just performance but also, flexibility, mobility, personalisation and so on.

Vassiliou et al [32] have surveyed objective and subjective QoS Measures for Video Quality Assessment as part of their study of requirements for the transmission of streaming video in mobile wireless networks. Objective measurements seek to compare the decoded received video sequence as a whole or specific features to the original transmitted video. Objective measurements are difficult to perform for live video streaming. Further these may not characterise the end-user QoE of the decoded received video. Subjective measurements that seek to measure the perceived quality of a video are done through human "grading" of streams. They state that the most popular subjective measures of network QoE are Double Stimulus Impairment Scale (DSIS), Double Stimulus Continuous Quality Scale (DSCQS), Stimulus Comparison Adjectival Categorical Judgement (SCAJ) and Subjective Assessment Method for Video Quality evaluation (SAMVIQ). From these they chose SAMVIQ although clear justification for this choice is lacking. Although SAMVIQ has been developed to apply to offline retrieval of video clips, it can be applied to online streaming through providing some accessible user action control that can be used to express the level of satisfaction of

some specific user feature such as the perceived quality of the received video stream.

A challenge with using a single metric for user satisfaction for a non-trivial user task, e.g., say when a user has done a local digital personalised zoom of some streamed content and expresses his satisfaction, is that multiple factors may be involved and it may not be so clear what the user is expressing his satisfaction about. We may need to differentiate between the satisfaction being about the target selection, or the zooming or the adaptive steaming, or the increase in resolution, or the whole UI, or the personalisation, or the whole My eDirector system, or the ISP connection, or the core network QoS. If we can't differentiate these then the causal interpretation of the user satisfaction is unclear. The solution to this challenge will be to study this satisfaction under more limited viewing conditions and more controlled interaction studies.

For example, with respect to the project, end-users may attribute a high QoE for the following possible network-aware user scenarios:

1. Received video content remains smooth when the bandwidth fluctuates due to changing core network conditions.
2. Received video content remains smooth when the bandwidth fluctuates due to changing access network use conditions.
3. Higher resolution video can be triggered to be streamed on demand in response to a user task context change, e.g., zooming.
4. Narrow-field higher resolution versus wide-field low resolution can be requested
5. Seamless Roaming User Scenario.

Some of these scenarios are discussed in more detail below.

2.3.1 Streaming Adapting to changing Access Network Conditions User Scenario

Multiple users may share the same (xSDL) internet connection distributed via a Wi-Fi access point and the access link may get congested as users undertake different tasks: downloading emails (some with attachments), uploading pictures to a social networking site, downloading video episodes from a portal providing time-shifted content. Now, an additional user enters and signs into the My eDirector 2012 platform to watch the proceedings of an athletic event on the My eDirector 2012 player. Instead of being denied access due to a low bandwidth, or providing a high quality stream with sub-optimal experience by intermittently interrupting the live streams to buffer for a few seconds, the My eDirector 2012 player detects the network condition and delivers a the live stream(s) at a lower bit-rate and resolution. As the other family members complete their tasks and free up the internet connection, the My eDirector 2012 player will detect this change in network quality and increase the offered bit-rate and resolution respectively.

2.3.2 High-resolution Streaming On-demand User Scenario

The My eDirector 2012 system can provide 'selective' high bit rate audio visual content. This can be demonstrated while a user is consuming a given live stream of an event; the user could choose to zoom into an area of interest. Such action triggers a change in the user's context as well as the network context. This is because it serves as an implicit request for a higher resolution of the given stream for the selected zoom area. Although this places some demand on the network, the My eDirector 2012 system handles this demand by adapting to the network's context while serving the high resolution stream.

2.3.3 Higher resolution of Smaller Views User Scenario

The My eDirector 2012 system delivers content that adapts to the size of view available on a user's device. For example, if a mobile device is detected as the user's terminal, a mobile optimized stream is sent to service the user's requests. This mobile optimized stream is automatically generated by focusing in a region of interest within the each frame in the original stream. The way this region of interest is estimated is described in D4.2, Live Information Analysis [33]. Likewise, if a user originally connected to the My eDirector 2012 system with a mobile device via a Wi-Fi connection goes out of coverage or willingly connects via the mobile operator's mobile network, the My eDirector 2012 system will detect this change and automatically switch to a lower bit rate version of the given stream.

2.4 Terminal-aware Video and Text Streaming

Whereas yesterday's display screens vary by a factor of 2.5 in size and a factor of 4 in pixels, today's screens vary by factors of 100 in size and a factor of 625 in pixels, from a mobile phone to wall a screen (4000 x 1500 pixels). In addition, display devices are more heterogeneous even with the same network, e.g., there are different display resolutions for mobile phones, for TVs (HDTV versus SDTV) and for computer displays. Hence, video content may be accessed via a proliferation of interactive devices with diverse capabilities. These may range in size, weight and mobility. They may have different display capabilities such as screen resolution, size and colour-depth. They may use different forms of input, including different types of keyboard, pointer devices, speech and gesture input.

Lack of information of the terminal capabilities can reduce the QoE for the end user. For example, if the resolution of the video is too high for a display, it can't be fully displayed or will be incorrectly displayed or may not even be displayed. If low resolution video content is transmitted to a high resolution video terminal, the problem is less severe. Either the video is displayed only in part of the screen or the video is digitally zoomed at the terminal with features lacking the ability to take into account the resolution in the display.

2.4.1 Terminal Specific Streaming User Scenario

2.4.2 The core user scenario is that for each terminal, a terminal profile is created, see Section 3.2.4. Knowledge of the terminal characteristics can be used to select the appropriate video frame resolution and rate for streaming. Seamless Roaming User Scenario

Users may move their access terminal across different access networks roaming between Wireless LAN and WANs using WiFi, GSM and DVB networks. Users may switch seamlessly between displays. Users may switch viewing their current personalised stream choices from a mobile device to a DVB-T connected TV or to a LAN connected PC. To support these scenarios requires capturing some session context, closing an existing session with one type of display and access device and using existing session information about a user's current viewing context in a new session with a different terminal and access network.

2.5 Location-aware Streaming

Location-awareness is a common context-awareness for mobile users to access information services when they are roaming, commonly to know where they are currently located in relation to a destination. There is an additional potential specific use of location awareness for video streaming of live sports events for viewers located at the event in relation to the location of the athletes participating and in relationship to multiple camera views that may be available at the event. Location awareness of athletes can be used to enable viewers to automatically track specific athletes, possibly across multiple camera views, keeping them in view across specific sections of an athletic event or as much of the whole event that is possible. The motivation for the location awareness of the viewer in relation to the athletes and the camera views at the event is to enable viewers at an event to see additional complementary views of athletes that they cannot see physically from where they are located. For example, viewers in a sports stadium may find their physical view of event less satisfactory when the athletes move furthest away from them. The multiple camera views use of the location context is discussed in deliverable D4.4 [27] and is not discussed further here.

2.5.1 Location-Aware Streaming Use case

When a user is in the sports arena, the My eDirector 2012 system determines the location of the user's mobile device using a location determination method specified in Section 3.3. This is then mapped to a view of the sports stadium and event, containing the tracks of the athletes and the spatial shapes corresponding to the view from each camera. Either, a rule based model can be used to predict when an athlete is out of a viewer's physical view or tracking can be used or the user can manually signal this is the case. This information will then be used to facilitate a recommendation to the user for an alternative camera view to be streamed to a user device while the athletes are out of the viewer's physical view. The system could also recommend an entirely different sports event based on the user's preferences.

2.6 Light-level aware Streaming

If the ambient light is too bright and the brightness of a display is too low it can be difficult to distinguish features in streamed video. Alternatively, in a dark ambient environment, the screen can appear too bright,

brightness can be reduced saving energy. The brightness of more indoor display devices can be adjusted by controlling the brightness of the artificial lighting and by shading natural outdoor light or by manual control of the display brightness. In addition, the brightness can be determined by inbuilt sensors in a device, see Section 3.3.2.

3 Context Representation and Acquisition

Detection of context changes and the acquisition of context data can improve the extensibility and reusability of the system. There are different methods of context-data acquisition associated with different system architectural designs. Chen [14] differentiates three different approaches to how to acquire contextual information: direct sensor access, middleware infrastructure based and context-data server. Each of these is discussed in turn.

Direct sensor access: This approach is often used in devices that contain built in sensors. Client software gathers the desired information directly from these sensors, i.e., there is no additional layer for gaining and processing sensor data. Drivers for the sensors are statically bound to specific local device application, so this tightly coupled method is usable only in rare cases. Sensor data can be accessed by 3rd party services such as those on the My eDirector 2012 platform.

Middleware infrastructure: Object-oriented software design uses encapsulation to separate e.g., business logic and graphical user interfaces. A middleware based approach introduces a layered architecture to context-aware systems with the intention of hiding low-level sensing details. Compared to direct sensor access this technique eases extensibility since the client code has not been modified any further and it simplifies the reusability of hardware dependent sensing code due to the strict encapsulation;

Context-data server: This permits multiple clients to access remote data sources. This approach extends the middleware based architecture by introducing a remote access management component. Context-data is moved to a server to facilitate concurrent multiple access. Moreover, the use of a server has the advantage of relieving clients of resource intensive data operations.

In the My eDirector system, contexts that affect users tasks and interaction are acquired from three environments, the human, ICT and physical environments. A context-data server approach is used to store users' context data in a centralised Profile Builder (server) accessed via a web service. My eDirector refers to this as a network centric architecture design [35]. Currently, a middleware infrastructure approach, My eDirector refers to this as a terminal centric architecture design [35], is used to stored the terminal context locally and to support remote access.

3.1 Context Representation

Context information needs to be represented and modelled for being machine interpretable and exchangeable using well-defined interfaces to allow access by services external to the context acquisition sub-system. The context use requirements are to support easy manipulation, i.e. low overhead in keeping the model up-to-date, easy extension, i.e. cheap and simple mechanism for adding new types of information, efficient search and query access and scalability. Context representations may be tightly coupled to the inference mechanism, e.g. probabilistic logic requires the modelling of probabilities or tailored to the problem domain and to a specific goal of a system. Strang and Linnhoff-Popien [10] identified generic requirements for context modelling:

- be able to cope with high dynamics and distributed processing and composition,
- allow for partial validation independently of complex interrelationships,
- enable rich expressiveness and formalism for a shared understanding,
- indicate richness and quality of information,
- must not assume completeness and unambiguousness,
- be applicable to existing infrastructures and frameworks.

Context models can be classified into six different model categories, namely Key-value models, mark-up scheme models, graphical models, object oriented models, logic based models and ontology based models [69].

1. Key-value pairs form a simple tuple of information. The context information is assigned to a unique key in order to allow for easy lookup by applying a matching algorithm. These pairs are easy to manage but lack structuring and therefore do not allow for efficient context retrieval;
2. Markup scheme models incorporate a hierarchical data structure of mark-up tags, attributes and content. Well known examples are the User Agent Profile and the Composite Capabilities/Preference Profile (CC/PP) [15] being based on XML (Extendible Mark-up Language). For a rather lightweight modelling of contextual information, some projects and labs have developed proprietary context exchange formats like ContextML (Context Meta Language) [16][17];
3. Graphical models (e.g. based on the Unified Modelling Language (UML)) allow for a picturesque description of a context model [18] and for deriving an Entity Relationship (ER) model as required in relational databases. An extension is proposed by Henriksen et al. [13], introducing Object-Role Modelling (ORM);
4. Object oriented models offer powerful capabilities of inheritance, reusability and encapsulation. Access to contextual information is provided by well defined interfaces [19];
5. Logic based models offer a high degree of formalism and typically comprise facts, expressions and rules. They enable formal inference, e.g. by means of general probability logic, description logic, functional logic or first-order predicate logic;
6. Ontological modelling intends to capture an abstract conceptual vision of the world. The relations within could also be described by object oriented methods. Ontology is commonly described by using languages standardised by the World Wide Web Consortium (W3C) in the context of the semantic web. Most relevant are the Resource Description Framework Schema (RDF-S) and the Web Ontology Language (OWL) [21]. Available reasoning engines currently do not support complete interpretation of OWL Full [20]. Researchers in [23][21][24] concluded that ontologies are in principle the best way to represent and model context due to their extendibility and unambiguousness. However, with the size of the ontology, querying and processing the information embedded within becomes slow, in particular if performed at resource constraint mobile devices. Full featured ontological representations tend to decrease the inference performance and may not be suited for highly dynamic systems.

In the My eDirector system, the terminal and user context is represented using an XML based mark-up scheme such that the user's context in terms of the user's preferences is defined using an XML schema as most My eDirector platform services are Web services. In a real-time video streaming system, only simpler context data representations could be processed on the fly so this rules out the use of much richer context representations, except in limited cases in a more off-line mode (see the end of this section). Many contexts are based upon MPEG-21 Usage Environment Description (UED) descriptors, an XML extension. The UED is defined in Part 7 of MPEG-21, i.e., Digital Item Adaptation (DIA) [36]. The UED covers four categories of environment including: user, terminal, network and natural or physical environment. MPEG-21 terminal capabilities can be used to describe the device class; encoding and decoding capabilities; display capabilities; audio output capabilities; benchmark information; user interaction possibilities; power characteristics; IPMP tools; data input/output characteristics; and storage possibilities. The My eDirector use of the MPEG-21 terminal capabilities is described in [34]. MPEG-21 User characteristics describes user information; usage preferences; usage history; display, audio, and graphic presentation preferences; conversion preference; presentation priority preference; focus of attention; visual and auditory impairment; destination; and mobility characteristics. The My eDirector user profile is defined in [28]. My eDirector does not use MPEG- to describe the network characteristics as the main use of the network context is handled by the adaptive HTTP protocol itself. MPEG-21 can be used to describe the natural environment characteristics: lightning conditions, noise level, time, and location. However, richer logic-based and semantic context representations are used for specific types of context adaptation such as personalised zooming and camera selection, see D5.1.2 [34].

3.2 ICT Context Environment Acquisition

The state of a user's network connection is a factor that can significantly affect the level of Quality of Experience (QoE) the user perceives from the My eDirector 2012 system. Hence, the acquisition and subsequent

utilization of a user's network context is needed.

3.2.1 Network Context Acquisition – Network QoS Conditions

An approach for acquiring a user's network context might involve the extraction of detailed information about the network's QoS conditions. This approach might involve periodic measurement of network traffic parameters at specific nodes so as to determine the context. In carrying out these measurements, internal network elements such as routers will have to be probed actively with network packets and saving timestamps periodically as well as information about packet loss and delay [3]. Hence, this approach might be able to provide network context information related to a specific network path in terms of the network delay, jitter, loss rate and available bandwidth.

3.2.2 Network Context Acquisition – Wireless Networks

An approach for acquiring context (and capabilities) in wireless networks can involve the installation of a Wireless Access context provider which resides in a wireless network device [3] and it is used in association with other elements in acquiring information about the diverse landscape of heterogeneous radio access technologies in addition to selecting the appropriate network during handovers. Such context provider is able to provide both static and dynamic contextual information which consists of different parameters depending on the respective radio access network. The static context comprises information such as network type (e.g WLAN, UMTS, WIMAX, etc), Point of Attachment (PoA) ID (e.g SSID, call ID), the supported air interface (e.g. 802.11 a/b/g/n, UMTS/HSDPA/HSUPA/HSPA+, 802.16e/m) and their respective rates, and the network's multicast capabilities (e.g. UMTS/MBMS).

For WLANs, dynamic context can include a list of associated MAC addresses and respective RSSI (Received Signal Strength Indication) values, noise levels at access points and neighbour access points within reach. While in the case of UMTS, the dynamic context information will be the list of attached devices, their reported signal-to-noise ratio and neighbour base stations, and all these will have to be retrieved from the NodeBs and/or the RNCs(Radio Network Controller).

3.2.3 Network Context Acquisition from User Terminal

Apart from using the network elements to acquire the network context, the user's terminal can also provide vital information from which the network context can be acquired. A variety of metrics can be retrieved from the user terminal to deduce the network context. This could include the basic RAT (Radio Access Technology) related parameters from the terminal point of view such as RATs within reach, and their respective RSSIs. The user terminal can be said to possess both static and dynamic network context information. The static context information involves the MAC address, IMEI, IMSI, a list of RATs the terminal is able to support, whereas the dynamic context comprises parameters such as RSSI, call state, data connection activity or battery status.

This concept of utilizing the user terminal's perception of the network as a means to acquire the network context is applied in the My eDirector 2012 project. This approach presented the most plausible option due to the lack of access to (the internal elements of) the end user network infrastructure.

3.2.4 Terminal Context Acquisition

The main techniques for terminal profiling have already been defined in Deliverable D5.1.2 [34]. From these MPEG-21 was selected as the representation for the terminal context. Either this is created manually off-line and loaded into each user's terminal or the terminal type needs to be specified by the end-users when they login and a terminal profile is assigned by the My eDirector 2012 system within the network. Once the appropriate terminal profile is assigned to a user's terminal, the video stream suited to that terminal is selected for streaming.

3.3 Physical Environment Context Acquisition

3.3.1 Location Context Acquisition

Location and position have been widely investigated for their use in context-aware systems. In ascertaining an object's location, common sources of location information have included Global Positioning System

(GPS) and location determined using WiFi measurements and proximity to Bluetooth, Radio-Frequency Identification (RFID) readers, or IR transmitters at known locations. Location determination based upon GSM/CDMA cell IDs and, possibly dynamically assigned, IP addresses are too coarsely grained to determine the location of users as well as user input. These sources of location information all have performance limits, for instance GPS systems are for use in outdoor environments (depending on the number of satellites) while they could be virtually of no use in indoor environments or in partially roofed sports stadiums as they lack a line of sight to the GPS transmission satellites. To best of the authors' knowledge virtually no sports stadiums are instrumented heavily with Bluetooth, RFID or IR readers to be used to determine the position of viewers. Many modern stadiums which are used for high-level national and international events may have WiFi access.

Location can be determined from packet round trip travelling time (RTT) and Received Signal Strength Indication (RSSI) measurements. However, the RTT is highly time-variant due to the impact of multipath and clock quantization errors. This leaves RSSI which involves a specific technique called finger-printing. Finger prints are created and used in two phases: offline and online phase. The offline phase involves creating a radio map. The radio map will store distributions of RSSI values from all detected WiFi Access points (APs) at known locations. These AP locations together with the MAC address of each detected APs and their corresponding RSS values will be stored in a database to create the radio map consisting of a set of RSSI samples taken at specific location on a map called fingerprints. In the online phase, when the user will initialize the mobile application, the sniffer component will collect samples of RSS values from all detected APs. To estimate the position of the user, the data collected will be compared in the radio map. The issue is whether or not there are a sufficient number of APs available and if a WiFi radio map has been compiled. In addition, each of these techniques needs to handle some uncertainty in position determination [25].

If WiFi location determination is not available there remains one further technique that can be used to determine a viewer's position in a stadium. Users can manually enter their position either by entering their seat number mapped to a stadium plan or by marking their position on a view of a stadium.

3.3.2 Light level Context Acquisition

Some commercial consumer device displays are starting to have automatically configurable light brightness. For example the Android Nexus 1, HTC TouchDiamond and iPhone mobile phones and smart TVs such as the LCD Sony Bravia and Samsung 550. Embedded system devices such as smart TVs that run application specific operating systems tend to limit direct access to sensors to fixed applications, i.e., the sensor data is not accessible to remote applications. Mobile devices do define APIs to access sensors. In some cases, e.g. Windows Mobile, there is even a Unified Sensor API that allows developers generic access to the hardware sensors that are available on various phones such as Light Sensors, capacitive touch pad and accelerometer.

4 Context Adaptation Design

This section presents an architectural overview of the core My eDirector 2012 system components that provide the context awareness properties.

4.1 User Context Adaptation

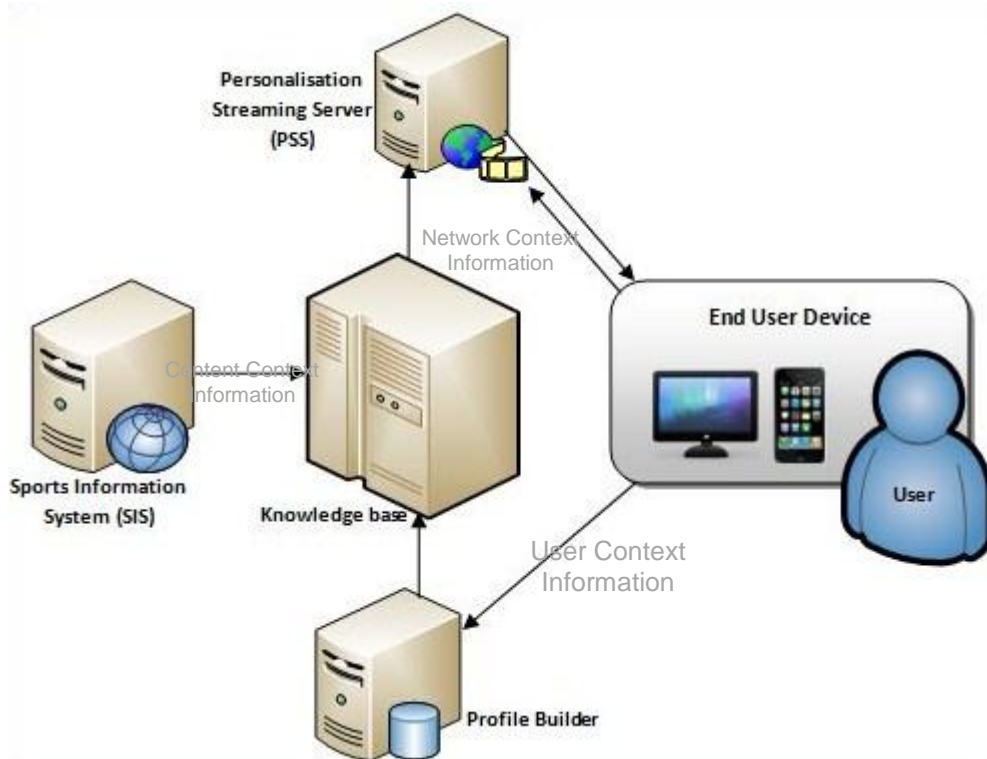


Figure 2 Overview of User Context adaptation in the My eDirector 2012 system

During system operation, the user context information is cached on the user terminal and then periodically transferred via a SOAP interface to the Profile Builder (Figure 2). The profile builder is responsible for clustering user contexts into user groups. The Sport Information System (SIS) provides some content metadata via a Java Messaging Service (JMS) interface to the Knowledge Base. The Knowledge Base then reasons over the user context information (received from the Profile Builder input) and the content metadata for the SIS to produce a context-based stream to each user according to their group. This reasoning technique is as specified in Deliverable D4.4[26]). The Personalized Streaming Server (PSS) receives the user context information from the knowledge base and streams the personalised views of events to the user terminal.

4.2 Network-Context Adaptation

The network context information retrievable from the user terminal in the My eDirector system include: the bandwidth, bit rate, location (using IP address of the terminal), and the download rate in fps (frame per second) from the client which is a reflection of both the network and terminal CPU context. While the IP address of the terminal can be said to be more of a static context at least during a user session, the others represent dynamic contexts.

The network context is a contributory factor to the overall QoE that is delivered from any networked multimedia platform. Hence the utilization of the network context in the delivery of content is imperative, but most elements of the network context are dynamic and fickle. For example, if a user connects to an ISP with a laptop at 56 Kbps, that does not mean that 56 Kbps of bandwidth is available at all times. Bandwidth can vary, meaning that a 56-Kbps connection may decrease or increase, causing video quality to fluctuate. User tasks

also influence the network context. For instance, a user could choose to watch the broadcast sports content in full screen instead of in a reduced screen size., In another case, the user could choose to zoom into a particular area of interest in the video content. Executing these tasks has a direct effect on the (dynamic) network context. These changing network contexts place a variable demand on the network resources and could adversely affect the user's QoE (Quality of Experience) if not adequately catered for.

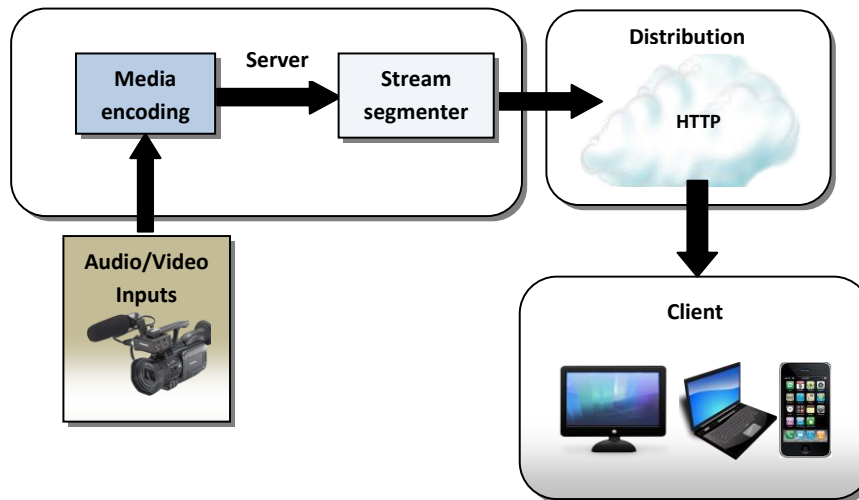


Figure 3: Overview of the HTTP Adaptive Streaming scheme

The My eDirector approach to handling the network context is based on the adoption of an HTTP Adaptive Streaming scheme (see Figure 3). HTTP Adaptive (video) streaming is a process that adjusts the quality of a video stream delivered to a terminal based on changing network contexts in order to ensure a better QoE for the user. In a nutshell, adaptive streaming adjusts the bit rate of the video content to adapt to changing network contexts. The HTTP Adaptive Streaming uses a simple principle which involves the delivery of small video content fragments - about two seconds worth of video - and verifying that each has arrived within the required time and played back on the client at the expected bit rate. If one fragment does not meet these requirements, the following fragment will be delivered at a lower bit rate. If the network conditions are more favourable, subsequent fragments will be sent at a higher bit rate.

For video content served using HTTP Adaptive Streaming, the content will have to be encoded at several bit rates, typically with each encoding level in its own complete file using a compression tool. The content is then delivered using an HTTP Adaptive Streaming-enabled server. Once the content server receives a request for media, it will dynamically create cacheable virtual fragments from the video files and deliver the best content possible to each end user. The benefit of this virtual fragment approach is that the content only needs to manage complete files rather than thousands of pre-segmented content files.

5 Conclusions

Context aware services have the potential to enhance the QoE for viewing streamed video content for end-users of the My eDirector 2012 system. The potential benefits to viewers of live video streams is that they offer more relevant and selective content to their interests and can react and handle environment changes that are not directly under management control of the system, e.g., a public Internet infrastructure. Two main types of context awareness are supported: by the system: end-user context awareness (personalisation) and network context awareness. End-user context awareness is discussed in detail in the WP5 deliverable D5.2 and in a future WP5 deliverable D5.3 (currently being drafted). The My eDirector 2012 system uses a network context awareness approach that is specific to the streaming protocol used based upon adaptive HTTP. Rather than just using an application agnostic adaptive video streaming approach, this provides the added benefit for My eDirector 2102 that viewers can enjoy user-task specific adaptive streaming, e.g., higher resolution videos can be triggered when viewers zoom in at specific events (providing the source content and bandwidth are available to leverage this). In addition, an awareness of the physical environment context in terms of the viewer location, in relation to athletes' and cameras' locations, and of the ambient lighting conditions can be supported by the My eDirector 2012 system to enhance the viewer experience. Streamed (e.g., far field) camera views could be offered to mobile devices of viewers at sports events to complement their (e.g., near field) natural view.

6 References

- [1] Brown, P.J. (1996). The stick-e document: a framework for creating context-aware applications", Proceedings of the Electronic Publishing, Palo Alto, pp.259–272.
- [2] Brown, Peter J., John D. Bovey and Xian Chen (1997). Context-aware applications: From the laboratory to the marketplace. IEEE Personal Communications 4(5): pp. 58-64. October 1997.
- [3] C-CAST Project Deliverable D 13 (2009) Specification of context detection and context-aware multi-party transport, June 2009
- [4] Dey, A.K. (1998) Context-aware computing: the CyberDesk project", Proceedings of the AAAI, Spring Symposium on Intelligent Environments, Menlo Park, CA, pp.51–54.
- [5] Dey, A.K. and Abowd, G.D. (2000) „Towards a better understanding of context and context-awareness", Proceedings of the Workshop on the What, Who, Where, When and How of Context-Awareness, ACM Press, New York.
- [6] Poslad, S. (2009) Ubiquitous Computing: Smart Devices, Environments and Interactions, Wiley, ISBN: 978-0-470-03560-3, 213-244.
- [7] Hull, R., Neaves, P. and Bedford-Roberts, J. (1997) „Towards situated computing", Proceedings of the First International Symposium on Wearable Computers (ISWC '97), pp.146.
- [8] Leiberman, H., et al., (2000) Out of Context: Systems That Adapt to, and Learn from, Context. IBM Systems Journal 39, 3&4, pp. 617-632.
- [9] Ryan, N., Pascoe, J. and Morse, D. (1997). Enhanced reality fieldwork: the context-aware archaeological assistant", Proceeding of the 25th Anniversary Computer Applications in Archaeology .
- [10] Schilit, Bill N. and Marvin M. Theimer (1994). Disseminating active map information to mobile hosts. IEEE Network 8(5): pp. 22-32. September/October 1994.
- [11] Strang, T. & Linnhoff-Popien, C. (2004). A Context Modeling Survey. In: Workshop on Advanced Context Modelling, Reasoning and Management, at UbiComp 2004, Sixth International Conference on Ubiquitous Computing, Nottingham/England.
- [12] Gonzalez A., Ahlers R. (1999) Context based representation of intelligent behavior in training simulations", Trans. of the Society for Computer Simulation International, Vol. 15, No. 4, p. 153-166.
- [13] Bucur, O., Beaune, P. and Boissier, O. (2005) Representing Context in an Agent Architecture for Context-Based Decision Making", proceedings of CRR'05 Workshop on Context Representation and Reasoning.
- [14] Henricksen, K. et al. (2005). Middleware for Distributed Context-Aware Systems. In On the Move to Meaningful Internet Systems 2005: CoopIS, DOA, and ODBASE. pp. 846-863. Available at: http://dx.doi.org/10.1007/11575771_53 [Accessed July 13, 2010]
- [15] Chen, H. (2004) An Intelligent Broker Architecture for Pervasive Context-Aware Systems, PhD Thesis, University of Maryland, Baltimore County.
- [16] Composite Capabilities/Preferences Profile Home Page, <http://www.w3.org/Mobile/CCPP>
- [17] Goix, L. et al., (2007) Situation Inference for Mobile Users: A Rule Based Approach. In Mobile Data Management, 2007 International Conference on. pp. 299-303.
- [18] Moltchanov, B. et al. (2008) Context-Aware Content Sharing and Casting. ICIN 2008 Proceedings, Bordeaux, France.
- [19] Sheng, Q. & Benatallah, B. (2005) ContextUML: a UML-based modeling language for model-driven development of context-aware Web services. In Mobile Business, 2005. ICMB 2005. International Conference on. pp. 206-212.
- [20] Hofer, T. et al. (2003) Context-awareness on mobile devices - the hydrogen approach. In System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on.

- [21] Eiter, T. et al. (2008) Rules and Ontologies for the Semantic Web. In Reasoning Web. pp. 1-53.
- [22] Ejigu, D., Scuturici, M. & Brunie, L. (2007) Semantic approach to context management and reasoning in ubiquitous context-aware systems. In Digital Information Management, 2007. ICDIM '07. 2nd International Conference on. pp. 500-505.
- [23] Pils, C., Roussaki, I. & Strimpakou, M. (2007) Distributed Spatial Database Management for Context Aware Computing Systems. In Mobile and Wireless Communications Summit, 2007. 16th IST. pp. 1-5.
- [24] Hyunjun Chang, Seokkyoo Shin & Changshin Chung (2007) Context Life Cycle Management Scheme in Ubiquitous Computing Environments. In Mobile Data Management, 2007 International Conference on. pp. 315-319.
- [25] Anagnostopoulos, C. & Hadjiefthymiades, S. (2008) Enhancing Situation-Aware Systems through Imprecise Reasoning. Mobile Computing, IEEE Transactions on, 7(10), 1153-1168.
- [26] HTML5 Geolocation API Specification - W3C Working Draft 07 July 2009 (<http://www.w3.org/TR/geolocation-API/>)
- [27] My eDirector 2012 (2009), "Configurable Camera Selection Metrics and Algorithms". Project Deliverable, D4.4.
- [28] My eDirector 2012 (2009), "Personal Preference Model Report". Project Deliverable, D5.2.
- [29] Poslad S, Laamanen H, Malaka R et al. (2001) CRUMPET: Creation of User-friendly Mobile services PErsonalised for Tourism. Proc. IEE 3G2001 Conf. Mobile Communication Technologies, pp 28-32.
- [30] S L. Amundsen, K. Lund, C. Griwodz, et al.(2005) QoS-aware Mobile Middleware for Video Streaming, Proceedings of the 2005 31st EUROMICRO Conference on Software Engineering and Advanced Applications (EUROMICRO-SEAA'05).
- [31] My eDirector 2012. Handling the media stream on the My eDirector 2012 platform (or RTSP vs. HTTP). Internal Project Technical Report. May 2009.
- [32] V. Vassiliou, P. Antoniou, I. Giannakou, et al. (2006) Requirements for the Transmission of Streaming Video in Mobile Wireless Networks Proc. ICANN 2006, Part II, LNCS 4132, pp. 528–537.
- [33] My eDirector 2012 (2010) Live Information Analysis. Project Deliverable, D4.2,
- [34] My eDirector 2012 (2009), User Terminal and User Task Interface Report Extension to D5.1. Project Deliverable, D5.1.2
- [35] My eDirector 2012 (2009), System Architecture and Specifications. Project Deliverable, D2.3
- [36] Bormans, J.; Gelissen, J.; Perkis, A. (2003) MPEG-21: The 21st century multimedia framework. IEEE Signal Processing, 20(2), 53 – 62.