

Mobile TV becomes Social – Integrating Content with Communications

Raimund Schatz, Siegfried Wagner, Sebastian Egger, Norbert Jordan
Telecommunications Research Center Vienna – ftw., A-1220 Vienna, Austria
{schatz, wagner, egger, jordan}@ftw.at

Abstract. *Mobile TV is a class of pervasive multimedia services which currently enjoys considerable scientific and commercial attention. In this paper, we propose the integration of peer-to-peer interactivity as one likely Mobile TV evolution path. We integrate findings from related fields to develop a definition of Mobile Social TV and outline its design space in order to discuss opportunities of integrating content and communication. Furthermore, we illustrate the concept by presenting our ongoing mobile TV research, technological and user aspects as well as future research directions.*

Keywords. Mobile TV, DVB-H, Social TV, Sociability, Mobile HCI.

1. Introduction

Mobile broadcast TV combines the two best-selling consumer products in history - TVs and mobile phones. The growing number of technological and commercial trials in Europe confirms the following qualities of Mobile TV as key user benefits: mobility and flexibility [12]. Furthermore, the results of user trials state that interactivity will add significant value to mobile broadcast service offers [17]. However, interactivity in this context is mostly understood as human-service interaction in the shape of quizzes, voting as well as electronic service guides (ESGs).

Nonetheless, considering the *social* interactivity perspective, studies on living-room TV viewing patterns show that the majority people favor watching TV together with friends over watching alone [4]. In addition, recent advances in interactive TV technology such as IPTV have paved the way for a new class of interactive TV services: Social TV which aims to provide multiple remote viewers with a joint TV watching experience by adding communication features. Social TV User studies have confirmed high acceptance of the concept [6],[13] because it satisfies a fundamental user need: social interaction within a shared context - the jointly

watched programme. These observations combined with the fact that so far the most successful services on mobile phones (voice calls, SMS) are targeted to peer-to-peer-communications [19], indicate that a hitherto neglected class of interactive Mobile TV services exists, which demands further investigation: *Mobile Social TV*.

In this paper, we extend common notions of interactive Mobile TV towards Mobile Social TV by integrating concepts of peer-to-peer (P2P) interaction. It is structured as follows: Chapter 2 outlines relevant concepts from related fields in order to provide a definition of Mobile Social TV and discussion of its design space in Chapter 3. Chapter 4 discusses Mobile Social TV research issues and related requirements. Chapter 5 presents our Mobile TV research infrastructure and its application to Mobile Social TV scenarios. Finally, we discuss conclusions and future research in Chapter 6.

2. Towards Mobile Social TV

The most efficient way to deliver TV to the mobile is a pure re-broadcast of ordinary TV programming. Such kind of “TV mobile” creates the same user experience that older types of analogue mobile TV receivers like the “Sony Watchman” already provided. This first generation of broadcast services rolled out in 2006 offers little interactivity or content adaption towards mobile consumption requirements, a condition that may be temporarily sufficient and cost-effective for content and service providers. However, the first-generation rebroadcast TV approach remains unsatisfactory, as it does not harness the full potential of Mobile TV.

2.1 Mobile TV: Content and Interactivity

Mobile TV is supposed to evolve beyond small-screen TV due to native differentiating features of mobiles: portability, immediacy and connectivity. Mobile TV therefore can be expected to develop beyond normal television

(above described as “TV mobile”) because of these differentiating elements. New or adapted *content* programming for mobiles is one approach. A second dimension is *interactivity*, similar to living-room interactive TV. Both concepts (content-adaption and interaction) constitute the characteristic design space for mobile TV evolution: it starts from TV mobile up to full content-adaption and interaction (i.e., true “Mobile TV”) as illustrated by the Fig. 1 below.

Interactivity	None	Medium	Full
Content-adaption			
Full	Dedicated production of programmes		TV made for mobile, e.g. mobisodes with chat portal
Medium	Re-editing	COSTS & COMPLEXITY	
Light	Re-mastering, re-rendering of text		
None	Rebroadcast without modifications (TV mobile, 1 st generation)	Access to related content (portal), feedback (e.g. voting)	Multimedia portal, feedback, plus person-to-person interactivity

Figure 1. Evolution of TV made for mobile

Above differentiating elements - interactivity, P2P interaction, ubiquitous service access and adapted or made-for-mobile content formats - characterize the upcoming future generations of Mobile TV services. In the following sections we present our arguments that mobile phones are prime candidates for delivering not only single-user interactive Mobile TV, but also for its fusion with multi-user P2P-interactivity.

2.3 The Mobile Communications Perspective

With the use of TV on mobile phones two additional factors appear, beyond content and interactivity: *social connectivity* and *mobile context*.

Social connectivity. In contrast to living-room TV sets, mobile phones are natively social devices: since their introduction, users have learned to use them as social connectivity tools, with voice-calls and SMS as key functions. Therefore, the Mobile TV user experience is intrinsically shaped by people’s usage of mobile communications and services. Particularly younger mobile phone users tend to use mobile communication services to enhance what Carsten

Sørensen refers to as ‘mobile intimacy’ and ‘new networking’ [21]: extending personal relationships, socializing and networking through personal cellular devices. His study shows that mobile is changing the love lives of youngsters who exchange invitations for dates and sexually explicit texts.

Contextual factors. Mobility means that content consumption takes place in various dynamic mobile contexts e.g. on the go, in the bus or at work [12], a direct contrast to the static embeddedness of living-room TV. Mobility thus means limited attention spans, but with increased user readiness for interruptions and interaction. Furthermore, mobile displays are considerably smaller than living-room screens. The consequence of these contextual factors is that static TV is superior in creating immersive, passive media experiences. Mobile TV might never be able to lull people in the same way as high definition television, but on the other hand it allows for more *interactive* and *intimate* experiences.

On behalf of above connectivity and context aspects Vincent and Harper [19] argue that so far, the most successful services and products that satisfy emotional needs have consisted of *person-to-person connectivity applications*. The social value that people attribute to these services currently is much higher than the value given to pure information services (i.e. data services). This is also reflected in the market success of mobile communication compared to data services. Consequently, the mobile communications perspective indicates that broadcast services such as re-broadcast Mobile TV might face limited user acceptance and market success because of being conceived as *non-personal single-user-infotainment*. The implication is that mobile TV service designers should consider the integration of emotional factors and person-to-person connectivity.

2.4 The Social Presence Perspective

While co-presence is most commonly used in face to face interaction analysis, two of its dimensions are present in any kind of mediated presence. The first is *knowledge about peers* through the presence function or the chat window, and message exchange with these peers regarding the program watched. Knowledge about others, watching the same program is vital to the feeling of being involved in a social setting. The second dimension is usage of *phatic*

responses. Their appearance is usually not intended as they are reactions related to certain occurrences on TV. Hence, phatic responses do not convey relevant information but they mediate participation of other peers. Therefore, they contribute vitally to a sociable atmosphere [13]. Providing these two dimensions of co-presence in the context of TV viewing enables a social experience close to the joint living room setting – Social TV.

2.5 The Social TV Perspective

In the traditional living-room context, TV viewing mostly is an immersive, social co-located group activity. The two key reasons are: in most TV watching situations, users just want to relax with content ready for passive consumption TV programming – the classic “lean back” condition. The second reason is the social gratification aspect when watching TV together and having the agenda setting function of popular TV shows, sports and news formats [1].

This agenda setting function has been further investigated by Douglas Rushkoff who redefines the role of content as catalyst for communication and social interaction: *‘The driving force behind all of the authorship and creative energy of the networked age is the need to create [...] social currency.’* [15]. The concept of Social Currency proposes, that in networked societies, content is not something which is simply consumed. Content rather functions as a *medium* for social interaction between people (like a good party joke), providing common ground and shared interests. This perspective redefines the role of broadcast: instead of being a plain consumable, TV content serves as conduit of social interaction i.e. socializing around the content might be more important than the content itself. Exactly this intrinsic *sociability* of content has given rise to the concept of Social TV.

Social TV itself is a form of computer mediated communication that primarily addresses solitary viewers feeling short of co-located peers: it aims to *provide multiple remote viewers with a joint watching experience*. The strategy is to provide mediated *social presence*, a “sense of being together” [8]. Its sociability is twofold: direct sociability parallel to watching a program, and indirect sociability, taking place far after the related event e.g. chatting with colleagues at work about the TV program from yesterday evening.

A prime example of living-room Social TV is AmigoTV, a set-top box based system that integrates television programming with communication and community functions. It allows users to communicate with their buddies while watching television by using audio conferencing, graphic symbols, and avatars [3]. User studies performed at our centre have shown a high acceptance rate of the Social TV concept, with almost equivalent ratings of the co-located group and the mediated-presence conditions [6].

3. Mobile Social TV

This chapter proposes a definition of the term “Mobile Social TV”, our research motivation and an initial exploration of its design space.

3.1 Term Definition and Research Motivation

Our observations and user study results have encouraged us to transfer the Social TV concept to the mobile domain. The rationale behind is that if users can communicate with each other and can consume TV via their mobile devices, there is a certain propensity to use both together. In this sense we define Mobile Social TV (MSTV) as *“a form of computer mediated communication which creates a joint Mobile TV watching experience”* – the integration of Social and Mobile TV.

Even though the idea of Mobile Social TV itself seems intriguing, there are potential problems: for example, concurrent conversation also means competition for the user’s limited attention resources. Geerts [8] compared Social TV text and audio chat and found out that applications need to balance watching and communication activities in order to minimize distraction from the shown TV programme.

Although a considerable amount of research on Mobile Multimedia Services or Social TV in static contexts (see [2],[3],[8],[13]) does exist, to our knowledge there is only little research on live *mobile* Social TV. Open MSTV research issues address technical and user aspects, opportunities vs. challenges, and MSTV interaction design. This situation has motivated us to investigate the MSTV design space and set up a research infrastructure for user studies in the context of live Mobile TV.

3.2 MSTV Interactivity Design Space

MSTV services are characterized by their capacity to provide mediated P2P interaction. The resulting social interaction design space of MSTV features several dimensions:

Table 1. MSTV Design Space Dimensions

Dimension	Example Service
Asynchronous vs. Synchronous/real-time	Show community blog vs. Voice conferencing
Verbal vs. non-verbal/emotional	Text chat vs. Emoticons/Symbols
Integrated with TV content vs. Separated	Shared content ratings vs. Independent Channel-chatroom
One-to-one vs. One-to-many	Peer Show-notifier via SMS vs. Shared bookmark
Push vs. Pull	Text-Chat/notifier vs. Show community blog
High-Bandwidth vs. Low-Bandwidth	Rich communications vs. Text chat

Since mobile devices are characterized by immediacy, intimacy and considerable technical constraints, we propose that any practical investigation of Mobile Social TV should start with *real-time low-bandwidth communications* in order to avoid unduly implementation complexity. Furthermore, real-time communication is a main driver of social presence, because it creates a level of intimacy unlike asynchronous forms of communication, particularly in mobile contexts [19].

For these reasons, we chose to prioritize the following two dimensions for our initial investigations: *verbal/non-verbal* and *one-to-one/many*. Representative interactivity features are text-messaging/chat with presence and awareness, emoticons, JointZapping and ShareMarks:

Table 2. MSTV Interactivity Features as prioritized by the authors.

Dimensions	Verbal	Non-verbal
One-to-one	SMS Private text chat	Joint Zapping ShareMarks
One-to-few		
One-to-many	Text chat in public TV-chatroom	Presence, Emoticons

While chat and presence are common P2P features, the matrix also exhibits two more unusual elements: *JointZapping* and *ShareMarks*. JointZapping synchronizes the current channel of two receivers, in order to ensure that peers keep on watching the same content. Associated with communication, this offers users the possibility to negotiate about the

jointly watched program. This would address the social activity of joint TV planning reported by Turner and Cairns [18]. For looser coupling, ShareMarks serve as “links” to the show/channel currently watched, which a user sends to another peer e.g. via SMS/MMS. If accepted by the receiving user, her TV viewer application switches to the referenced programme.

Although not in the focus of this paper, we want to point out two more advanced MSTV services to exemplify the remaining design space dimensions: *blogs* and *Rich Communications*. Blogs (and other persistent forms of text communication) support asynchronous long-term community building, due to (on average) longer messages and their lack of immediacy. Nonetheless, blogs might well complement real-time communications, e.g. in the shape of collaboratively annotated EPGs for community building. In contrast, Rich Communications (e.g. video/audio-conferencing) enables live high-bandwidth communications. However, parallel integrated TV and Rich Communications on today’s mobile devices pose considerable technological and usability challenges such as graceful handling of multiple audio channels (voice chat vs. TV). Therefore, we consider these dimensions as advanced MSTV research issues.

4. Mobile Social TV Research Issues and Requirements

This chapter discusses MSTV research issues as well as related MSTV user and system requirements. To summarize our previous discussion and related work, the most critical MSTV *research issues* are:

Genre suitability. Existing investigations regarding accepted content for mobile TV have to be critically analyzed for applicability to the idea of MSTV. Mobile TV user studies have identified certain content (e.g. news) as highly suitable for mobile TV due to genuine conciseness and modest attention requirements (e.g. music clips), while other types of content have been identified to be not very suitable for mobile TV. For example, Knoche and McCarthy [11] identified a limited suitability of sports TV due to sport’s strong dependence on the sociable context it is normally viewed in. Since MSTV exactly adds the required sociability, dedicated studies concerning content suitability for MSTV are required.

Interaction channel fit. As mentioned in the previous chapter, an MSTV key issue is attention distraction. Interaction channels differ in terms

of distraction potential and ease of use. While audio chat is easy to use even for novice users, it also can consume excessive amounts of user attention because speech utterances are ephemeral. On the other hand, text chat can provide limited history log ability but it is much more circumstantial to use. In respect to this trade off, assessing which type of interaction channels offers the best balance between these inherently converse properties is needed

Contextual assignment of comments/utterances. The history log property of text chat allows users to catch up with the discussion thread in times of content breaks [20], giving rise to features for assigning utterances made to their related triggering events. Easy assignment would enhance correct emotional interpretation of the utterances shown in the chat history, thus increasing perceived social warmth. One implementation option are small TV screenshots of the trigger events depicting the contextual background of the utterance.

The features necessary for P2P-interaction present *additional requirements* for Mobile TV delivery systems. Concerning *user* aspects, key requirements are the following:

Interaction channel. For socializing mobile TV, an interaction channel has to be offered enabling the users to interact. Realisations for an interaction channel can be text-, audio-, or even audiovisual chat functions.

Minimum distraction. As user perception of MSTV may be negatively affected by interruption caused from social TV features, those features should not interfere with the user's TV viewing activity (e.g. with unwanted message pop-ups). The system design should minimize user annoyances and distraction from the main TV content in order to maintain a seamlessly integrated user-experience [8].

Further *technical platform* requirements are:

Group communication service enablers. MSTV requires a suite of service enablers (e.g. presence, awareness, identity management) similar to groupware, instant messaging and telecommunications systems. These enablers should be based on established communication standards such as XMPP [9].

Low-latency event processing and transmission. Beyond satisfying basic user requirements such as rapid interface response times, latency is a critical issue for MSTV applications. Particularly real-time communications (e.g. text chat) and

synchronised states (e.g. JointZapping) require fast event processing as well as efficient mechanisms for routing and pushing data to mobile clients.

5. Test Platform and Client Architecture

In this chapter we discuss architectural and technological aspects as relevant for MSTV related research activities. Our research focus on telecommunications [7] and the evaluation of advanced service paradigms such as MSTV required us to develop a universal research testbed for interactive Mobile TV based on the DVB-H, the European de facto standard for mobile broadcast [14]. Concerning user interaction, we focus on the integration of personalized, interactive content and navigation with broadcast content on Symbian *Smartphones*. For prototyping interactive TV services we utilize a browser-based web-application approach. However, in contrast to the current practice of using separated terminal-applications for web-interactivity and video, we integrate the browser directly in our custom media player.

5.1 Architectural Overview

Fig. 2 overleaf shows the basic communication networks and architectural components of our hybrid testbed. Content, metadata and events are transmitted via both the broadcast and unicast paths. We use open source software tools as basis for almost all required components. Linux-based tools for Video encoding, IP-Datacast and multimedia content allow for content modifications in real-time.

The unicast path via 3G/UMTS is used for user-specific content and backchannel-interactivity which cannot be provided by the broadcast path. A J2EE 5.0 application server [10] hosts the business and presentation components which process the HTTP requests from the mobile iTV client. We realize data push to the mobile client (e.g. chat events, JointZap channel switches) using lazy HTTP connections. With Lazy HTTP the software client leaves the connection open in order to receive XML-data chunks asynchronously sent by the server application. Alternative methods for wireless data push are SMS or WAP-Push. However, both approaches require tight integration with the MNO infrastructure and are better suited for message types which occur less frequently (e.g., ShareMarks).

The application server also hosts the mediation components which communicate with the broadcast media processing tools via SOAP. This link enables the media encoding and data carousel generator to integrate public information and interaction events (e.g. channel chat messages) generated by the interactive service in the media and data-cast streamed via the broadcast path.

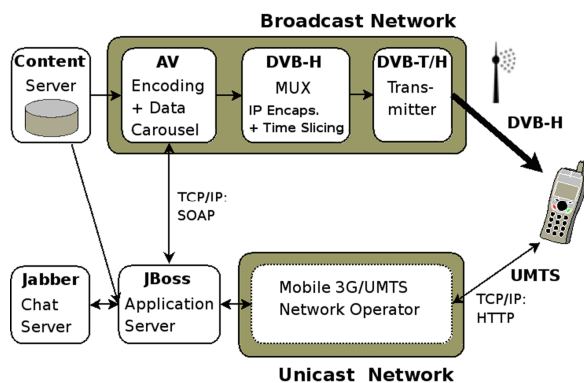


Figure 2. The AMUSE DVB-H Testbed Setup

To ensure the flexibility and extensibility required for prototype service development, we utilize Seam [10]. Seam is a Service-oriented architecture (SOA) framework which reduces the inherent complexity of J2EE by unifying popular web technologies such as AJAX, Java Server Faces (JSF), Enterprise Java Beans (EJB3) and workflows. In addition, the application server also hosts the gateway to P2P service components: a Jabber IM (Instant-messaging) server [9] provides real-time communication features such as instant messaging, presence, awareness, as required by MSTV. For further architecture and implementation details of the testbed, please refer to [16].

5.2 Interactive Smartphone Client

Our client media player is based on Symbian OS platform for current 3rd edition S60 mobile devices such as the Nokia E61 and N92. The Symbian APIs provide basic UI widgets, including the MultiMedia Framework (MMF) which supports playback and rendering of audio/video content. Additional types of multimedia content can be processed and rendered by the respective APIs of the Nokia S60 Webkit, the open source browser development kit for the Symbian S60 platform [5].

Above APIs enable the realization of a split-screen solution with content and navigation frames to create the integrated user experience

demanded for MSTV (see Fig. 3 below). A major requirement is efficient utilization of screen real estate: the interactive mobile TV user experience runs the risk of being impaired by the small screen size and resolution of mobiles. Therefore, our MSTV client's window component management allows for screen layouts that accommodates flexible component placing and sizing.



Figure 3. MSTV chat client during Sports TV

We are aware of the fact that the capabilities of our browser-based approach does not match interface implementations in Macromedia Flash or native Symbian C++, despite recent developments such as AJAX. On the other hand, simplicity remains a key priority for Mobile TV service GUIs due to mobile usability constraints anyway. Second, using web-technologies not only allows for rapid prototyping of interactive Mobile TV applications, but also represents a standard mode of service deployment within mobile telecommunications.

6. Conclusions and Future Work

In this paper we presented the concept of Mobile Social TV (MSTV) as the integration of Mobile TV and Social TV. We proposed MSTV as a highly relevant research topic, since P2P-interaction features create novel ways to integrate mobile video content with communication and mediated presence. We laid out some of these ways with an initial exploration of the design space for MSTV services. We have also shown that the integration of P2P interaction has been successful in other fields such as living-room Social TV. Nonetheless, MSTV comes with a number of potential issues (distractions, privacy, mobile context, sociability) that need to be addressed

adequately. Therefore, more research is required on MSTV concepts, interaction design and technologies. To this end, we presented a testbed that allows for flexible prototyping and evaluation of interactive hybrid multimedia services and technologies. Concerning the Smartphone client, we integrated video display with browser functionality to realize a unified rich media player for the S60 mobile platform. This architecture not only enables the required seamless integration of TV viewing and P2P-communication, it also allows for rapid prototyping of MSTV concepts.

Regarding future work, next steps include MSTV evaluations based on a representative suite of prototypes. User studies in the lab and in the field should not only address user acceptance of the concept in mobile contexts, but also critical MSTV interaction design issues. In this context, the key challenge remains split attention, requiring interface concepts for user distraction management and content adaption strategies such as interspersing intermissions.

7. Acknowledgements

This research has been performed within the projects M2 AMUSE 2.0 and U0 SUPRA++ at the Telecommunications Research Center Vienna (ftw.) and has been funded in the framework of the Austrian Kplus Competence Center Programme. We'd like to thank M2 team members Andreas Berger, Michael Jank and Erwin Wittowetz for their fruitful work on the hybrid testbed and interactive clients.

8. References

- [1] Barbara L & Robert S L. How and Why People Watch TV: Implications for the Future of Interactive Television. *Journal of Advertising Research*, 35/6, 1995.
- [2] Chuah, M. Reality instant messaging: injecting a dose of reality into online chat. In *CHI'03 Extended Abstracts*. ACM Press, New York, NY, 926-927.
- [3] Coppens, T, Handekyn, K & Vanparijs, F. AmigoTV: A Social TV Experience through Triple-Play Convergence. www.telecomreview.ca/epic/internet/intprp-geert.nsf/vwapj/AmigoTV.pdf.
- [4] Engel, B & Windgasse, T. Mediennutzung und Lebenswelten. Ergebnisse der 9. Welle der ARD/ZDF Langzeitstudie ‚Massenkommunikation‘. *Perspektiven* 9/2005.
- [5] Forum Nokia Homepage, <http://forum.nokia.com/>
- [6] Fröhlich, P, Baillie L & Schatz R. Exploring the Joint iTV Experience, FTW-Technical Report-2006-005, January 2006.
- [7] ftw, AMUSE 2.0 Project Homepage, <http://amuse.ftw.at>
- [8] Geerts, D & Calvi, L. Investigating new user experience challenges in iTV: Sociability at CHI 2006, Nov 2006.
- [9] Jabber Software Foundation, <http://www.jabber.org/>
- [10] JBoss Seam, <http://labs.jboss.com/portal/jbossseam>
- [11] Knoche, H & McCarthy, J. Mobile Users' Needs and Expectations of Future Multimedia Services in *Proceedings of the WWRF'12*, 2004.
- [12] Mason, S. Mobile TV – Results from the BT Movio DAB-IP trial in Oxford. *EBU Technical Review*, 2006.
- [13] Oehlberg, L et al. Social TV: Designing for Distributed, Sociable Television viewing. In *Proc. EuroITV 2006*, Athens, Greece.
- [14] Reimers, U. *DVB: The International Standard for DTV*. Springer, 2001.
- [15] Rushkoff, D. *Get Back in the Box: Innovation from the Inside Out*, Harper Collins, 2005, ISBN 0060758694.
- [16] Schatz, R, Jordan, N & Wagner, S. Beyond Broadcast – A Hybrid Testbed for Mobile TV 2.0 Services. In *Proc. ICN'07*, Martinique, April 2007.
- [17] Sodergard, C. *Mobile Television – Technology and User Experiences*, VTT Information Technology Espoo. Finland: VTT Publications, 2003.
- [18] Turner, N, Cairns, P & Jones, M. Dispersing the Interactivity: Mobiles and Electronic Programme Guides. *ACM CHI Workshop on Investigating new user experience challenges in iTV*.
- [19] Vincent, J & Harper, R. The Social Shaping of UMTS-Preparing the 3G Customer, *UMTS Forum Report* 26, 2003.
- [20] Weisz, J D et al. 2007. Watching Together: Integrating Text Chat with Video. In *Proc. in of CHI 2007*, San Jose, California, USA.
- [21] YouGov Mobile life report 2006, <http://www.mobilelife2006.co.uk>