AMBULANT

Phase I: A Multi-Profile SMIL Player for Mobile and Desktop Systems

Project Overview

Project leader: Dick C.A. Bulterman

Version 2.1W
15 April 2003

Table of Contents

1 Proposal Abstract

The past five years have seen an explosive growth in the installation and acceptance of multimedia technology on desktop computers. While the deployment of media presentation technology has been impressive, its use for other than the most simple applications has been disappointing. In spite of language facilities in SMIL-2 [1] and MPEG-4 [2] that allow for the construction of complex presentations in which media objects can be linked into a related web of information based on the capabilities of the underlying network infrastructure, the media experience for users has never grown beyond that provided on the lowest-common denominator component within the infrastructure.

The AMBULANT project has as its goal the study of the network and infrastructure aspects of multimedia delivery. The project is defined in three phases. Phase I will provide a stable, reliable and open multimedia player that can be used at CWI and elsewhere for research in networked multimedia; it will implement the main standards of the W3C without intellectual property restrictions and will be available under open source licensing to member of the research and development community. Phases II and III will consist of research into models and protocols for application-influenced network layer support of multiple end-to-end multimedia streams; this work will study issues in accessing, transforming and delivering presentations in a heterogeneous environment based on device, network infrastructure and end-to-end constraints. Note that Phases II and III fall outside the scope of this proposal.

Phase I will produce its deliverables in three releases: the first release (after three months) will provide a basic protocol engine, the second release (after approximately 9 months from project start) will provide support for a broad range of codec's, and the final release (at the conclusion of this phase of the project) will consist of an optimized player, testset and documentation.

The project will be developed by a team of researchers at CWI and elsewhere, each of whom has some experience in implementing (partial) SMIL engines. CWI will play a coordinating and integrating role in the project and deliver the bulk of the implementation support. The project has a duration of 12 months and the CWI component will require external funding of € 122K.

2   Project Description

This section discusses the project scope of Phase I of AMBULANT. In particular, it considers the project's background, proposed workplan, deliverables, results dissemination, importance and partnership opportunities.

2.1   Background

The past five years have seen a tremendous growth in the installation and acceptance of multimedia technology on desktop computers. Ubiquitous software such as the RealPlayer, InternetExplorer's HTML+TIME Player and QuickTime — using languages and algorithms that, in large part, were developed at CWI — have made the presentation of atomic bits of media such as a single video or a simple slideshow commonplace. With the emergence of video capabilities on mobile telephones, the migration of atomic media objects to the mobile world seems imminent.

While the deployment of media presentation technology has been impressive, its use for other than the most simple applications has been disappointing. In spite of language facilities in SMIL-2 [1] and MPEG-4 [2] that allow for the construction of complex presentations in which media objects can be linked into a related web of information based on the capabilities of the underlying network infrastructure, the media experience for users has never grown beyond that provided on the lowest-common denominator component within the infrastructure.

From a network perspective, this means that multimedia has had little impact on the network layer infrastructure. Some streaming control takes place end-to-end, but there is little or no use made of facilities for controlling the flow of information from multiple sources, and no attempt is made to integrate network-level quality of service control for presenting complex media. Since there is little use of the network layer facilities, there has been little incentive to further develop network support beyond that found in current (generally non-synchronized) packet delivery protocols.

The environment for supporting rich media presentations is caught in a Catch-22 situation: the key market players (from Microsoft and RealNetworks) support only a single media stream because there is no network level support for efficient multi-stream management — but there is no network level support

of multiple streams because the main market players only support a single stream.

The primary reason that the major market players have limited their scope to single-stream transmission is that there is a high degree of complexity involved in each of the component areas of defining, synchronizing and presenting rich media presentations, and that there is a high degree of interaction necessary among the solutions provided at the user, media player and runtime infrastructure that all need to be addressed in a consistent manner to provide a useful end-to-end media experience. Much research needs to be done in the processing of media streams dynamically within the network infrastructure [3], [5], [6], [7] but this research has had only limited practical impact since an open, common reference platform does not exist to support such research and development.

The AMBULANT project proposes to study various network and systems aspects of supporting rich media in a heterogeneous, networked environment. The project is expected to consist of at least the following three phases:

- *Phase I*: development of a common playback environment based on current standards for multimedia presentation; this environment can be used as a basis for experimental research in later phases of AMBULANT and by other researchers who desire an open, tailorable player.
- *Phase II*: development of models to study predictive support for multi-sourced media presentations though a dynamic network infrastructure for delivery to heterogeneous clients; and
- *Phase III*: development of end-to-end models to support predictive multi-sourced hypermedia, including the processing of links and alternative content based on network infrastructure and client dynamics

Note that the first phase of the AMBULANT project is the subject of this funding proposal. Only this phase will be addressed in detail in this proposal.

2.2   Proposed Work

CWI proposes to coordinate the development of a multi-profile, multi-platform SMIL-2.0 player that provides full support for the main W3C standards used in multimedia applications. The player will provide a complete implementation of the various profiles defined or used with the W3C SMIL language (including the major mobile and desktop in use.) The player will be architected to run on

the major platforms available to, and in use by, the research community. The software will be developed by a team of research institutes, each of whom has experience in developing (partial) implementations of the SMIL language or associated W3C standards. As far as we know, this will be the only player available anywhere to support these three profiles within a single unit.

CWI will assume the responsibilities for overall system architecture, a major component of system implementation and the final publishing and initial distribution of results. Other distribution channels will be sought within the mobile and desktop communities during the course of the project. All of the software and test suites provided by the project will be available under GNU open source licensing terms (or equivalent).

The purpose of this platform is to provide a common base for research into new network protocols and infrastructure support for multimedia. The project will be geared to providing a well-documented player implementation that can be extended or further instrumented by external parties.

2.3   Deliverables

We propose to develop and distribute the following deliverables:

- *target platforms*[1]:
    - — PDAs running WinCE
    - — Tablet PCs running XP-Tablet
    - — Desktop PCs running Linux
- *target standards*:
    - — MMS (basic SMIL support for Mobile Multimedia Messaging),
    - — 3GPP-PSS5 [9] (enhanced 3G mobile multimedia support), and
    - — SMIL 2.0 Language profile (the full SMIL specification).

    Note that all three protocols are subsets of the W3C SMIL specification.
- *supported media types*:
    - — streaming video (in the open, DIV-X implementation),
    - — streaming audio and MP-3
    - — streaming timed text, and
    - — streaming images (PNG and JPEG).

    Only those codecs which are available under open source licensing terms will be integrated into the project.

---

1. These are the minimum set of platforms that the project intends to support. We expect to be able to offer support for other platforms (such as Linux on PDAs, optimized Win32 or MacOS support at no additional project cost, depending on the resources of the external development team.

- *supported transfer protocols*:
  — RTP/RTSP [4], [8]
- *implementation language*:
  — C / C++.
- *other deliverables*:
  The project will construct demonstrators and a test set to verify profile and platform implementations. Conference articles discussing scheduling and control issues for heterogeneous platforms will also be produced.

The project will produce its deliverables in three releases:

- *Release A*: this release will contain an implementation of the basic SMIL engine for each of the MMS, 3GPP and SMIL profiles. It will be provided within three months of project start. The initial implementation platforms will be: desktop (Linux) and TabletPC (XP/T).
- *Release B*: this release will provide support for a broad range of codec's for image, text, audio and video media. It will be available approximately 9 months into the project. The target implementation platforms will be: desktop (Linux), TabletPC (XP/Twirlers), PocketPC (WinCE).
- *Release C*: the final release will consist of an optimized player, plus testbed and documentation. This release will be available twelve months after project start. All of the supported platforms will be included.

Each release will be seen as a go/no-go point for further releases. A licensing scheme will be developed to coincide with the provision of Release A.

2.4 Importance

Despite impressive advances in underlying multimedia technology, the quality of media presentations provided to users has hardly changed during the past five years. If this project is completed in a timely manner, it could provide a significant stimulant to the study of new network infrastructure support; this, in turn, could have a major impact on the quality of the end-user media experience and in the way in which society shares information. We expect that this will stimulate other projects in networked information sharing, as well as projects that study the allocation of network resources during the end-to-end presentation of complex, synchronized media.

At recent conferences (ACM Multimedia in late 2002 and SMIL/Europe in early 2003), a strong desire was articulated for the development of an open-

source SMIL player. The lack of such a player was seen as a major impediment to the development of networked multimedia.

2.5   Results Dissemination

The results produced by this project will be disseminated under GNU open source licensing (or equivalent) by the members of the project and by various community sites for PDA, mobile and desktop users. The results of the project will also immediately be integrated into other research being performed at CWI, including the Topia e-learning project being undertaken by CWI and the Telematica Institute. Finally, the player is expected to be distributed in binary form in an upcoming book on SMIL by L. Rutledge and D. Bulterman, to be published by Pearson/Addison Wesley in 2003.

The final distribution and monitoring of results will be performed in association with major distribution partners, including W3C and 3rd-Party platform sites. These site will be identified and arranged after the release of the first set of product deliverables.

2.6   Project Development Team

Given its experience with developing the SMIL standard and in producing complete SMIL engine implementations, CWI has assumed the leadership role in this project. Contacts have been made with leading research universities and institutes in Europe and the USA for project partnering. Once the project schedule is finalized, the development team will be formalized.

3 References

[1] D. Bulterman, *SMIL 2.0 Part 2: Examples and Comparisons*. IEEE MultiMedia 9(1): 74-84 (2002)

[2] ISO/IEC, *MPEG-4 Part 1: Systems (IS 14496-1)*, Atlantic City, N.J., USA, Oct. 1998.

[3] A. El Saddik, *Multimedia Technologies and Applications*, Lecture notes, U. Ottawa, 2002. Available at: http://www.site.uottawa.ca/~elsaddik/abedweb/teaching/elg5121/pdf/06_MMQoS.pdf

[4] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, *RTP: A Transport Protocol for Real-Time Applications*, RFC 1889, 01/25/1996.

[5] A. Kind, R. Pletka, and B. Stiller. *The potential of just-in-time compilation in active networks based on network processors*. In Proc IEEE OPENARCH'02, June 2002

[6] H.K. Wang and J.L.C. Wu, *Object model for hypermedia applications*, In: Computer Communications, Vol. 18, No. 7, pages 475-485. 1995.

[7] F.B. Paulo, M.A.S. Turine, M.C.F. De Oliveira, and P.C. Masiero, *XHMBS: A Formal Model to Support Hypermedia Specification*, In Proceedings of Hypertext '98 (Pittsburgh PA, USA, 1998), ACM, pp. 161170

[8] H. Schulzrinne, A. Rao and R. Lanphier, *Real Time Streaming Protocol (RTSP),* IETF RCF 2326, April 1998. Available at ftp://ftp.isi.edu/in-notes/rfc2326.txt

[9] 3GPP, *Transparent end-to-end packet switched streaming services: protocols and codecs (version 5)*, 3GPP TS 26.234 v5.2.0, September 2002.