RadioActive: enabling mobile-based audio forums

Aaron Zinman
Sociable Media Group
MIT Media Lab
20 Ames St, Cambridge, MA 02138
azinman@media.mit.edu

Judith Donath
Sociable Media Group
MIT Media Lab
20 Ames St, Cambridge, MA 02138
judith@media.mit.edu

ABSTRACT
In this paper we describe a new kind of asynchronous discussion using audio instead of text over mobile devices. The format of the discussion surrounds both real content and Usenet-style topics. Users engage in public discussion using a combination of spatial and audio navigation.

Author Keywords
Asynchronous communication, audio communication, audio navigation, public discussion, virtual spaces

ACM Classification Keywords
H. Information Systems, H.4 INFORMATION SYSTEMS APPLICATIONS, H.4.3 Communications Applications

INTRODUCTION

Internet-based conversations typically use text as their primary medium. While this works well for many types of discussions using desktop computers, text is particularly difficult in mobile phones. Small keyboards restrict input, and the screen resolution limits legible text space. We believe that asynchronous audio conversation is a strong candidate for large-scale mobile-based discussions. Audio allows a more natural and engaging format, especially when coupled against a small form factor. However, such radical change in medium requires a departure from traditional navigational models.

RadioActive's primary platform is mobile devices. Such devices allow users to engage in public discussion anywhere. This is particularly important with changes in content delivery such as Podcasting. Podcasting is a technique of time-shifting content so that a user's computer downloads content from subscribed channels ahead of time, syncing with devices such as Apple's iPod. The user is then free to listen or view the content at any future point. By targeting future mobile devices, users could engage in public discussion surrounding real content in addition to Usenet-style topics. The previously passive devices then become interactive.

The system we are proposing uses many of the familiar features of current discussions. Threaded discussions contain posts by individual authors categorized by arbitrary topics, have *n* replies to replies, a subject, and a body. The threads are archived, creating persistent dialogues. However, since we are replacing text input with audio, the presentation of a thread must change. Messages are represented graphically inside a topic space. A "radio" mode enables the platform to be similar to dynamic talk radio without requiring input. The user may also chose to navigate the conversations using directional buttons common across mobile devices.

RELATED WORK

Arons's Hyperspeech [1] and Resnick's Hypervoice [5] were systems designed to navigate hyperlinked audio data. Both were the first major research systems designed to demonstrate how content could be delivered using a telephone-based interface. The focus of these works was enabling applications with large databases to be accessed and controlled remotely without a real computer. Since a mobile phone provides both a display and processing power, we can now overcome the problems of standard telephones with modern affordances in our interface.

TattleTrail [4] allows users to share a conversation space using mobile devices. It is primarily designed for enabling small work groups to converse both synchronously and asynchronously. Since archived messages are linear, it is not appropriate for large-scale discussions of arbitrary size.

Espaces2 [6] was an attempt to provide an audio-only environment using an "acoustic bubble" metaphor. Users could navigate hyperlinked audio data using 3D spatialized audio and audio icons [3] to provide awareness to potential nodes. Thus Espaces2 is attempting to create, maintain, and subtly manipulate a cognitive representational map of the data space using only audio. Three-dimensional audio performs well at informing us of relative distances of nearby objects, however it is difficult to give an overview of a large abstract space. We propose to augment audio content using spatial navigation.

DESIGN

When users log into the system, they are presented with replies to threads they have participated in. This encourages discussions to continue without requiring the extra work Slashdot or Usenet requires to discover replies. The user can choose to navigate through a hierarchy of topics of interest and saved content streams. Once a topic is selected, they are presented with either a series of threads or the selected content. Each post is represented by a simple geometric shape, using color and size to indicate message length, age, and read/unread status. Users navigate the hierarchy of posts using the arrow controls standard across

modern mobile devices. If streamed content was chosen, audio icons are overlaid to indicate the presence of threads branching from that point. Pressing down would go into the thread represented by the last heard audio icon. If the user is not listening to content, a short "subject" audio clip plays as the user hovers over individual posts. Subjects are created after an author finishes his/her post. This familiar message format allows users to quickly skim around text-less posts without requiring the user to listen or jump around potentially lengthy posts.

To give a more natural conversation or talk radio feel, RadioActive can continue playing the selected thread without the need for interaction. With each post the user would hear the name of the author, the subject, and then the message body. Users can use the left or right buttons to speed playback using SpeechSkimmer's audio-compression algorithms [2], while up and down navigate across the thread hierarchy. After the message has been fully played, the user has a chance to reply before continuing to the next post. At the end of the thread, the user would hear a musical transition before automatically going into the next thread in the spirit of National Public Radio. This helps alert the user that the thread has ended and the next message has no semantic relationship to the previous messages.

DISCUSSION

One difficult issue is meaningfully using space for navigation. Within a topic, what does it mean to present a thread to the left of another within a virtual space? If the space is separated into sub-topics, the individual posts would need to be classified into their proper section. This extra work is both a burden for the user and difficult as many posts would fit into multiple categories. To help alleviate this problem, RadioActive forces the start of each thread to grow downward from a one-dimensional line at the top of the screen. The width between individual threads is constant. Placement on the line can be chronological, or chosen by filters the user may interactively select. For example, the user might order messages by number of responses, audience size, or ratings. Since messages may contain many branches (taking up arbitrary width), only the first post is shown at the top. A line descends from the initial post indicating replies with line thickness to represent the number of messages in the thread. Upon selecting a post, all other threads disappear in order to guarantee enough room for display.

Another issue is the level of interactivity required for radio mode. Real talk radio is linear, requiring callers to respond to recent remarks rather than those from earlier in the program. RadioActive uses archived messages on a computer to enable more powerful non-linear discussions. However, it is difficult to know which path to choose when playing a thread with many branches. If the heuristic is breadth-first, it may be difficult to mentally retrace the path without visual aid. This is because the user is required to remember at level $n+1$ which message is being addressed from level n. Using depth-first traversal relieves this problem, but will pass related messages, making large semantic leaps as it transitions from the end of one path to the middle of another. One potential solution is to use the number of replies as a heuristic. Audio icons indicating alternative paths could alert the user of other paths they may interactively choose.

FUTURE WORK

As a new discussion medium, RadioActive can be augmented with modern ideas such as collaborative filtering, hyperlinked content, user-based moderation, and recommendation systems. Discussions can be encouraged to have faster responses by sending participants SMS messages as the thread grows. If the responses are fast enough, the asynchronous messaging could automatically switch into synchronous discussion in a similar manner to TattleTrail.

CONCLUSION

RadioActive, while still in the design phase, could create a novel environment for public discourse. Using voice, many characteristics of personality and gesture can be communicated in a way that is not possible using text. Furthermore, encouraging public discourse on universal content helps the democratic process.

ACKNOWLEDGMENTS

We would like to thank CHI and Workshop volunteers for putting on the conference. We also thank the Sociable Media group for their helpful input as well at the MIT Media Lab for providing its wonderful resources.

REFERENCES

1. Arons, B. Hyperspeech: Navigating in speech-only hypermedia. In Proceedings of Hypertext '91, pp. 133-146. ACM, 1991.
2. Arons, B. SpeechSkimmer: A system for interactively skimming recorded speech. *ACM Transactions on Computer-Human Interaction*, 4(1):3-38. March 1997.
3. Blattner M., Sumikawa D., Greenberg R. Earcons and icons: Their structure and common design principles. Human Computer Interaction 4, 1 (1989), 11-44.
4. Kim J.S. TattleTrail: An Archiving Voice Chat System for Mobile Users Over Internet Protocol. MIT M.Eng Thesis, May 2002.
5. Resnick P. HyperVoice: A Phone-Based CSCW Platform. In *CSCW '92: Conference of Computer-Supported Cooperative Work* (Toronto, 1992), ACM, pp. 218-225.
6. Sawhney N., Murphy A. ESPACE 2: An Experimental HyperAudio Environment. Proceedings of CHI '96. April 1996.