Annotating Mobile Multimedia Messages with Spatio-temporal Information

Misund Gunnar, Lindh Mats

Faculty of Computer Science, Østfold University College, Halden, Norway

E-mail: gunnar.misund@hiof.no

Abstract

Mobile messaging started with Short Messaging Service (SMS). Multimedia Messaging Service (MMS) was recently introduced, allowing users to send and receive messages composed of text, audio, images and even movie clips. In this paper, we propose a new step in the evolution of mobile messaging by introducing the Annotated Multimedia Messaging Service (AMMS). AMMS is in short an MMS document augmented with metadata that facilitates efficient structuring, storage, search and retrieval of mobile media content. We currently restrict the metadata to describe time and place, and show how to include this information in an MMS document by using the existing specification without changes. We suggest some applications that may be realized on the foundations of this simple, yet powerful, concept, both in the personal market and in the professional domain. A prototype framework called Been-There-Done-That has been implemented, where users can generate AMMS documents and upload them to an Internet server. The messages are accessible through a web based map application.

Keywords

Annotated Multimedia Messaging Service, Metadata, Mobile Content Creation, Mobile Positioning, Smart Phones, Web Map Service.

I. INTRODUCTION

The mobile phone has become an integral part of everyday life in many countries. In 2002, the number of cellular subscribers in Taiwan reached 22.6 million, meaning that there is more than one handset for each person (Possi, 2004). From being simple voice call devices, cell phones have developed into advanced multimedia communication tools. High-end mobile phones, often called smart phones, have the storage and processing capabilities like a five years old desktop computer (P900, 2003). Smart phones run full-blown operating systems facilitating development of a multitude of third party solutions, such as games and infotainment applications. Protocols and standards facilitate access to Internet services, and enable the users to roam freely, independent of local operators. At last, but not least, contemporary handsets offer a range of multimedia features, like built-in cameras, camcorders and audio recorders.

The Short Messaging Service (SMS), which allows subscribers to send and receive short textual messages, was first demonstrated in 1992 (Karuturi, 2002) and was after a while embraced by the users and ignited an explosion in network traffic. In 2002, approximately 30 billion SMS messages were sent globally each month (GSM, 2004). The SMS protocol was later extended to include transfer of small pictures, sounds and animations. This version of SMS is known as Enhanced Messaging Service (EMS).

SMS and EMS were originally designed for use in Global System for Mobile Communications (GSM) networks, also referred to as second-generation (2G) mobile networks. It is currently the most widely used system, and the bandwidth is around 10 kbps. The specification of a faster network system,

General Packet Radio Service (GPRS), was released in 1997 as an additional GSM service. The speed is typically 150 kbps, or more than 250 kbps in the case of the enhanced GPRS service (EDGE). GPRS is available from most GSM operators, and is often characterized as 2.5G. The third-generation networks (3G) offer a significant jump in bandwidth, ranging from 150 kbps when driving a car and up to two mbps for stationary usage (Oney, 2001). 3G networks were first rolled out in Japan 2001 (CNN, 2001), and presently there are more than 132 million 3G subscribers worldwide (UMTS, 2004).

The advent of faster networks, together with the boost in capabilities and performance of handsets, led to the specification of the Multimedia Messaging Service (MMS). MMS is a natural evolution from text messaging, and makes it possible to exchange messages with rich media content such as graphics, images, audio and video (3GPP, 2003). An MMS message has a number of slides and may be viewed as a "PowerPoint-style" presentation on the mobile device. The first MMS service was launched in Norway in 2002 (ITU, 2004). MMS usage is growing fast. In second quarter of 2004, 20 million MMS messages were sent in Norway, or approximately five messages for each inhabitant, as opposed to 6 million in first quarter (VG, 2004).

Most MMS applications assume that the users are producing simple MMS documents, like a single image or a single video clip, optionally followed by a text note. Typically, the users are invited to take advantage of some central server, offering functionality for storage, structuring and presentation of mobile snapshots or small video clips. An example of such "your

personal mobile album on the web" is the award-winning Foneblog from NewBay (NewBay, 2003). The majority of these types of services are structuring the committed MMS documents according to one single metadata property, however perhaps the most significant one; namely the time of capture.

In this paper, we take advantage of the new and coming generations of mobile handsets, high-speed networks and open formats and standards. Our main contribution is a simple enhancement of the MMS specification (CMG, 2002) which we have termed Annotated Multimedia Messaging Service (AMMS). We review some of the previous work on multimedia and metadata relevant to MMS applications in Section II. The specification of the Annotated Multimedia Messaging Service is given in Section III. We then outline some scenarios based on the AMMS concept. In Section V we describe a proof-of-concept implementation called "Been-There-Done-That", and finally give some concluding remarks.

II. MOBILE MULTIMEDIA AND METADATA

Metadata is essentially additional information about some artifact. Metadata facilitates organizing, searching, browsing, and sharing, see for instance (Naaman, et al., 2004a), where the role of metadata in searching large image repositories is investigated. Metadata management has for a long time been an important aspect of the work in museums, libraries and archives. The rise of the World Wide Web has added another important arena for metadata management. The problems of searching and browsing the massive amounts of information on Internet are the main focus in the large field of research and development referred to as the "The Semantic Web". The cornerstone of this effort is the Resource Description Framework (RDF), in essence a tool for describing and interchanging metadata about web resources. RDF is frequently used together with a multipurpose, general metadata specification called The Dublin Core Metadata Initiative (DCMI, 2003). For an overview of the semantic web, RDF, DCMI and related topics, see (Miller, 2004). The semantic web relies heavily on the use of Extensible Markup Language (XML) as the underlying metalanguage.

The authors are of the opinion that consistent use of metadata is a key to successful deployment of future MMS applications. One of the reasons for this is that it is far more difficult to handle multimedia content than traditional (well-)structured information such as text-only data. However, little, if any at all, attention has been paid to MMS metadata management. As a backdrop for our proposed specification of annotated multimedia messages, we outline some relevant specifications and then review some work concerning generation of media metadata.

Many metadata standards and frameworks may be applied to multimedia. However, most of them are designed for use in museum collections, libraries and archives of analogue media items. In the following, we briefly outline a few research and standardization initiatives supporting digital media metadata.

Not surprisingly, the major bulk of media metadata work seems to be applied to still images, as for instance the "DIG35 Specification" from International Imaging Industry Association (I3A, 2001). This standard defines a flexible XML based framework to describe location, date and time of capture, focus distance, light levels, GPS location, image type, copyright, subject matter, etc.

Another large area of metadata research concerns audio and video, where the major projects are conducted under the Moving Picture Experts Group (MPEG) umbrella. MPEG is a working group of ISO/IEC in charge of the development of standards for coded representation of digital audio and video. Their MPEG-7 specification, "Multimedia content description interface" (ISO/IEC, 2002), offers a rich XML vocabulary for providing metadata covering both content and context. MPEG-7 is an approved ISO standard. It is highly complex, and it is difficult to find examples of real-life use. Another video metadata specification, far more simple in use than MPEG-7, is the "Dublin Core Application Profile for Digital Video" from the Video Development Initiative (ViDe, 2001).

The above-mentioned initiatives all target a single media type, and there are not many projects covering mixed media. The MMS is indeed an example of mixed media, combining audio, video, images and text. MMS uses a subset of Synchronized Multimedia Integration Language (SMIL) as encoding and presentation (3GPP, 2003). SMIL is an XML grammar similar to HTML. It is essentially a way of choreographing rich, interactive multimedia content for real-time presentation over the web and over low bandwidth connections (W3C, 2001).

SMIL comes in two flavors, SMIL 1 and SMIL 2, the latter an extension of SMIL 1. Both versions offer metadata capabilities. SMIL 1 has a *meta* element with two attributes, a *name* and a *content*. There may be several meta elements in one document. The element is generic in the sense that users may define their own metadata properties. Each property has a corresponding meta element with a given name/content pair. The following XML snippet states that there is a metadata type termed "Creator", and that the value of the type is "Mats":

```
<meta name="creator" content="Mats"/>
```

This is a very simplistic way of modeling metadata, and it is not supporting interoperability. Without detailed knowledge of the application that generated the SMIL document, it is impossible for external components to treat the annotations properly. In addition, this is an example of "bad" XML modeling. There is widespread agreement on *not* using the attributes of the elements to carry essential information, but rather embed it in the content of the elements.

SMIL 2 offers a more advanced metadata model. In addition to the SMIL 1 *meta* element, there is also a *metadata* element. This element is a container for well-structured information formatted according to the Resource Description Framework (RDF) model and the element set defined by the Dublin Core Metadata Initiative (DCMI). This enables a growing number of Semantic Web based applications to read and "understand" the metadata information.

It is one thing to find an appropriate metadata specification or standard; another issue, perhaps a more difficult one, is to generate the actual metadata. In traditional media collections, annotations are added after the time of creation, and typically in an archiving context. For personal media applications, such as private digital image collections, growing with hundreds of entries each month, this might become an overwhelming and daunting task.

Recent work has been focusing on using time and location of capture as major metadata components for digital images, see for instance (Naaman, et al., 2004b), where two systems for metadata based browsing in large photo collections are compared. Two of these projects work with advanced frameworks for semi-automatic metadata generation for digital images, the LOCALE system from Stanford University, California (Naaman, et al., 2003), and the MMM (Mobile Media Metadata) from Berkeley, California (Davis, et al., 2004). In the LOCALE project, the users have digital cameras connected to GPS devices so that each image is stamped with time and position. In addition, each snapshot is labeled with an arbitrary string that the user may or may not fill in. When empty labeled images are uploaded to the storage server, the system automatically assigns a label based on other photographs taken in the same area. The MMM system manages metadata for images from mobile camera phones. The initial annotations are username, date, time and the cell ID retrieved from the GSM service provider. Then the system searches a central storage server for images with "similar" metadata, and suggests additional annotations. The process is iterative such that the metadata may be refined in a user/system loop.

The major bulk of research and development in media metadata management is focusing on the personal market segment, typically snapshot collections. In addition, only simple multimedia content is considered, the overwhelming majority being single images. Work on complex multimedia documents, such as messages composed of images, video clips, voice notes and text is non-existing in the literature, as to the knowledge of the authors. Further, examples on work targeting professional applications seem to be scarce.

III. ANNOTATED MULTIMEDIA MESSAGING SERVICE

An MMS message is coded by using a subset of SMIL (see Section II). In order to ensure interoperability among different network operators and handset vendors, a group of major companies has published a conformance document describing the subset of SMIL allowed in MMS (CMG, et al., 2001). One of the optional components of an MMS document is the *meta* element, as defined in the SMIL 1 specification. The conformance document does *not* include the richer RDF based metadata module in SMIL 2.

In this work, we restrict the metadata to describe the most essential information, where and when. Several different technologies make it possible to retrieve the position of a given cellular unit (Chan, 2003). There are reasons to believe that all mobile phones in the near future will be location enabled. As an example, the introduction of the enhanced 911 system (E911) implies that in the near future, all US wireless carriers must provide precise location information for all emergency calls from mobile phones (FCC, 2004).

We define an annotated multimedia message (AMMS) as an MMS document augmented with spatiotemporal metadata. One of our design goals is to take full advantage of existing MMS applications and services, and the only way to do this is to make sure that the AMMS is fully compliant with the current conformance document. As pointed out in Section 2, the metadata capability of SMIL 1 is inferior to SMIL 2, but to keep the AMMS conformant, our only option is to use the *meta* element and its attributes to carry the necessary spatiotemporal information. The three proposed meta elements are defined as follows:

Name	Content
time	The time of the recording, given in the same notation as used by existing e-mail applications.
place	A text string describing the place where the MMS was composed. This name may later be resolved through a gazetteer service to an actual position.
position	A set of one or more coordinate pairs describing a point or a track somewhere in the world. Currently we assume geographic coordinates and WGS84 datum. One or more white space characters delimit each coordinate pair, and a comma separates each two coordinates.

An example of an AMMS follows, where we have expanded the sample document from the MMS conformance specification with spatiotemporal annotations:

```
<smil>
    <head>
        <meta name="time" content="Thu, 24 Jun 2004 13:37:
00 +0200" />
        <meta name="position" content="59.00000,11.00000" />
        <layout>
            <root-layout width="352" height="144" />
            <region id="Image" width="176" height="144"
left="0" top="0" />
            <region id="Text" width="176" height="144"
left="176" top="0" />
        </layout>
    </head>
```

```
<body>
    <par dur="8000ms">
        <img src="FirstImage.jpg" region="Image" />
        <text src="FirstText.txt" region="Text" />
        <audio src="FirstSound.amr" />
    </par>
    <par dur="7000ms">
        ...
    </par>
</body>
</smil>
```

When embedding the metadata as specified, any MMS compatible device may treat the AMMS as a standard MMS, by just ignoring the additional meta elements. On the other hand, all AMMS enabled applications must understand the metadata and be able to process the spatiotemporal content.

We strongly recommend that the next version of the MMS conformance document should include the SMIL 2 metadata module, which would bring MMS metadata management a significant step forward.

IV. AMMS APPLICATIONS

Annotated multimedia messages may be used in wide variety of application. We are of the opinion that additional metadata, in particular spatiotemporal information, may lay the foundation for novel and interesting usage. Up until now, MMS services have in general targeted the private sphere, in particular picture messaging and mobile gaming. However, by enhancing MMS messages with significant metadata, many professionals might consider this as a useful tool in their daily work.

Apart from picture messaging, i.e. when subscribers take a picture and upload it to a central server, the majority of MMS applications pushes content to the users, like games and advertisements. In the following overview of AMMS scenarios, we will focus on the user as a content creator, rather than a content consumer.

A. Blogging

The most obvious usage of the AMMS is as input to a *web log*. A web log is a semi-interactive personal diary that a user keeps on his or her website. The blog (as they are called in short form) is updated with small notes about what goes on in that person's life (Blogger, 2004). Recently the traditional blogs have been extended to handle richer content by adding multimedia elements, as outlined below.

B. Photo blogs

A photo blog builds on the same concept as a regular blog. The photo blog does however put images in focus, and an entry in the blog is rather a single photograph with comments. Photographer Emese Gaal (Gaal, 2001) hosts one prominent

example.

C. Moblogs

While regular blogs usually are operated from a personal computer, there is a trend to extend the blogging activity to include messages from mobile phones. The main point is still the same; persons create posts about what happens in their every day lives, take pictures of views and events that interest them and so on. The main difference is that everything is operated from a cell phone, but the blog is still viewed from a personal computer. The concept is not widespread at the moment, but is apparently catching momentum. Nokia has developed a tool called Lifeblog, a sort of a scrapbook used to keep track of photos, videos, text message and other mobile media uploaded to a website (Nokia, 2004). Joi Ito maintains an informal site about the status of mobile blogging (Ito, 2004). One (of many) commercial actors in the field is New Bay Software with their award-winning FoneBlog service (Newbay, 2003).

D. Location blogs

The natural evolution of blogs and the ever-increasing development efforts put into these projects have resulted in location aware blogs. In addition to being able to store multimedia content, they offer functionality for storing the location of where the content was recorded. The JPEG image format makes it possible to store the geographical coordinates of the moment of capture, provided manually or by the device in real time. Several systems exploit this option, for instance WWMX (Toyoma, et al., 2003) and WaveBlog (WaveMarket, 2003). In addition a system called Blogmapper has been developed, which provides additional location services to already existing blogs (Harlan, 2003).

E. AMMS blogs

Obviously, the AMMS concept may support advanced blogging applications, combining the functionality of traditional blogs, moblogs and location blogs. In Section V, we present an enhanced mobile blogging framework, called Been-There-Done-That.

F. Field reports

While advanced blogging mainly targets the personal market segment, it is easy to apply the AMMS concept to various professional domains. A typical application area is field reporting. We briefly outline two candidate cases.

G. Power line maintenance

Recently a prototype has been presented where power line malfunctions are reported using a helicopter, a PDA (Personal Digital Assistant) and a digital camera. The camera is transferring images to the PDA, and a description of the

situation is recorded using a registration form on the PDA. Finally, the complete report is transmitted to the headquarter using a GPRS enabled network (Langdahl, 2004).

In our opinion, this procedure would be greatly simplified using the AMMS concept. First of all, only one recording device would be needed, and the standardized AMMS message would make it simpler to build a modular system, in particular at the server/browser end. However, it would be necessary to extend our proposed simple metadata capabilities to include error messages relevant to the power line case. This could easily be achieved by using the SMIL 2 specification and its RDF/DCMI extension (see Section II).

H. Biodiversity research

A group at Stanford University, California, has used a location blogging tool to archive records of plants in the surroundings of the Rocky Mountain Biological Laboratory (Garcia-Molina, et al., 2004). The browser uses a wide range of metadata to facilitate easy access to the photographs of the plants and their environment. However, the majority of the additional information is added manually in retrospect, including the location. Clearly, a more automated approach, for example by using the AMMS protocol, would improve and simplify the registration procedure.

I. Where R U

As an educated guess, a large number of all mobile conversations contain one or two questions of the type "Where are you?" By applying the AMMS principles, such questions could be answered very precisely. Assume the recipient of a request takes a snapshot of her surroundings, and then sends it as an AMMS to a "Where R U" server, together with the calling number of the asking unit. Then the server would retrieve the location and produce an appropriate image map with the help of a WMS server (OGC, 2002). Then it would forward the map and the snapshots to the person asking. Clearly, a map and an image (or more) of the location would, in general, be much more informative than a verbal description.

J. Mobile journalism

Journalists have gradually entered the digital domain, and are now using digital cameras (stills and video), digital audio recorders and digital communications such and GSM, GPRS and WLAN. However, the current way of working is characterized by transferring disparate and unsynchronized items covering the same event. Considering the fact the capabilities of handheld devices soon will satisfy the newspaper (and to some extent radio and television network) standards with regard to technical quality; it is easy to image the mobile unit as the main, not to say the single, tool for a field reporter. In fact, taking full advantage of the SMIL capabilities, a news AMMS might provide a full story of an event, including the

location. When received at the news desk server, the story could be automatically converted to fit various channels, such as conventional newspapers, online newspapers, radio and television, and, of course, be accompanied with an informative map produced by a WMS server.

K. Collaborative mapping

While most of today's maps are generated and provided by central mapping authorities, small groups of people or individuals see the need for a more "grass-root" oriented mapping. Building and updating maps from a collection of GPS points and annotations is a conceivable concept. Figure 1 shows a map of the walkways in a park in Gävle, Sweden, generated by the experimental Been-There-Done-That application presented in Section V.
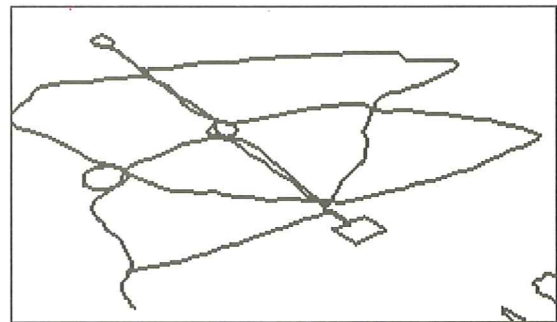

Figure 1. Collaborative mapping

Another, and more elaborated example, is the Amsterdam Real-Time project. During a two months project, a number of Amsterdam citizens were equipped with portable devices connected to GPS receivers, and their movements were traced (Figure 2):

> "This way an ever-changing, very recent, and very subjective map of Amsterdam will come about" (Waag, 2002).

Figure 2. Amsterdam realtime (from Waag, 2002)

One of many rationales for the community mapping approach is the fact that in some parts of the world, it is difficult and/or expensive to obtain good and accurate geospatial data. Collaborative mapping makes it possible to generate highly

customized maps, taking in account the specific location and particular needs of the persons involved.

L. Mobile games

One of the major areas of large revenue expectations is mobile gaming. By adding the spatiotemporal aspect, the expectations should grow even larger, in particular if one does not restrict the domain to the large bulk of "shoot-and-kill" applications. In the following, we briefly outline an application under development by the Norwegian Automobile Association (NAF) together with a major GIS vendor, Bravida Geomatikk (Høseggen, 2004). For a long time NAF has provided their members with a "Michelin Guide" type book of maps and roadside information like restaurants, tourist attractions, scenic routes etc. Gradually, the content of this guide has become digitally available, and the organization wants to exploit the digital content beyond the services offered in the hardcopy version. One proposal is to design a suite of "back seat games", mainly targeting children in various ages during longer car trips. The idea is to combine local based information, quizzes, track visualization, communication with buddies in cars on the same (or other) route, and similar "keep-the-kids-in-the-backseat-happy-and-active-with-some-meaningful-challenges" activities.

An AMMS application connected to various databases, like attractions, historical information, etc. could be the backbone is such a system.

V. BEEN-THERE-DONE-THAT

For real-life testing of the AMMS concept, we have developed a framework called Been-There-Done-That. The infrastructure consists of three main modules, as illustrated in Figure 3. A client application runs on a smart phone linked to a Bluetooth GPS unit. Bluetooth is a short-distance radio communication protocol (Kardach, 2000). The phone client assists the user in composing annotated messages, and sends them to a server-side dispatcher. The dispatcher processes the messages and stores the content for later use. Finally, the user may access the stored messages through a web browser interface.

The modules are relatively easy to implement, and highly interchangeable, due to extensive use of open standards and specifications. We give some details in the following sections.

A. Smart phone client

The smart phone client provides an interface where the user can compose and send annotated messages. The implementation consists of two simple applications, an AMMS manager called BTDT and a location acquisition program called Where. The software is implemented on a Sony-Ericsson P900 (P900, 2003). This device runs the Symbian operating system, and C++ is used as programming language.

B. BTDT

BTDT is responsible for composing the annotated message and transferring it to the dispatcher. The user may choose one of two methods for composing the message. In manual modus, the user specifies the media elements (images, movie clips, audio recordings or text notes) by browsing the file system, as illustrated in Figure 4. In order to provide the spatiotemporal metadata, the user has to select a file that contains the appropriate annotations.

The message may also be generated in a more automatic way, which can be viewed as a recording session. The user pushes the start button, and then generates the multimedia elements by using the built-in standard applications (camera, camcorder, audio recorder). The spatiotemporal annotations are created by the Where program (see next section). When the user pushes the send button, the program traverses the file system on the phone and retrieves all media files that have been created since the application was started. The files can be text, images, sound or video, depending on the media capabilities of the phone. This "two-click" process is illustrated in Figure 5.
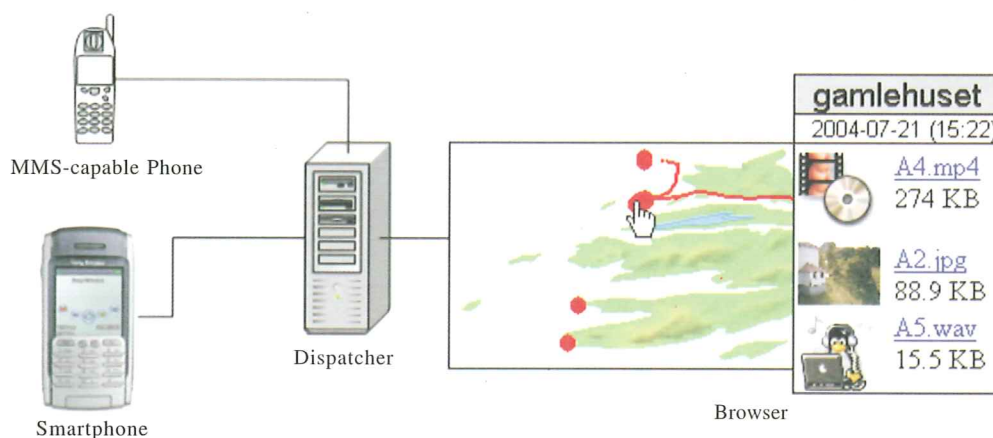

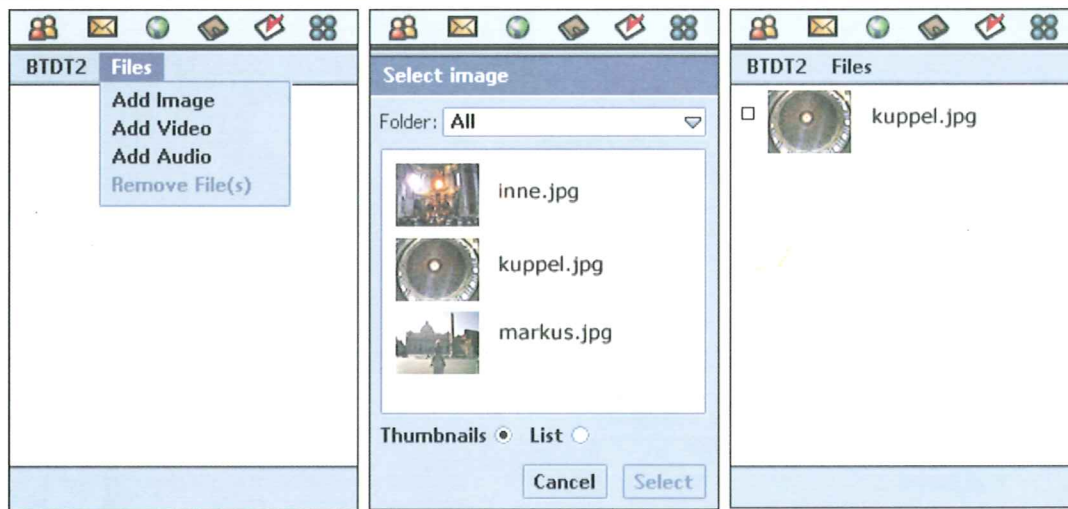**Figure 3.** Been-there-done-that architecture

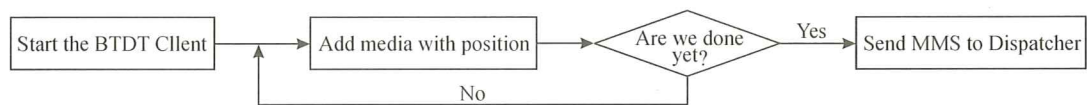**Figure 4.** Manual selection of multimedia elements

Figure 5. Message recording

When it comes to media creation, different smart phones offer completely different services for both users and developers, concerning available media types and types of multimedia applications. By using such a simple method as checking the creation date of a file and comparing it to the time when the application was started, one is able to solve a set of complex problems concerning media retrieval across a variety of platforms.

When the message composition is completed, it is transferred to the dispatcher using a regular TCP/IP connection. On current GSM-based phones, this transmission takes place by GPRS. The BTDT application also provides functions for setting options such as the URL address of the dispatcher.

C. Where

The Where application is responsible for the spatiotemporal component of the AMMS. The location can be specified by a place name or as a point or track of geographic coordinates. Coordinates may be provided either manually or from an optional Bluetooth GPS module. The program may easily be extended to retrieve location information from other sources, such as the network operator. Date and time of day is obtained either from the operating system or from the GPS data. The application generates the appropriate SMIL meta elements, and stores them in a text file. In the case where a place name has been given as location, it is assumed that the dispatcher or another component will use a gazetteer service to look up the geographic position.

D. Dispatcher

The dispatcher works as a server towards the smart phone clients and provides a central storage system for the annotated messages. When a message arrives on the dispatcher, the content is examined and processed according to a given set of rules. The positional data is retrieved from the meta elements. If the location is specified as a textual description, the geographic position is obtained by calling an external gazetteer service. Then the Dispatcher examines each multimedia entry and associates a single coordinate or a collection of coordinates (a track) with the submitted media.

The chosen method of transportation can be any regular TCP/IP-connection. This provides us with a flexible interface separating the dispatcher and the clients. In our case, a client is a smart phone, but it could be any unit capable of generating an AMMS and sending it over TCP/IP.

E. Browser

The browser is an online, web based map interface where users may inspect and retrieve the content of the messages stored in the repository. The main component is a zoomable and clickable map of the world, where areas that have messages assigned to them are highlighted. By clicking on a hot spot, the user is presented with a list of all messages from that particular area. If there are several local areas with information, they will be grouped into one larger area, depending on the size of the current view in the browser. The users are then able to retrieve the information they are looking for by simply selecting it from the list of available media. Predefined areas of interest may be created in order to facilitate easy navigation. The various elements of the browser are shown in Figure 6. Here the user has zoomed in on some hotspots in Rome, and selected two images and a soundtrack from St. Peter's
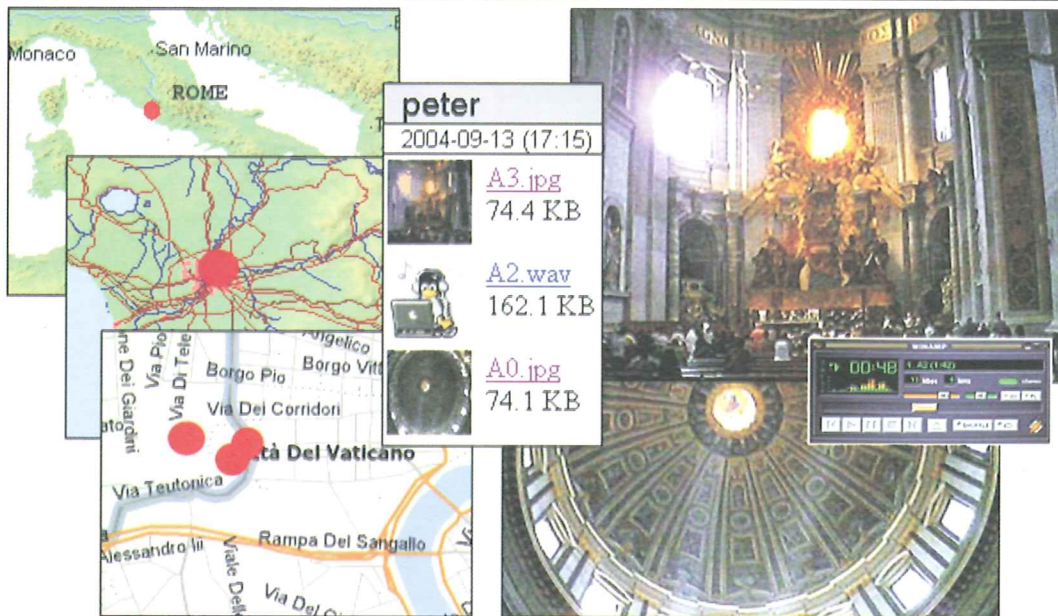
Figure 6. Message browsing

cathedral.

The background map is fetched from external map servers using the Web Map Service (WMS) protocol (OGC, 2002). Hence, the map creation process is completely separated from the rest of the infrastructure. Further, it is easy to customize the map, either by choosing another map service or request another set of layers. The browser is open for public use (OneMap, 2004).

F. Lessons learned

Been-There-Done-That have been in use during a five months test period. Approximately one hundred AMMS messages have been transmitted from various locations in Norway, Sweden and Italy, using five different service providers. The operating conditions have ranged from sitting indoors at a table to standing in an open fishing boat in rough weather.

The smart phone client has shown to be relatively stable, taking in account that it is a rapidly developed prototype. During the field-testing, we learned that extreme attention has to be paid to make the handset user interface as simple as possible. A mobile user has neither the time nor the patience to interact with a complex interface. Ideally, composing and sending AMMS messages should be as easy as sending text messages.

The GPRS communication has functioned very well on most occasions, even when sending messages in the megabyte range. However, we have experienced that network speed varies greatly, depending on operator, location and time.

The Bluetooth GPS has most of the time been carried in a belt holster, and has worked satisfactory. When starting the unit, it typically takes a minute or two to get a lock on the required

number of satellites. Problems have occurred in dense forests or in city areas with tall buildings, a phenomenon, however, common for all GPS receivers. We have experienced that the Bluetooth communication between the phone and GPS unit has been slightly unstable, even when being very close (the distance between Bluetooth devices should not exceed 10 meters).

The dispatcher module has been working flawlessly during the test period, and so have the browser and the underlying map services. Since we have paid little attention to implementing an efficient and appealing user interface, this module would obviously benefit from improvements, preferably based on extensive user centered research.

VI. FINAL REMARKS

We have demonstrated that it is simple and straightforward to extend the MMS format to include spatiotemporal information, and have outlined how to include even richer metadata. The proposed format, Annotated Multimedia Messaging Service (AMMS), is fully MMS compatible according the current MMS conformance document. We have also tested the concept successfully by implementing a test bed called Been-There-Done-That.

It is worth noting that the proposed annotation of mobile media content might have been formatted differently and transmitted in other ways then in the context of the MMS framework. However, we believe that incorporating the annotations in an existing, well known and widely used format, offers significant benefits.

By extensive use of open standards and specifications, in

particular from the 3rd Generation Partnership Project (3GPP) and Open Geospatial Consortium (OGC), we have shown that it is easy for software and content vendors to implement location aware solutions based on the AMMS approach. Two students developed the core parts of the software in a half semester project in their first year of their master's program. The students had no prior knowledge of mobile units, wireless protocols or geographic information.

Currently, there are several projects at Østfold University College aiming at extending the work presented in this paper. We are in particular working on expanding the metadata management by adopting the SMIL 2 specification, and leveraging current research in semi-automatic metadata generation, as reported in (Davis, et al., 2004; Naaman, et al., 2003; Naaman, et al., 2004a). In addition, we are developing applications in the area of mobile gaming and collaborative mapping (Section IV). We welcome comments, contributions and proposals for collaboration. Source code and additional documentation are available on request from the authors.

ACKNOWLEDGEMENTS

REFERENCES

[1] 3G (3G Today). 2004. Over 132 million reported 3G CDMA subscribers. http://www.3gtoday.com/subscribers/.

[2] 3GPP (The 3rd Generation Partnership Project). 2003. 3GPP TS 23.140 Multimedia Messaging Service (MMS); Functional description. http://www.3gpp.org/ftp/Specs/archive/23_series/23.140/.

[3] Grace Agnew, Markus Buchhorn, Dan Kniesner, Jean Hudgins, Douglas King, Mary-Frances Panettiere, Manjula Patel. 2001. ViDe User's Guide: Dublin Core Application Profile for Digital Video. http://www.vide.net/workgroups/videoaccess/resources/vide_dc_userguide_20010909.pdf

[4] Blogger. 2004. blogger.com, 2004. http://www.blogger.com/about/.

[5] Neil Chan. 2003. Introduction to Location-Based Services. http://www.giscentrum.lu.se/www_summeruniversity/projects2003/Chan.pdf.

[6] CMG, Comverse, Sony Ericsson, Logica, Motorola, Nokia and Siemens. 2002. MMS Conformance Document Version 2.0.0. http://www.ia.hiof.no/~gunnarmi/MMS_Conformance_v2_0_0.pdf.

[7] CNN International. 2001. DoCoMo unveils 3G, with caution. http://edition.cnn.com/2001/BUSINESS/asia/09/30/tokyo.docomo3Gdebut.

[8] Marc Davis, Simon King, Nathan Good, Risto Sarvas. 2004. From Context to Content: Leveraging Context to Infer Media Metadata, Proceedings of 12th Annual ACM International Conference on Multimedia (MM 2004). Forthcoming 2004. http: //fusion.sims.berkeley.edu/GarageCinema/pubs/pdf/pdf_63900590-3243-4FA0-845E4BF832AA8BCC.pdf

[9] DCMI (The Dublin Core Metadata Initiative). 2003. Dublin Core Metadata Element Set, Version 1.1: Reference Description. http:// dublincore.org/documents/dces/.

[10] FCC (Federal Communications Commission). 2004. Enhanced 911. http://www.fcc.gov/911/enhanced.

[11] Emese Gaal. 2001. Emese's photo blog. http://www.sciencemeetsart.com/emese/blog/.

[12] Hector Garcia-Molina. 2004. BioAct! http://shark.stanford.edu:4230/cgi-bin/flamenco/bio/Flamenco?username=default.

[13] GSM Association. 2004. SMS (Short Messaging Service). http://www.gsmworld.com/technology/sms/index.shtml.

[14] Arne Enger Hansen. 2004. A Location Bound Media Client for Sony Ericsson P800/P900. Project Report, Østfold University College, Faculty of Computer Sciences, Norway.

[15] Jason Harlan (Map Bureau). 2003. Blogmapper. http://www.blogmapper.com/.

[16] Stein Høseggen. 2004. Mobile gaming based on tourist information. Personal communication.

[17] I3A (International Imaging Industry Association). 2001. DIG35: Metadata – A smarter way to look at digital images. http://www.i3a.org/i_dig35.html.

[18] ISO/IEC. 2002. ISO/IEC 15938: Multimedia content description interface. Forthcoming 2004. http://www.iso.org.

[19] Joi Ito. 2004. Joi Ito's Moblogging, Blogmapping and Moblogmapping related resources. http://radio.weblogs.com/0114939/outlines/moblog.html.

[20] ITU (International Telecommunication Union). 2004. Shaping he future mobile information society: The case of the Kingdom of Norway. http://www.itu.int/osg/spu/ni/futuremobile/general/casestudies/norwaycaseE.pdf.

[21] James Kardach. 2000. Bluetooth Architecture, *Intel Technology Journal*, 2nd Quarter 2000. http://www.intel.com/technology/itj/q22000/articles/art_1.htm.

[22] Subrahmanyam Karuturi. 2002. What is SMS? http://www.funsms.net/what_is_sms.htm.

[23] Bjørn Inge Langdahl. 2004. Registrering av nettfeil ved bruk av PDA og GPS fra helikopter. ItEnergi 2004 (in Norwegian). http://www.itenergi.com/2004/program/.

[24] Eric Miller. 2004. The Semantic Web. http://www.w3.org/2004/Talks/0120-semweb-umich.

[25] Gunnar Misund, Knut-Erik Johnsen. 2003. The OneMap Project. http://www.ia.hiof.no/~gunnarmi/omd/gmldev_02/.

[26] Mor Naaman, Andreas Paepcke, Hector Garcia-Molina. 2003. From Where to What: Metadata Sharing for Digital Photographs with Geographic Coordinates, Proceedings of CoopIS/DOA/ODBASE 2003.

[27] Mor Naaman, Susumu Harada, QianYing Wang, Hector Garcia-Molina, Andreas Paepcke. 2004. Context data in geo-referenced digital photo collections, Proceedings of the 12th annual ACM international conference on Multimedia.

[28] Naaman M, Harada S, Wang Q, Paepcke A. 2004. Adventures in space and time: Browsing personal collections of geo-referenced digital photographs. Technical report, Stanford University. Submitted for Publication. http://dbpubs.stanford.edu:8090/pub/2004-26.

[29] NewBay Software. 2003. FoneBlog - A Web Site for Your Mobile Phone. http://www.newbay.com/whitepapers/FoneBlog%201.0%20White%20Paper.pdf.

[30] Nokia. 2004. Nokia Lifeblog. http://www.nokia.com/nokia/0,1522,00.html?orig=/lifeblog.

[31] OGC (Open GIS Consortium, Inc.). 2002. Web Map Service Implementation Specification (OGC 01-068r3). http://www.opengis.org/docs/01-068r3.pdf.

[32] Project OneMap. 2002. Project OneMap. http://www.onemap.org.

[33] Project OneMap. 2004. Been-There-Done-That Browser. http://www.onemap.org/ch/geometa/browse.

[34] Oney J T. 2001. Wireless Protocols. http://www.nvcc.edu/home/joney/Wireless%20Protocol.ppt.

[35] P900 (Sony Ericsson). 2003. P900 Overview. http://www.sonyericsson.com/p900/index.htm.

[36] Petri Possi (UMTS World). 2004. UMTS/3G History and Future Milestones. http://www.umtsworld.com/umts/history.htm.

[37] Christer Stenbrenden. 2004. A Location Bound Media Server, Project Report, Østfold University College, Faculty of Computer Sciences, Norway.

[38] Symbian (Symbian Ltd.). 2004. Symbian OS - the mobile operating system. http://www.symbian.com/.

[39] Kentaro Toyama, Ron Logan, Asta Roseway, Anandan P. 2003. Geographic Location Tags on Digital Images. Proceedings of ACM Multimedia 2003, http://wwmx.org/docs/wwmx_acm2003.pdf.

[40] VG. 2004. MMS tar helt av. (in Norwegian) http://www.vg.no/pub/vgart.hbs?artid=251070.

[41] W3C (World Wide Web Consortium). 2001. Synchronized Multimedia Integration Language (SMIL 2.0). http://www.opengis.org/docs/01-068r3.pdf.

[42] Waag Society. 2004. Amsterdam Realtime. http://www.waag.org/realtime/.

[43] Wave Market. 2003. WaveBlog. http://www.waveblog.com/.