

Making a Case for Spoken Web as the Mobile Web for Developing Countries

**Sheetal K Agarwal, Anupam Jain, Arun Kumar¹, Priyanka Manwani,
Amit A. Nanavati, Nitendra Rajput, Aaditeshwar Seth**

IBM India Research Laboratory
4, Block C, Vasant Kunj, Institutional Area,
New Delhi - 110070, INDIA
{sheetaga, anupamjain, kkarun, pmanwani, namit, rnitendra, aaditeshwar}@in.ibm.com

Abstract

A Mobile Web for developing regions is being viewed as a potential mechanism for proliferating information and Internet-based services to the masses. While there are disparate technologies that can come together to deliver such services to different segments, an integrated platform capable of serving the specific needs of this population segment is missing. Often what we have are point solutions with limited applicability. We propose that the Spoken Web can be considered as a viable realization of the Mobile Web concept offering comprehensive capabilities and catering to semi-literate, bottom-of-the-pyramid population of developing countries. We present an overview of the Spoken Web along with some of the lessons learnt and experiences gained from the field studies and pilots.

Introduction

Information and Communication Technologies (ICT) are being viewed as a promising medium for aiding social development at the bottom-of-the-pyramid to scale in a cost effective and sustainable manner. However, some of the challenges that often arise while exploring the appropriate use of ICT include the lack of suitable ICT infrastructure, un-affordability of the technologies by the poor, low literacy levels among the people, the lack of suitable content, and even cultural inhibitions to the use of technology. Unless these issues are addressed in the right context of the people, the impact of ICTs for development can be quite limited.

Mobile-phone based applications that expand the reach of Web to the masses are able to overcome many of these challenges. Governments and policy makers have realized the need for mobile communication, and as evidenced by the growth of mobile phones in India, Africa and China, the infrastructure is becoming more ubiquitously available. The cost of mobile phones is also falling rapidly, making them increasingly affordable to the poor. Finally, the voice-based access medium automatically overcomes illiteracy barriers. Thus, the use of mobile phones as an ICT medium certainly seems attuned to the context of the end-users that constitute the bottom-of-the-pyramid. All of these motivate the need for a web based on the mobile phone as the access device. Such a Mobile Web concept is envisioned to

¹ Member of W3C Mobile Web for Social Development (MW4D) Interest Group

expand the reach of the Web to the masses, thus bridging the knowledge gap among societies that is vital for survival and success in today information driven world [1].

Traditionally, solutions have been designed keeping the educated and well-to-do users in mind as the primary consumers. Such solutions often rely on ever increasing advances in technology. However, solutions meant specifically for the underprivileged in developing countries need to be built with a different mindset. The innovation need not be in building complicated systems with the most advanced hardware and software, but in finding simple and novel ways of applying existing technology to enable hitherto unavailable solutions.

With this background, we make a case for the Spoken Web (also known as World Wide Telecom Web) [0] as a Mobile Web capable of serving the needs of various categories of underprivileged users in developing countries. Spoken Web is a network of VoiceSites, or voice applications, accessible over a phone call. The users of the Spoken Web include illiterate/semi-literate people, highly mobile workers or even those who cannot afford high cost solutions. The Spoken Web is equally applicable to rural as well as urban users.

In this paper, we explain how and why Spoken Web is emerging as a preferred candidate solution for fulfilling several ICT needs of the underprivileged in developing regions. We describe how it enables ordinary end users to get transformed into service providers rather than being mere service consumers. We will also touch upon some of the compelling applications that are beginning to be offered as applications over this new alternative web that is complementary to its established counterpart – the World Wide Web.

Technology

Several challenges exist in the path towards realization of the Mobile Web concept for developing regions. These include the vast social diversity present among developing countries, the resulting uncertainty in the migration path to data services as well as non obviousness of the transition from desktop to handhelds in developing countries [1].

IBM's India Research Laboratory envisions overcoming these challenges by building an alternate web for the masses in developing countries. It is called The Spoken Web, also known as The World Wide Telecom Web [0]. Keeping in view the diversity in language, culture, dialects and the varying levels of literacy, it is designed to work with voice as the primary user interface modality. To keep the acquisition and maintenance cost extremely low, this web only requires the possession of an ordinary telephone instrument capable of making voice calls and accepting Dual Tone Multi Frequency (DTMF)² inputs and can be offered to end users under a hosting model. Further, Spoken Web relies on the presence and availability of already established telecom infrastructure either wired or wireless. All the above characteristics put Spoken Web in a unique position to cater to the needs of the population segment yet untouched by IT, in a very simple manner.

² Navigation is performed by pressing appropriate keys on the phone (Press 1 for A, 2 for B,...), also known as touchtone dialing.

Technologically, the Spoken Web is a network of interlinked VoiceSites that are voice applications created by subscribers and hosted in the network. It uses voice (and DTMF) compared to text, as the primary mode of interaction and operates over existing telephony infrastructure. The key strength comes from the simplicity with which users can interact can create their VoiceSites in Spoken Web. Using Voigen [5], users can create their own VoiceSites simply by talking to a voice application over a phone call. The generated VoiceSite could mostly be in the local language of the VoiceSite owner, if so chosen. These VoiceSites can be hyperlinked to one another through Hyperspeech Transfer Protocol [7] and can be accessed through a simple phone call.

The Spoken Web ecosystem makes few important contributions. The first is the availability of a new, simple and pervasive channel for information access and dissemination. Second is the use of an intuitive user interface that crosses the linguistic and cultural barriers. Third is the powerful capability of enabling even the illiterate end users to contribute and upload content as well as create and offer their own services. Fourth is the ability to shield the end user from the complexity to manage the content and services in the network. Fifth is the use of existing network infrastructure. Lastly, the entry barrier to access Spoken Web is low due to its support for low end devices.

Field Pilots/Studies

In this section, we present some applications either studied or deployed in the field and provide a view of the potential uses of the platform in different settings.

VoiKiosk

A VoiKiosk is a VoiceSite that acts as an information and service portal for a village or a community. It is a central point of access where information relevant to the community can be posted / accessed directly by the users themselves. It does not rely on internet connectivity and allows end users to directly interact with the services removing the dependence on the kiosk operator [2].

There are two types of users for this system. First is the kiosk operator who is responsible for supplying local content on the VoiKiosk. She configures the VoiKiosk for the village through a voice interface. Second are the end users, who can either access or post information on the VoiKiosk depending on the services that it supports.

The information on the kiosk is uploaded, maintained and consumed by the local community. Illiterate and semi-literate users find it very convenient to talk to the system in their own local language. The highly interactive nature of the system empowers the local community to become information providers instead of just consumers. Also, this enables NGOs, government and other agencies to reach out to the rural areas in a way that has not been possible before.

Dial-An-Auto

Dial-An-Auto [5] is a spoken dialog matchmaking system offered as a VoiceSite that enables unorganized auto-drivers (taxis) to find the closest passenger and vice versa at any time.

Unlike taxis which have their designated stands, autos are hailed from the street and they rarely have designated parking space while waiting for customers. Auto, as we found out, have 7-8 hours of idle time per day (waiting for customers) from their working periods of 12-14 hours per day. From surveys we discovered that there are several inefficiencies in their current method of operation. These inefficiencies lead to problems such as dishonesty of the drivers, and refusal to go on non-lucrative routes.

Also, often while waiting for customers on roadside or making momentary stoppages on the roadside they get fined by the traffic police as they are not authorized to park on road. All of these ultimately result in harassment to the drivers and the passengers. These and other inefficiencies result from the fact that no central entity that can coordinate the operations of the autonomous autos exists.

Dial-An-Auto application provides a mechanism to organize auto-drivers into a virtual community while offering services to passengers. The fare prices and decision to accept a passenger can still rest with the auto driver thus retaining the autonomy of the drivers.

Avaaj Otlo

Avaaj Otlo [4] ("voice-based community forum"), is a VoiceSite in Gujarati language application allowing farmers to access agricultural information over the phone. It was designed specifically for the farming community of rural Gujarat. It allows farmers to access latest information regarding farming practices; it enables them to interact with agriculture experts by posting their questions on the VoiceSite. These questions may also be answered by other farmers, encouraging more peer interaction. The content for the VoiceSite is managed by a VoiceSite administrator who belongs to the local community. He is responsible for posting farming related information and monitor the questions that are posted on the site. Any post that is irrelevant or inappropriate is deleted. The administrator manages the site through a voice based interface. This application is currently available only to selected farmers registered with the site.

VoiAvatar

VoiAvatar [5] system enables individuals to create their own personal or business VoiceSites. Each of these VoiceSites acts as an online avatar or proxy of those individuals. These VoiAvatars can be accessed in the absence of the owner (or otherwise) to get information about them and their work in essence acting as the owners' automated secretary. Such information could include area of operation, service charges and work hours, in case of a micro-businessman (such as plumbers, electricians, carpenters, craftsmen etc) for instance.

Folksomaps

Sparsely populated semi-urban and vast rural areas of developing countries such as India do not have detailed map systems built for most locations. The semi-literate, low income, Non-ITsavvy population residing in these areas cannot use such services even if they were offered online over the internet. Furthermore, lack of stable infrastructure including electricity, internet connectivity and the lower purchase power of people, also complicate the situation further. Secondly, lack of structured addressing conventions and poor road signs makes it difficult to follow the maps, if any. So, even people comfortable with maps, often need to ask people on the streets to find their way.

Folksomaps [3] is a community driven map system offered as a VoiceSite and a website that leverages Semantic Web technologies to create and manage a community generated knowledge base and makes use of web and voice applications to provide access to its services. Folksomaps is designed to be populated by end users for their own consumption

Applications for Visually Impaired

Websites in World Wide Web are primarily meant for visual consumption. Accessibility tools such as screen readers that render the visual content in audio format enable the visually impaired to access information on the websites but they have their own limitations.

Since the access to VoiceSites is a simple phone call, it can become a pervasive and low cost IT access mechanism for the blind. To demonstrate the utility, we developed a sample Grievance VoiceSite that enables people to lodge their complaints at the Ministry of Social Justice, Government of India.

We conducted surveys at two institutions for the blind in New Delhi which indicate that the learning curve for using applications on the Spoken Web is relatively low and does not require extensive training.

Besides accessing content there are many potential applications based on the Spoken Web which enable the blind to publish their content and reach out to a wider community. VoiceSites designed as Audio Blogs and interlinked through VoiLinks while including features such as voice messaging, voice-mailing, voice chat, referrals etc., that enable social networking services for the visually blind and a business VoiceSite for the visually impaired micro-entrepreneurs, advertising the services offered by him and created using The Telecom Web framework are a few examples.

Conclusion

Spoken Web is gathering acceptance among academia and industry as a simple and viable candidate for delivering locally relevant content, information and services to the masses. However, it has some limitations and challenges. Voice being the primary modality of interaction, the user interface and functionality offered is sometimes constrained. From our field experiences we have learnt that voice does not become a road block as a user interface but offering images and pictures is

currently not supported over Spoken Web. Further, compared to the IT savvy population that has access to many communication channels, our target user segment is more patient with voice interfaces. This is primarily because often the only other option available to them is a few miles of walk or an entire day and significant money spent to get some information.

There are research challenges yet to be resolved. These include keeping the navigation and access to content simple as VoiceSites grow in size and number, enabling search over VoiceSites, and providing mechanisms for scalable deployment of Spoken Web over large geographies at a low cost. We hope to provide solutions to these and other issues in the near future.

References

1. Ravi Jain, **The Mobile Web in Developing Countries**, W3C Workshop on the *Mobile Web in Developing Countries*, Dec 2006, Bangalore, India.
2. Sheetal K Agarwal, Arun Kumar, Amit Anil Nanavati, Nitendra Rajput, **Content Creation and Dissemination by-and-for Users in Rural Areas**, To appear in proceedings of 3rd IEEE/ACM International Conference on Information and Communication Technologies and Development (ICTD2009), 17-19 April 2009, Doha, Qatar.
3. Arun Kumar, Dipanjan Chakraborty, Himanshu Chauhan, Sheetal K. Agarwal, Nitendra Rajput, **FOLKSOMAPS - Towards Community Driven Intelligent Maps for Developing Regions**, To appear in proceedings of 3rd IEEE/ACM International Conference on Information and Communication Technologies and Development (ICTD2009), 17-19 April 2009, Doha, Qatar.
4. Neil Patel, Sheetal Agarwal, Nitendra Rajput, Amit Nanavati, Paresh Dave and Tapan S. Parikh, **A Comparative Study of Speech and Dialed Input Voice Interfaces in Rural India**, To appear in proceedings of ACM Conference on Human Factors in Computing Systems (CHI), 4-9 April 2009, Boston, US
5. Arun Kumar, Nitendra Rajput, Sheetal K Agarwal, Dipanjan Chakraborty, Amit A Nanavati, **Organizing the Unorganized - Employing IT to Empower the Under-privileged**, In proceedings of *2008 International World Wide Web Conference (WWW)*, Beijing, China, 21-25 April 2008.
6. Arun Kumar, Nitendra Rajput, Dipanjan Chakraborty, Sheetal K. Agarwal and Amit A. Nanavati, **WWTW: The World Wide Telecom Web**, In *Proceedings of ACM SIGCOMM Workshop on Networked Systems for Developing Regions (NSDR)*, August 2007, Kyoto, Japan.
7. Sheetal K Agarwal, Dipanjan Chakraborty, Arun Kumar, Amit A Nanavati, and Nitendra Rajput **HSTP: Hyperspeech Transfer Protocol**, In *Proceedings of the Eighteenth International ACM Conference on Hypertext and Hypermedia (Hypertext 2007)*, Sept. 10-12th, 2007, Manchester, UK.