Designing and Evaluating Voice-Based Virtual Communities

Neil Patel

Stanford University

353 Serra Mall, Gates 382

Stanford, CA 94305

neilp@cs.stanford.edu

Figure 1. A Farmer in Gujarat, India, using Avaaj Otalo, a voice-based virtual community for agricultural information.

Abstract

Voice-based virtual communities offer new possibilities for information dissemination and sharing for billions of users who lack access to Internet-connected PCs. As interaction is only through voice, these systems are subject to different design constraints than web-based social software. Through lab experiments and fieldwork, we have identified three key design challenges for voice-based virtual communities: supporting threaded conversations; indexing and searching content; and managing identity. We discuss each of these issues and propose approaches to address them. We also present plans to evaluate the impact of voice-based virtual communities on knowledge access and sharing in rural India.

Keywords

Voice user interface, IVR, social media, forum, India, rural development, agriculture, literacy, ICTD

ACM Classification Keywords

H.5.2 User Interfaces: Voice I/O

General Terms

Human Factors, Experimentation

```
AO:      Welcome to Avaaj Otalo! You can get to
         information by dialing the number. To ask
         a question dial 1; to listen to
         announcements, dial 2; to listen to the
         radio program,dial 3.

User: (dials 1)

AO:      OK, you want to ask a question. To record
         your own question, press 1. To listen to the
         questions and answers of other farmer
         friends, press 2.'.

User:   (dials 1)

AO:    OK, you want to record a question. Please
       say
       your question slowly and clearly after the
       beep
.
User:  How can I protect my cotton crop from mealy
       bugs?
```

Figure 2. A sample interaction with Avaaj Otalo.

Introduction

Voice-based virtual communities are voice-only analogs to virtual communities on the Web (*e.g.*, forums, wikis, and social networks). Virtual communities accessible through voice can connect users who cannot afford or otherwise access Internet-connected PCs; all a user needs is a mobile phone. Voice is appealing as a communications medium for users with low literacy or who speak a regional language. In our work we have designed and deployed *Avaaj Otalo* (literally, "voice stoop"), a voice-based question and answer forum for farmers in rural India to share advice and experiences on agricultural topics. In a 6-month pilot of the system, Avaaj Otalo (Figure 2) became a lively and informational social space for its 50 users [1].

Much prior research on virtual communities has focused on web-based systems [2]. Voice has different properties than the Web. For example, voice content must be scanned sequentially, making it more time consuming to consume, search, and browse content. Prior research on voice-based social media has been concerned with simple voice bulletin boards or synchronous phone lines, with limited support for threaded conversations or searching archived content [3] [4].

We have explored the design of voice-based social media through lab experiments and fieldwork in India. In the following sections, we identify three key design challenges for voice-based virtual communities that have emerged: supporting threaded conversations, indexing and searching content, and managing identity. We conclude with our plans to evaluate the impact of voice-based virtual communities on knowledge access and sharing in rural India.

Supporting Threaded Conversations

On the Web, virtual communities provide the ability to scan, search, and participate in ongoing conversations through visual feedback. For example, on forums, conversations are organized as threads to separate parallel discussions, and on discussion lists, users respond to a previous comment by inserting their text next to the original. The ability to lay out content spatially helps organize the conversation. Voice-only content lacks a shared visual representation.

For Avaaj Otalo, threaded discussions were originally omitted to reduce interface complexity for novice users. As a next step, Avaaj Otalo will be extended to support threaded conversations, including non-linear threads and linking replies to content within a recording. Specific components include presentation of position or state within the thread, navigation to jump to different parts of the thread, and commands to record and playback replies embedded within recordings. Study of posting on web forums has shown that long threads with complex social dynamics are not uncommon [2]. I hypothesize that threads on a voice-based forum will tend toward much simpler, linear patterns, as voice content will encourage users to communicate in a style closer to face-to-face conversation. To test this hypothesis, we will study conversations on Avaaj Otalo over several months in the coming year, supplementing results of the data analysis with interviews to see how easy users found it to engage in conversations.

Indexing and Searching Content

Indexing of the voice content in the virtual space is necessary for later retrieval. It can be done automatically using speech recognition technology, by third-party transcribers, or having users tag content themselves. A combination of approaches may also be employed. However, all but the last require considerable extra cost to employ. This dissertation will explore techniques for indexing content through crowd-sourcing. On our live field deployment, we will compare various approaches for eliciting users to supply content tags, measuring quantity, accuracy, and intrusiveness (through user interviews). We will determine the best place in the interaction for tagging, and compare menu-based selection, audio tags, and collaborative game-like approaches (like ESP) for generating tags.

Once indexed, content may be searched or filtered for easy access. On the web, search is dominated by query-based interfaces. For voice-based virtual communities, an alternative approach is directory traversal. Organizing content into hierarchical directories allows a user to find content through traversal of touchtone menus, avoiding speech recognition and its associated constraints (misrecognitions, sensitivity to background noise, cost for development). To minimize long and complex traversals, directories and content can be presented based on recommendations gathered through traffic analysis, content rating systems, user profiles, and other techniques. Our prior work evaluating speech input to voice interfaces suggests that our mostly novice users in rural India have a low patience threshold for recognition-related errors [5]. Others have shown that novice web users prefer directory traversal over query-based search [6]. I hypothesize that the same will hold for voice-based searching of content, and I will test the hypothesis through lab experiments to be run in the coming year.

Managing Identity

With Avaaj Otalo, some users expressed hesitation to respond to others' questions out of fear for giving incorrect information or not having sufficient authority to address the problem. This seemed at least partially due to protectiveness of personal reputation amongst other users. While encouraging users to post "best guesses" can be done through establishing the appropriate culture, another approach is to provide ways to post anonymously. We will compare two approaches for addressing identity issues on Avaaj Otalo: anonymous posting and group posting. With anonymous posting, users can post content through a neutral, synthesized voice to mask personal identity. Group posting will allow users to form groups in which content can be internally vetted before being posted to the live forum under the name of the group. I hypothesize that group posting will be a more attractive option for Avaaj Otalo users, who are accustomed to group-based decision making.

Evaluating Impact

Beyond researching the design of voice-based virtual communities, we are interested in the impact they have on information access and sharing in rural developing regions. Drawing from developmental economics (particularly studies on information diffusion [7]), we plan to perform an impact analysis of voice-based virtual communities in rural India in the coming year. We will have two measures: knowledge retention, and adoption of new techniques. Initial pilot studies will establish a baseline for the measures and determine a suitable sample size. We will then conduct randomized

field trials comparing communities with access to the technology against a control group.

Related Work

This work was originally inspired by projects in India that use media such as the Web, video, and mobile phones to deliver information to rural communities [8] [9]. Prior work in voice social media demonstrates the viability of a voice-only social space [3] and offer initial directions in design and implementation of voice bulletin boards [10] [4]. Other work on navigating and skimming speech [11] [12] will be drawn upon in our design work going forward.

Conclusion

Voice-based virtual communities are a promising tool to connect previously disconnected communities around the world. Through careful design, they can offer a viable large-scale alternative to web social software in developing regions. This dissertation identifies and addresses a few of the most challenging aspects of designing voice-based virtual communities, complementing the design work with a real-world assessment of their impact from a development perspective.

Acknowledgements

References

[1] Patel, N., Chittamuru, D., Jain, A., Rajput, N., Dave, P. and Parikh, T. S. Avaaj Otalo - A Field Study of an Interactive Voice Forum for Small Farmers in Rural India. In Proceedings of the Proceedings of the 28th international conference on Human factors in computing systems (Atlanta, GA, USA, 2010). ACM.

[2] Turner, T. C., Smith, M. A., Fisher, D., and Welser, H. T. Picturing Usenet: Mapping computer-mediated collective action. Journal of Computer-Mediated Communication2005).

[3] Ackerman, M. S., Starr, B., Hindus, D. and Mainwaring, S. D. Hanging on the 'Wire: a field study of an audio-only media space. ACM Trans. Comput.-Hum. Interact., 4, 1 1997), 39-66.

[4] Resnick, P. Phone-based CSCW: tools and trials. ACM Transactions on Information Systems1993).

[5] Patel, N., Agarwal, S., Rajput, N., Nanavati, A., Dave, P. and Parikh, T. S. A comparative study of speech and dialed input voice interfaces in rural India. In Proceedings of the Proceedings of the 27th international conference on Human factors in computing systems (Boston, MA, USA, 2009). ACM, [insert City of Publication],[insert 2009 of Publication].

[6] Annabel Pollock, A. H. What's Wrong with Internet Searching. In Proceedings of the Designing for the Web: Empirical Studies (Microsoft Corporate Headquarters, Redmond, Washington, 1997).

[7] Erin Godtland, E. S., Alain de Janvry, Rinku Murgai, Oscar Ortiz The Impact of Farmer-Field-Schools on Knowledge and Productivity: A Study of Potato Farmers in the Peruvian Andes. Economic Development and Cultural Change2004).

[8] Rikin Gandhi, R. V., Kentaro Toyama, Vanaja Ramprasad Digital Green: Participatory Video and Mediated Instruction for Agricultural Extension. Information Technologies and International Development2007).

[9] Ramamritham, K., Bahuman, A. and Duttagupta, S. aAqua: a database-backended multilingual, multimedia community forum. In Proceedings of the Proceedings of the 2006 ACM SIGMOD international conference on Management of data (Chicago, IL, USA, 2006). ACM.

[10] Zinman, A. RadioActive: Enabling large-scale asynchronous audio discussions on mobile devices. 2006.

[11] Arons, B. SpeechSkimmer: a system for interactively skimming recorded speech. ACM Trans. Comput.-Hum. Interact., 4, 1 1997), 3-38.

[12] Arons, B. Hyperspeech: navigating in speech-only hypermedia. In Proceedings of the Proceedings of the third annual ACM conference on Hypertext (San Antonio, Texas, United States, 1991). ACM.