Demo: Audio QR Codes for Voice Service Position Sharing

Jennifer Pearson,[1] **Simon Robinson,**[1] **Nitendra Rajput,**[2] **Matt Jones,**[1] **Amit Nanavati,**[2]

[1] FIT Lab, Swansea University, SA2 8PP, UK
{ j.pearson, s.n.w.robinson, matt.jones } @swansea.ac.uk

[2] IBM India Research Lab, Vasant Kunj New Delhi, 110070, India
{ namit, rnitendra } @in.ibm.com

ABSTRACT

In this demonstration, we introduce ACQR (pronounced 'aqua') – audio-based sharing for the Spoken Web. By using a series of DTMF tones, we allow direct synchronisation of exact positions within the voice service hierarchy without the need for high end devices, or an internet connection.

This paper abstracts and summarises a full paper which describes the technique in more detail [3].

INTRODUCTION

In many underdeveloped rural regions, access to technology, particularly the internet, is scarce. Poor textual literacy, low income and intermittent power or data connections all contribute to the lack of technology availability in these areas. In recent years, however, there has been a high uptake of basic cellular phones [1] – devices that are the main focus of audio-based information services such as the Spoken Web [2]. Their purpose, is to create voice-driven alternatives to the internet that provide users of low-end phones with access to cheap, audio-based information without the need for textual literacy or an internet connection.

The Spoken Web is hierarchical in structure, similar in design to the Interactive Voice Response (IVR) systems commonly used for customer service telephone enquiries. To access information, users call via the public telecom network and navigate to specific locations within the hierarchy by using the phone's keypad. Prompts and previews help the user to traverse the menu structure and, after a series of these menus, find the content they are searching for.

Sharing this information with other people can be difficult, however. While internet addresses can be shared in a variety of ways—ranging from emailing a link, to searching for keywords, to scanning a QR code—sharing within voice services is far more difficult. The diversity in website sharing methods is possible because of the use of a textual address, and a well-indexed population of websites. In contrast, voice services are audio in nature and recorded in many different languages, which means that searchable indexes are not usually available. This situation makes novel mechanisms to enable sharing within these services particularly important to help the spread of useful content amongst users.

ACQR

To aid in the sharing of voice service positions, we developed an audio-based system that uses DTMF tones to synchronise people's positions within the Spoken Web hierarchy. Figure 1 illustrates the process of sharing from one phone to another –

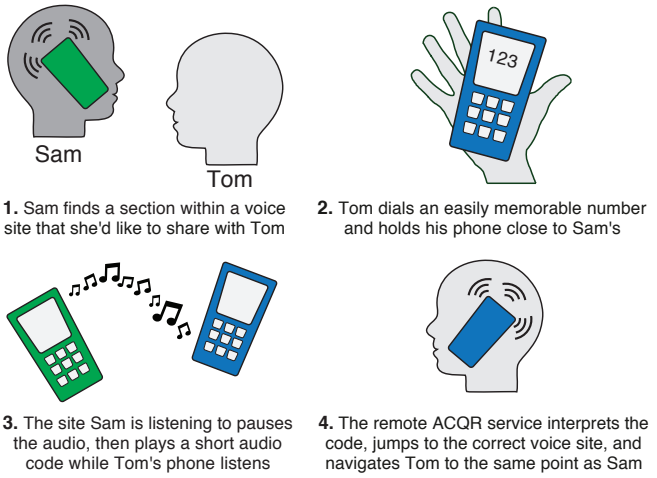

Figure 1. Audio sharing. Callers listening to a voice service can share their precise position with others by playing them an audio code. After sharing, both users can navigate within the original voice service from the same position. No additional hardware or infrastructure is required, and both the sharer and the recipient may use standard, low-end phones.

callers simply hold their phones together to synchronise; an audio tone from the sender is recognised at the server side through the receiving phone's microphone. Receiving users are automatically navigated to the same spot within the voice service hierarchy and are then able to browse independently of the original sharer. The process works as follows:

Sending: When activated (by pressing #), the call is paused, and the current location within the voice service is converted to a DTMF-based audio code. Pressing # plays or repeats this audio code whenever the user requires.

Receiving: The recipient dials a short telephone number, and the remote service begins listening. When a code is detected, the listening service parses and decodes the tones, then redirects the caller to the correct voice service and exact position the sender requested. If recognition fails, the user is prompted to replay the code.

The use of audio-based codes makes the ACQR technique accessible to callers with any type of phone. This is a key benefit for those living in developing, rural regions as it ensures that no additional hardware is required to use the service.

Using audio as a method of facilitating sharing, however, also provides us with another key benefit; namely, that codes can be "heard" by more than one device simultaneously. Figure 2 illustrates three possible relationships afforded by

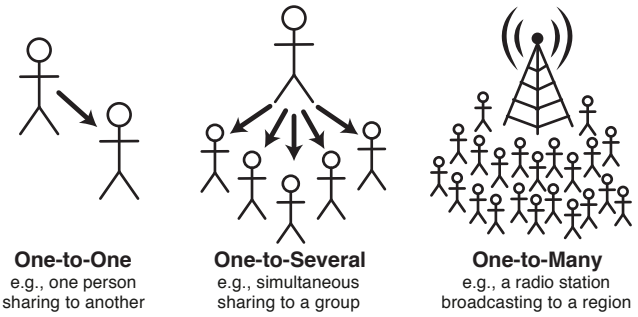

Figure 2. Sharing relationships possible with the ACQR technique.

our approach – sharing to a single person, sharing to several others (e.g., in a small local group) or even sharing to multiple people (e.g., over a radio broadcast).

SUMMARY

This demonstration has presented ACQR – an audio-based method of synchronising positions within large IVR hierarchies. We have described the interaction design of our approach and underlined its benefits for underdeveloped regions where textual literacy, income and data connections are low. We have also illustrated how our audio-based approach offers more than simply a one-to-one relationship by describing ways in which it could be used by multiple people at once. For a more in-depth discussion about the design and implementation of our approach, as well as a suite of evaluations highlighting its value to real Spoken Web users, see the full paper [3].

ACKNOWLEDGEMENTS

This work was funded by EPSRC grant number EP/J000604/2, and also supported by IBM Research India.

REFERENCES

1. *ICT Development Index*. ITU, 2002.
2. Kumar, A., Rajput, N., Chakraborty, D., Agarwal, S., and Nanavati, A. WWTW: The world wide telecom web. In *NSDR workshop, SIGCOMM '07*, ACM (2007).
3. Pearson, J., Robinson, S., Jones, M., Nanavati, A., and Rajput, N. ACQR: Acoustic quick response codes for content sharing on low end phones with no internet connectivity. In *Proc. MobileHCI '13*, ACM (2013).