

ENHANCED CAPABILITIES OF PROCESSOR-BASED DIGITAL RADIO SYSTEMS

Günther Schatter

`schatter@uni-weimar.de`

Bauhaus University Weimar, Faculty of Media, Germany

ABSTRACT

This paper presents a survey of our recent efforts for the improvement of a processor-based digital radio DAB. The aim is to overcome the restrictions of radio in availability of content, presentation format and airtime. The presented solutions are: an interest filter-based personal recommender/recorder for audio content, a voice control of the device, a speech-enabled information retrieval system and a solution for broadcasting of podcast content. The focus here is on interactivity, usability, cross-linkage, and on-demand qualities. The prototype is based on a flexible framework (Java).

INTRODUCTION

The advent of new media devices has changed the habit of information reception. Now younger audiences are taking an active role in personalising their content, schedule, and even devices. The traditional AM-FM-radio is the last representative of a media system which is not allowing to search interesting content and does not enable to immediately listen the items or store them electronically. The younger generation typically wants more choices, interactivity, and usage of content on demand. People will overcome the habits of a »rendezvous-listening«. It is to observe a clear trend to enjoy audio independent from temporal and geographical constraints. The audience expects those behaviours from all media devices or they turn away from an outdated radio. We think, future broadcasting devices will provide solutions for an extensive multi-channel content access with on-demand facilities combined with visual support, and huge storage capabilities.

In the last decade much effort has gone into the development of new interfaces (1), and additional data services for Digital Audio Broadcasting (DAB). The technical specifications for digital radio have been refined (2). Numerous qualified services were established (3) as we know them from similar TV- and WWW-applications (broadcast websites (BWS), dynamic label plus (DL), Intellitext, Journaline, moving images (DMB), slide shows, electronic programme guides (EPG), file caching, VoiceXML for DAB etc.). The Digital Audio Broadcasting is likely to turn from a classic audio receiver into a more sophisticated multimedia platform. These new capabilities could be used now to provide a better and more complete user experience as well (4).

Further, we think, a new set of applications could be introduced to provide increased usability and user friendliness. By doing so, radio can overcome its restrictions in content, presentation format and time and to offer more comprehensive and contemporary choices. Hence, it is necessary to look beyond the delivery of improved quality of audio content and to survey datacasting applications. However, preconditions are: adequate processor performance, storage volume, enhanced (graphical) user interfaces (GUI) and displays. Further improvements in hardware can be expected for many years to come, these new

capabilities could be used to provide sophisticated radios that will evolve into hybrid systems which are integrated into manifold handheld devices.

This report is a summary of selected activities and contributions concerning the advancements of an improved processor-based digital radio with the aim to attract particularly younger people. The presented system consists among other things of:

- an interest filter-based personal recommender/recorder for audio content,
- a voice control of the device,
- a speech-enabled information retrieval system and
- a solution for broadcasting of podcast content.

The focus here is on interactivity, usability, cross-linkage, and on-demand qualities. The prototypes were implemented on the basis of a DAB receiver (DR Box 1, Terratec) connected via USB with a standard laptop configuration (Core2Duo; 1.8 GHz; 4GB DDR3-RAM; 500 GB HDD). The System uses Java JDK 6/MySQL 5.1 and is operating portable or mobile. No major obstacles exist to apply other communication channels such as satellite and HD RadiosTM or Digital Radio Mondiale.

PERSONAL RECOMMENDER AND RECORDER

Media consumption is mostly related with an overwhelming set of choices and aesthetic details. Broadcast systems offer a limited control only: users select stations, not really content. Recommender systems help users to find out the content of interest. Our efforts are carried out to adaptively convey the methods of the WWW to search, classify and evaluate information to broadcast receivers. This is based on the development of a personalised DAB receiver that obtains user preferences explicitly and by autonomous recognition from condensed usage scenarios as well. The receiver automatically monitors a multiplicity of programmes to recommend, store and offer interesting contents (5).

In order to transmit information on planned programmes radio stations use the electronic programme guide EPG. This programme overview is based on a standardised XML format. The programme guides are consistently structured and can be processed by computers. It is essential to note that EPGs use the classifications schemes of the TV-Anytime standard. The EPG enables the radio stations to transmit information in various ways. The programme is described by elements that can be recursively structured. Numerous descriptors are available for the title, the point in time and the duration, short, middle and long descriptions of the programme, the genre, references to other information resources, relations to other programme elements etc. The aim is the development of a personal receiver including the following features:

- explicit definition of user preferences,
- automatic recognition of user preferences,
- permanent analysis of entire programme ensembles,
- creation of lists containing recommended programmes,
- recording of programmes relevant for the user.

There are two fundamental concepts to capture and analyse the preferences of users (6):

- The user explicitly enters preferences via an interface.
- By analysis of the user's behaviour the system deduces preferences implicitly.

A combined application of these principles may construct a profile of interest in either case. The user preferences in combination with the metadata constitute the fundamental information units for the recommendation of programme elements. Figure 1 illustrates the mentioned elements and outlines the necessary connections to provide the functionality of the application.

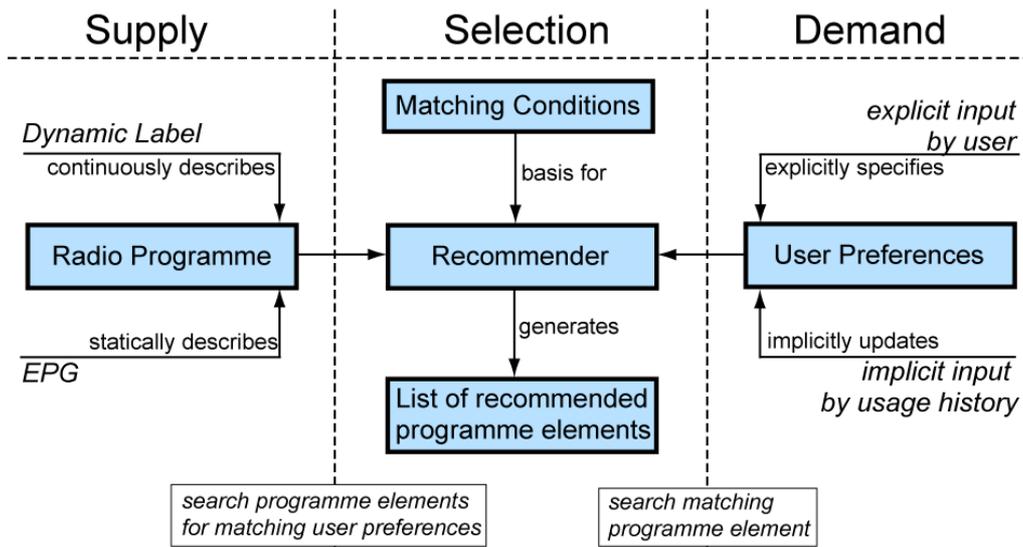


Figure 1 - Personalisation of the programme selection

The application extends the radio with the capability to clearly increase this density of information without necessarily demanding any additional efforts from the user. After some time of usage the radio potentially suggests interesting programmes on other stations or at other times. The longer this functionality is used the more reliable the recommendations of the radio will be. The radio may develop from a largely passively perceived medium to an individual media manager. Particularly the functionality of implicit preference definition by continuous analysis of the user's behaviour seems promising to the authors. By this means the recommendation of programme elements is possible without any additional efforts by the user. For the enhancements of the DAB-metadata system the following suggestions are proposed:

- Main radio stations must send DL and EPG in a proper and reliable manner,
- The time synchronisation of the DL with audio has to be defined,
- The DAB variable programme type PTy shall be dynamically used,
- TVA is to be defined radio- and country-specific.

Up to now, the radio was an easy medium: switch-on, choose a station and listen. This simplicity should not be sacrificed to annoying input procedures. The current state of our work did not allow a comprehensive evaluation of the system, however some people tested the hole system. The purpose was to identify weaknesses and to offer suggestions for improving the system. The appliance of the personal audio recorder attracted interest and is generally considered as easy and convenient. Most people show interest in a solution providing audio- and metadata simultaneously and believe that the system is encouraging to listen more attentive, for longer and at different times.

A specific problem is the granularity and depth of provided classifications to choose in the Interest Manager. Intentional our system provides only a small choice of content classification. The TV-Anytime standard allows more than 1000 different classification statements in several categories, what is difficult to provide without overstraining the user. On the other side the offered descriptors are often not suitable enough for a sophisticated audio classification. The number of interaction dialogues must be adapted well considered. A strict balance between disturbing the user and a pleasant personalisation of the system must be found. Another challenge is the adaptation to portable devices. Therefore the design of appropriate interfaces and principles of user interaction is required.

Each programme is shown as a table entry with title, start time and description, see Figure 2. It can be flagged for recording by using the radio button top right.

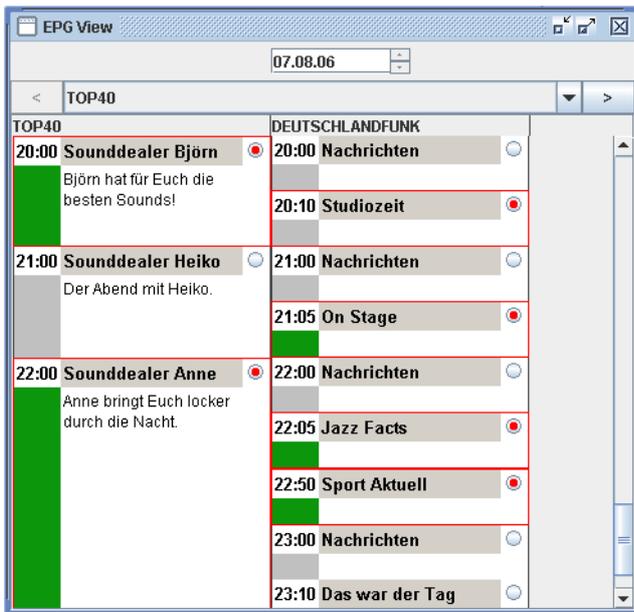


Figure 2 – Electronic Programme Guide



Figure 3 - Prototype of GUI for digital radio

The dark green bar at the bottom left indicates a recommended event which is automatically set for recording if the user specified it. The marked radio button indicates a programme which is flagged for recording. A GUI for the digital radio is shown in Figure 3.

SPEECH-BASED INTERFACES AND INFORMATION SERVICES

Human speech is the most natural and appropriate way of communication, notably if the involved devices produce audible content. The benefits of a speech-conversation system have for instance visual impaired persons, people who possess a miniaturised device without a display, and drivers. On the other hand, the radio is widely conceived as a live, passive, and volatile medium. The obvious advantages of high topicality and mostly free of charge reception are related to the lack of a search and memory function as we know it from the seemingly endless resources of the Internet. This fact is significantly important with respect to content users might want to listen to, like general and traffic news, miscellaneous forecasts etc.

For these two reasons our aim was to develop the concept of an entirely speech-based user interface for a digital radio in combination with the possibility to specifically search for stored audio and data content in a DAB receiver (7). The system utilizes a multichannel receiver with the possibility of simultaneously monitoring a variety of information sources with the purpose to generate spoken replies to verbal user requests. Furthermore, the system embraces an audio repository storing speech-based content elements to diminish the volatile property of broadcast content and to enable a memory function.

Because a lot of information is simultaneously transmitted as text messages in a DAB receiver environment, it is obvious to apply text-to-speech (TTS) and automatic speech recognition solutions (ASR) for a communication by speech assistance, because displays are often not appropriate. After recognising speech, the information is available as text. Internet-based information is applicable as well. BWS may contain multifaceted news, press reviews etc. and are much easier to handle for information retrieval. The structure of the entire system is shown in Figure 4. The multimodal user interface enables the user to interact in two ways, via a:

- Speech-based User Interface (SUI),
- Graphical User Interface (GUI).

It is important to note that both interfaces could be utilised for either case. They offer much more flexibility to the user in terms of navigation and information access, content can be presented with the visual and the audible interfaces concurrently. The user is able to specify a query by voice or via an interface which is similar to a web based query field of a browser. Subsequently the system parses resp. interprets the input of the user and searches for corresponding data in the database. The results are listed into the GUI as shown in Figure 5 and the user is able to select a content element and to listen to the associated audio file. Although our prototype utilizes underlying textual representation and employs text-based information retrieval techniques, this mechanism is hidden to a great extent from the user, only snippets are displayed. The results could be presented either as spoken response utilizing text-to-speech technologies.

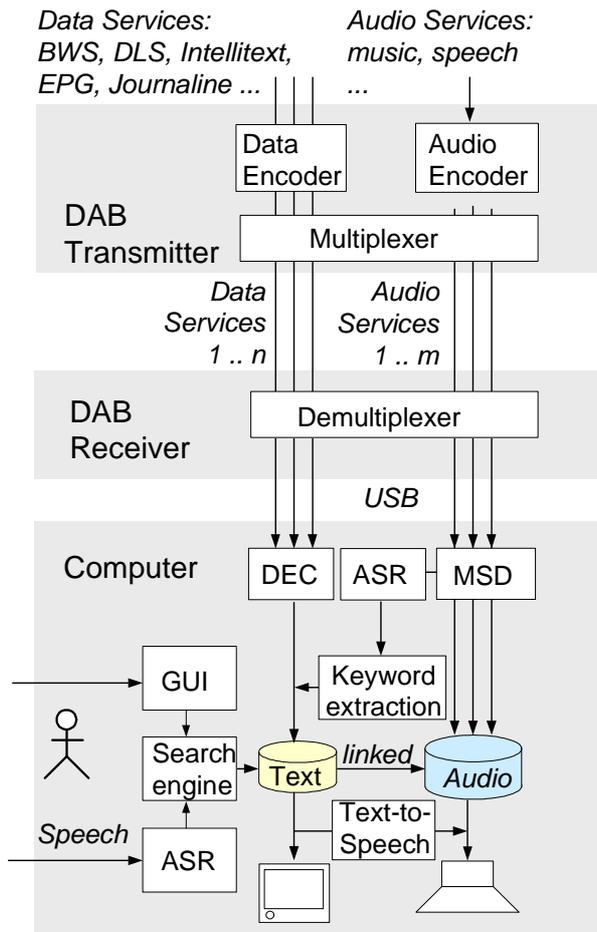


Figure 4 - Overview of the system (DEC: Decoder)



Figure 5 - Design of a GUI and results for the query "Berlin"

Doubtless, through continuous improvements in large-vocabulary ASR systems, recognition of broadcast materials has now become possible in real-time. Though, problems such as the use of abbreviations, elements of foreign languages, and acoustic interferences are complicating the recognition process significantly. The goal is not to obtain a correct text, but simply to gather enough semantic information to generate a characterization that the system can employ to find relevant content. The interface offers primary the user the original audible content from recordings, because audio is a much richer medium of communication.

The main result is a holistic approach to store speech-based audio content and – for the first time – the entire range of functionalities contained in DAB data services. All these information units are combined with the capability to search for specific content on the base of an efficient MSD and a speech retrieval system for spoken content as well.

Furthermore, the development of a speech-based user interface enables users to operate the functional scope of a digital radio solely by vocal commands. Most importantly, this development was focused on a dialogue system that conforms to the concepts of human-centred dialogues in an user-oriented manner, see Figure 6. In this way a sophisticated radio may offer with modest ambition similar favorable characteristics referred to multimodal content like internet services do for text.

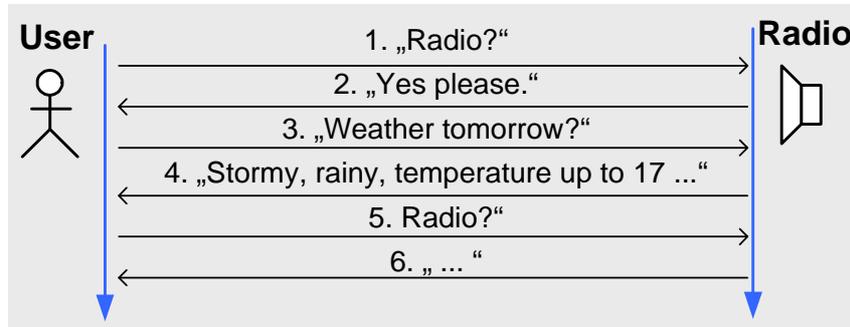


Figure 6 - Example of spoken dialogue about weather

The extension of a digital radio with the possibility of speech-based interaction in combination with multichannel monitoring, music-to-speech discrimination (MSD) and a memory function significantly adds value to a regular radio by promoting the evolution towards an embedded device providing the following functionalities:

- interactive search for content from audio and data information sources,
- speech-based output of content,
- completely controllable by spoken commands,
- personalisation and basic personification.

The current limitations of the introduced system have to be handled by more efficient speech recognizer, sophisticated semantic retrieval algorithms, and a higher degree of parallel processing.

PODCAST CONTENT BROADCAST ON DIGITAL RADIO

The podcast-broadcast hybridisation is a new example of the large number of possible use cases of transportation technologies beyond the Internet: traditional linear broadcasting, time shifted delivery, content on demand, and various cellular links such as point-to-point resp. point-to-multipoint. Due to the availability of push and pull technology on one device and the possibility of shifted media consumption user habits changed remarkable. The emerging trend of podcasting has shown the request for highly personalised and on-demand media on PCs and mobile media players. Podcasting augments the importance of niche entertainment as well thematic special content with a strong and unique style. An alternative solution for time shifted media consumption are additional devices: personal or digital video recorders (PVR, DVR) on the base of EPGs.

Our approach develops a technology how the podcasting concept can be expanded for non-web-based applications making them available to broadcast media with the tendency to generalising the approach (8). The two main objectives are: At first general methods of linking and distributing podcast content in broadcast and cellular systems are discussed. Secondly, the embedding of podcasting in broadcast systems is demonstrated in order to increase the comfort of media consumption, to lower the costs of networking, and to avoid load peaks while delivering podcast content on the Internet.

The solution uses a content description with the powerful tools EPG and TV-Anytime (TVA), they consider adequately the needs of radio broadcasting. It is indicated by own analyses that for the audio domain a systematic and detailed description of contents is

largely possible although national conversion and consolidation is required. TVA is widely used for EPG applications, both are based on a XML description. Only a part of the comprehensive standard was used for own test purposes: *genreType*, *IntentionCS*, and *ContentCS*. The so prepared metadata files are combined with the appropriate RSS formats. They describe fundamental characteristics of the subsequently by MOT transferred podcast files, see Figure 7.

It is necessary to design and implement a suitable user interface for activating the stored podcasts because most of the receivers are not equipped with large display and the familiar mouse/pointer/icon paradigm. The costs of the XML description are significant compared with simple ID3 tags. However, the schemes are machine-readable without difficulty, expandable, and possess a fine-grained hierarchical description. A solution on the basis of DL is possible as well. The prototype software shows, that integration of podcast content into digital broadcast environments as DVB or DAB/DMB is possible on the base of available standards and techniques.

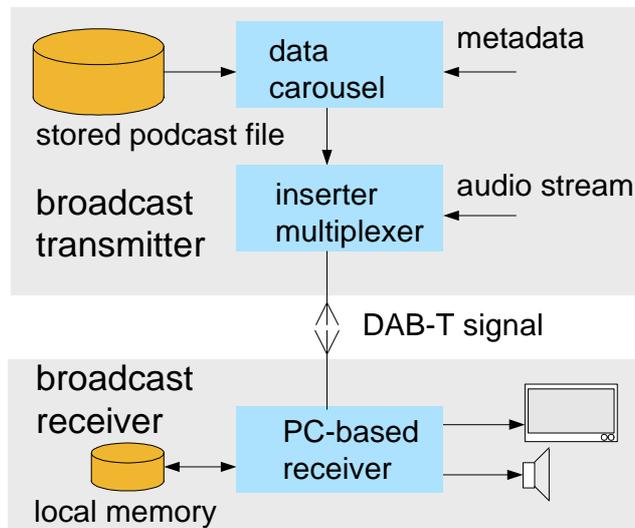


Figure 7 - DAB podcasting setup

The user interface of the prototypical software is a first draft and offers a lot of room for improvements regarding design and user interaction. The integration of podcast content into broadcast and cellular systems is very promising, because the metadata and media content is already produced and available. With a minimum amount of editorial work a broadcaster can enrich linear TV and radio content with on-demand media and adapt to changing consumer habits. Podcasts are also appropriate for using collaborative filtering methods to find unknown but interesting and well-rated content in an ad-hoc environment as well. It can be promising to apply methods and tools of ubiquitous computing, computer supported cooperative work and other branches to improve the social character of podcasting in public places supported by broadcast and cellular systems.

CONCLUSIONS AND OUTLOOK

The combination of above-named elements represents a fundamental modernisation of conventional usage patterns with radio devices. Hence, the development of radio usage from passive hearing towards listening and an interactive and individual dialogue is strongly supported. The improved functionalities render the radio to be an appropriate device to satisfy much more multifarious necessities for information than before. As a result users are capable of selecting desired audio contents more systematically, with higher concentration and with higher density of information from future, current, and even past programmes. These presented functionalities and listener focused benefits were perceived very positively by test users and became generally accepted under the condition

that the usage of the radio would be a fairly simple task. However, radio should not lose its role in this magic audio world, at where you can have your own privacy, at where you can listen with simple devices to useful and entertaining information everywhere. One should always bear in mind that the value added messages or images should always be supplementary to the content and the audio effect of a radio programme (9).

As a consequence the cost-benefit-ratio is to discuss. Most of the results related to new services are reasonable. At the present time, a speech-based data base for mobile devices is possible but makes not good economic sense in every aspect. Furthermore, there are several requests for the near future, among others:

- Broadcasters should send reliable and proper metadata, and EPGs as well,
- DAB must become a reality in mobile phones,
- Receivers should be equipped with color graphic displays resp. touch screens,
- Interfaces should consider the mood of the user as well.

We believe that digital radio has yet a huge development potential for numerous and still unknown applications especially as an integrated part of cell phones, navigation systems, personal information managers, and enhanced car radios etc.

ACKNOWLEDGEMENTS

Many thanks to all the contributors for collaboration, advice, and information that has made this article possible: C. Bräutigam, A. Eiselt, M. Neumann, C. Rotzoll, S. Schmolke, H. Tönnies, and B. Zeller. Special thanks to Thomas Kreuz of Radio Antenne Thüringen for the permanent support with the experimental usage of data services.

REFERENCES

- (1) Marks, B.: HuMIDAB: A user interface for digital broadcasting. EBU Technical Review Winter 1998.
- (2) Kozamernik, F.: DAB – from Digital Radio towards Mobile Multimedia. EBU Technical Review January 2004.
- (3) Herrmann, F.; Erismann, L.; Prosch, M.: The Evolution of DAB. EBU Technical Review July 2007.
- (4) Crawford, C.; King-Smith, T.: Digital Radio – a receiver manufacturer's viewpoint. EBU Technical Review 2008-Q2.
- (5) Schatter, G.; Bräutigam, C.; Neumann, M.: Personal Digital Audio Recording via DAB. Enhanced Radio as Interface. 7th Workshop Digital Broadcasting. Fraunhofer IIS Erlangen/Germany, September 14-15, 2006, pp. 146-153.
- (6) Schatter, G.; Zeller, B.: Design and Implementation of an Adaptive Digital Radio DAB using Content Personalization on the Basis of Standards. IEEE Transactions on Consumer Electronics. 53 (2007) November, No. 4, pp. 1353-1361.
- (7) Schatter, G.; Eiselt, A.; Zeller, B.: A multichannel monitoring Digital Radio DAB utilizing a memory function and verbal queries to search for audio and data content. IEEE Transactions on Consumer Electronics. 54 (2008) No. 3, August, pp. 1082-1090.
- (8) Schatter, G.; Rotzoll, C.: Subscription of Digital Broadcast Content Utilizing Syndication Feeds Mechanism. IEEE Transactions on Consumer Electronics. 54 (2008) No. 4, November, pp. 2015-2022.
- (9) Han, Y: Possibilities of Value-added Digital Radio Broadcasting. Asia Pacific Broadcasting Union Conferences 2007.