

An Audio Wiki for Building Local Repositories of Knowledge in the Developing World

Pratik Kotkar

Indian Institute of Technology, Guwahati
Guwahati, India
Email: pratik@iitg.ernet.in

William Thies

Massachusetts Institute of Technology
Cambridge, USA
Email: thies@mit.edu

Saman Amarasinghe

Massachusetts Institute of Technology
Cambridge, USA
Email: saman@mit.edu

Abstract—Mobile phone technology has spread across the globe, especially in developing countries where the usage of mobile phones far exceeds that of desktop computers. Here we describe the idea of an Audio Wiki, a localized repository of information that can be accessed using the mobile phone. We aim to create systems that provide users with information in the form of audio recordings of other users' experiences and knowledge on various topics. Herein we present the design and implementation of such a system empowering the Audio Wiki. The information server can be accessed by any telephone and the user can use speech as well as the keypad to navigate menus, search and edit information on specific topics there in. We envision a broad array of applications for the audio wiki, spanning diverse areas such as agriculture, health, government, and local entrepreneurship.

I. INTRODUCTION

While in India an increasing use of computers for data processing and use of the Internet for information purposes has been seen, these technologies are mostly restricted to urban areas. In rural areas the barriers for computer usage and Internet access are mostly economic (the cost of computers and network access often exceeds the average wage). However, the deeper problem is sociological in nature: the default languages and user interfaces of the Internet are a mismatch for the developing world. Over 60% of today's Web pages are in English, while less than 10% of the world's population speaks English as their native tongue. Even if content was available in the local language, in written form it would remain inaccessible to the illiterate population - 40% of India's population and over a billion people worldwide. Languages and literacy are even more significant in preventing local populations from creating and sharing their own content.

Despite these challenges, mobile phones offer an innovative way to overcome the hurdles in a developing world. Though cell phones are inferior to existing desktop computers for high-tech applications, in the developing world cell phones are leap-frogging ahead of computers as the primary personal computing device. As illustrated in Fig. 1, cell phone penetration has far surpassed computer penetration in the developing world. In India, the margin stands at approximately 5x while the same figure is up to 25x in Ghana. Also, worldwide there are 1.4 billion people that have access to a cell phone but not to a computer. In 2004, the global mobile phone penetration was 32% and the growth of the mobile phone subscribers is fast, especially in developing countries [1]. In 2008, the global

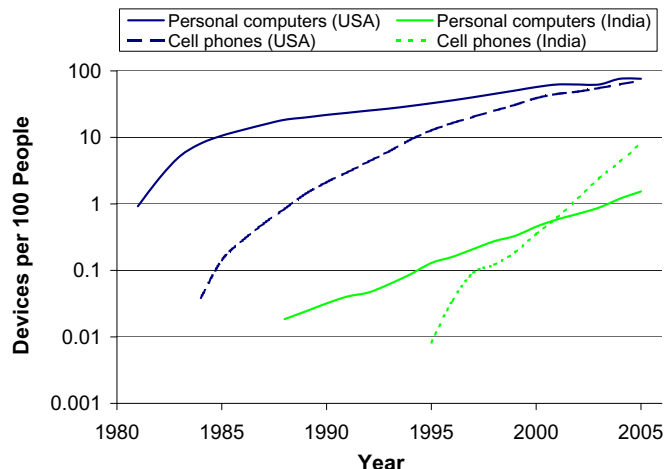


Fig. 1. Penetration rates for personal computers and cell phones in the United States and India. While both devices are ubiquitous in the United States, cell phone penetration exceeds computer penetration by 5x in India.

mobile phone penetration is expected to be more than 50%.

This gives impetus for the development of a voice-based interface for the phones to overcome the language and literacy barriers traditionally preventing scalable information services from reaching rural communities.

II. RELATED WORKS

The MobileED group has developed an audio wiki and evaluated its use in schools in South Africa [2], [3]. In their system, users send an SMS query to the server, which calls the user back and uses text-to-speech technology to dictate existing Wikipedia content. Users can also annotate the content with their own spoken tags. Our system (which was conceived independently) aims to apply a similar concept to the Indian context, targeting a broad set of users in agriculture and health in addition to education. We also aim to develop novel technologies for speech and user interfaces that will be widely applicable across audio wiki projects.

When Internet access is available, Wikipedia [4] is popular for allowing users to add, view and edit information regarding specific topics. For this concept to be extended usefully to developing regions the communication medium must be appropriate and a speech-oriented interface would make the system accessible to the illiterate population as well. While the Spoken

Wikipedia Project [5] aims to translate existing articles into speech, they do not support a speech-driven editing system.

The World Wide Telecom Web (WWTW) [6] aims to enable users to create “voice sites” that are linked together analogously to Internet websites, except that all content is spoken and is accessed over the phone. There are two main differences compared to our audio wiki. First, the audio wiki is a centralized information store (potentially with many local access numbers), while the WWTW is distributed across many independent hosts and requires a new “hyperspeech transfer protocol” to redirect calls between different sites [7]. Secondly, proposed sites on the WWTW are organized by user (analogous to a home page), while sites on the audio wiki are organized by topic (with many users contributing to a single site). Coordinating edits from multiple users will be a unique challenge and capability of the audio wiki.

The PeopleNet [8] project also aims to satisfy local information needs using mobile devices, but connects pairs of matching queries (e.g., buy/sell) in a given locale rather than accumulating a persistent repository of knowledge.

III. SYSTEM SETUP

Our system is based on a central server which is interfaced with a speech recognition engine, an information storage database and an audio recording and playback system. The system allows users to listen to content on any topic added by other users as well as edit the content or leave their own views about the topic. The audio recordings are stored on the server side hence the only requirement on the user side is a telephone.

A. Asterisk server

The server is built on Asterisk, an open source IPBX [9], [10]. Asterisk has inbuilt functionality that allows easy interface to other utility blocks. It allows interface to the fixed land line telephones (touch tone type) as well as mobile phones by use of Digium hardware cards [9].

B. Speech Recognition

As all of the Wiki content will be spoken language, there is an immense need to deploy novel speech technologies. Speech recognition in the form of “keyword matching” will be required for the user interface.

Also, as the audio samples will be spoken into a low-end cell phone with varying amounts of ambient noise, the server will employ spectral subtraction techniques to remove noise, esp. additive Gaussian noise. Volume equalizers are also being tested to provide a uniform playback level for all speakers.

C. User interface

The primary interface involved is that of a mobile phone keypad to the server. Since voice-driven menus may be tedious, the correct amount of voice and keypad usage for navigation will be tested to maximize user efficiency. Currently we have implemented a scheme analogous to an adaptive Huffman encoding for input symbols in which the most commonly

accessed topics are directed to the top level menu. This method was found to increase user efficiency in navigating to a particular topic. User interfaces can also be customized using caller identification to minimize audio prompting for regular users.

IV. APPLICATIONS

We envision a broad array of immediate, high-impact applications for the audio wiki. In the developing world, farmers can call in and describe their experience with various crops, diseases, and pesticides in the current year. Health workers can share up-to-date and locally-tuned medical information. Craftsmen and artisans can trade instructions and best practices. Entrepreneurs can describe the scope and operations of their local business. All this can only be possible on a localized information sharing system in which the user is empowered to add, edit and listen to content.

V. CURRENT AND FUTURE WORK PLANS

A. Current Work

Currently a prototype for the Audio Wiki has been built in Boston, USA. It has been developed currently to suit more of the developed country requirements. The prototype will be put to test to make the user interfacing efficient before deploying in a developing region.

B. Future Directions

After a rigorous testing of the prototype, we plan to deploy a similar system in India. The system will be designed in a local language. We are also simultaneously working on training available speech recognition kits to languages and accents used in the particular region of deployment of the Audio Wiki.

In the long term, we aim to employ language-independent speech recognition [11] to allow users to search audio content using audio keywords, independent of the language used. This will enable the wiki to apply uniformly across regions without needing to specialize it to a given language or accent.

REFERENCES

- [1] “ITU report,” <http://itu.int/newsroom/wtdc/2006/stats/index.html>.
- [2] T. Leinonen, E. Sari, and F. Aucamp, “Audio wiki for mobile communities: Information system for the rest of us,” in *ACM Conference on Human Computer Interaction with Mobile Devices and Services*, 2006.
- [3] M. Ford and T. Leinonen, “MobilED - A Mobile Tools & Services Platform for Formal & Informal Learning,” in *mLearn*, 2006.
- [4] “Wikipedia,” <http://www.wikipedia.org/>.
- [5] “Spoken wikipedia project,” http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Spoken_Wikipedia.
- [6] A. Kumar, N. Rajput, D. Chakraborty, S. K. Agarwal, and A. A. Nanavati, “WWTW : The World Wide Telecom Web,” in *Workshop on Networked Systems for Developing Regions*, 2007.
- [7] S. K. Agarwal, D. Chakraborty, A. Kumar, A. A. Nanavati, and N. Rajput, “HSTP - Hyperspeech Transfer Protocol,” in *Hypertext and Hypermedia*, 2007.
- [8] M. Motani, V. Srinivasan, and P. S. Nuggehalli, “PeopleNet: Engineering A Wireless Virtual Social Network,” in *MobiCom*, 2005.
- [9] J. Meggelen, J. Smith, and L. Madsen, *Asterisk - the Future of Telephony*. O’Reilly Media, 2005.
- [10] “Asterisk - Open Source IPBX,” <http://www.asterisk.org>.
- [11] M. Viikki, I. Kiss, and J. Tian, “Speaker and Language-Independent Speech Recognition in Mobile Systems,” in *Acoustics, Speech, and Signal Processing*, 2001.