A Picture and a Thousand Words: Visual Scaffolding for Mobile Communication in the Developing World*

Robert Farrell, Catalina Danis, Thomas Erickson,
Jason Ellis, Jim Christensen, Mark Bailey, and Wendy A. Kellogg

IBM T J Watson Research Center

Abstract

Mobile communication is a key enabler for economic, social and political change in developing regions of the world. Today's internet-enabled, multimedia, and touch-screen mobile smartphones could become the future platform for delivering information and communication technology (ICT) to these regions. We describe Picture Talk, a smartphone application framework designed to facilitate local information sharing in regions with sparse Internet connectivity, low literacy rates and having users with little prior experience with information technology. We argue that engaging citizens in developing regions in information creation and information sharing leverages peoples' existing social networks to facilitate transmission of critical information, exchange of ideas, and distributed problem solving, all of which can promote economic development.

Keywords: Mobile technologies, User Interface, Computing in Developing Countries, Social Issues

Introduction

We are interested in designing applications that enable people at the base of the economic pyramid (BoP) to create, share, and discuss information as is commonly done on the World-Wide Web today, but through mobile technologies. The BoP includes over one billion people with little access to computer technology living on less than $1US per day in some of the least developed countries in sub-Saharan Africa, the Indian Sub-continent, and parts of Asia and South/Central America. As others have recognized (Prahalad, 2004; Kumar et al, 2008), enabling connections among a wide spectrum of people can lead to the empowerment of the disenfranchised and enable people at the BoP to express their entrepreneurial tendencies. This could result, for example, in the creation of broader markets for local goods and services. The global reach of mobile communication networks offers, for the first time, a broad platform for delivering applications and software services in BoP regions.

We have three long-term goals for the mobile applications we build. First, we want applications we develop and deploy to be usable by even the most disadvantaged users. Second, we want to

* To appear in *The International Journal of Handheld Computing Research (in press, 2009)*.

enable these users to document local needs, problems, and issues by creating, storing, and sharing digital artifacts (e.g., maps, photos, graphics, radio news reports, music, games, TV segments, informal news). Third, we want to enable these users to engage in conversation about these digital artifacts to offer solutions, share perspectives, or to engage in social exchanges.

Our initial implementation toward these goals is Picture Talk, a social computing application framework that enhances persistent conversations with visual scaffolding. The social computing features support social behavior and social connections between users (Danis et al., 2009) through mobile phone conversations. The persistent conversation feature allows users to engage in spoken discussion asynchronously. Visual scaffolding provides structure for these asynchronous voice-based communications, enabling parallel access rather than requiring serial access as is done in voice-only messaging systems. Participants in Picture Talk conversations can engage in topics of shared interest using multiple access channels: telephone (voice-only), web browser or mobile smartphone (w/data connection), and mobile phones with Multimedia Messaging Service (MMS).

This paper first discusses some of the obstacles that BoP communities face in trying to access information technology, then introduces the Picture Talk application framework design and an implementation, and then discusses some of the particular challenges of the BoP environment for application developers.

Background

In this section we provide background on some of the obstacles that BoP populations currently face in becoming part of the global community with access to information technology.

In the economically developed world, access to information technology has been largely through Internet-connected computers. An important benefit of access to the Internet has been the potential for contact with the worldwide community of users. The Usenet network, one of the earliest online discussion venues (created in 1979), supported threaded discussion on a wide variety of topics among participants distributed worldwide. Online communities became very popular in the 1980s and 1990s. For example, the WELL ("Whole Earth 'Lectronic Link") was a hybrid face-to-face and online group that served participants in the Bay area of San Francisco, California (Rheingold, 1993). Members of the WELL engaged in discussions of topics of common interest and the forum also served as a means of self-expression. Similar applications could be deployed to BoP communities to enable discussions on topics of local interest, provide a voice for individuals who would otherwise have no forum for their ideas, and enable solutions to communal problems through information exchange.

The rapid uptake of mobile phones in developing regions has yielded examples that demonstrate the feasibility of giving individuals at the BoP a voice and aggregating their contribution to provide value to a broader audience. For example, Ushahidi, meaning "testimony" in Swahili, is a platform for crowdsourcing crisis information. Ushahidi allows anyone to transmit geo-coded data via Short Message Service (SMS), email or web and visualize it on a map or timeline. Time-sensitive information from the public is aggregated and distributed widely (Ushahidi.com,

2009). Another example, the AfriGadget site (AfriGadget, 2009) aggregates reports "showcasing African ingenuity" that are provided through emails by individuals throughout Africa. We would like to include both citizen journalism and reader comments in our designs.

While these examples illustrate that people in economically developing regions are beginning to participate in the production and consumption of information, particularly as it is enabled by mobile telephones, they also illustrate some of the obstacles to their widespread use in developing regions. Three obstacles are germane to our arguments. First, despite initiatives such as One Laptop Per Child (OLPC, 2009), computing technology remains out of reach of the large majority of the BoP population. Lack of reliable networks to access the Internet further limits the ability of people in these regions to access information even where capable devices are available. BoP users are necessarily very cost conscious, driving a need for a low cost platform comprised of both a mobile wireless infrastructure and low-cost mobile access devices. Second, low literacy rates prevent significant portions of the BoP from using the Internet's predominantly textual interaction mode. Third, despite skills and experience in the social use of mobile phones, many BoP users may have little familiarity with or motivation to use the device to access information services, preferring face-to-face interaction. In this section we examine each of these obstacles in more detail.

Technology Landscape

The statistics for even basic access to electricity in developing regions are alarming. According to the Open Society Initiative for Southern Africa (OSISA, 2009), approximately 90% of Africa's one billion people have no regular access to electricity. Where power to homes is not available, people often travel to a centrally located solar-powered, wind-powered, or coin-operated charging station to maintain use of their mobile phone. Inventors are working on providing ways of generating electricity from personal movement (AfriGadget, 2009).

Global statistics for computer usage demonstrate huge differences between developed and developing countries. For example, the highest rates of access to the Internet in 2007 were in Sweden (82%), the US (81%) South Korea (81%), and other developed countries, whereas the lowest access rates included Tanzania (6%), Kenya (12%), and Uganda (11%). Similarly, computer ownership is lowest in Uganda and Tanzania, both at 2% (The Pew Research Center for the People and the Press, 2007).

The picture changes radically when considering mobile phones rather than computers. A recent survey by the International Telecommunications Union found that while only one quarter of the earth's population of 6.7 billion uses the Internet, nearly two thirds of the population uses mobile phones (ITU, 2009). Wireless phone use is exploding in the developing world: Sixty-eight percent of mobile phone subscribers worldwide are outside of North America and Europe. In Africa, mobile subscribers have jumped from 10 million to 400 million in the last five years (2003-2008) and the growth is still accelerating (ITU, 2009). The rate of mobile phone ownership in the Ivory Coast, Mali, Nigeria and South Africa is over 60%, higher than in Canada (The Pew Global Attitudes Project, 2007). In 2009, India will become the second largest wireless phone subscriber base in the world, after China (EE Times, 2008). Mobile phone use in economically developing regions crosses the barriers of gender, age, and education (Samuel et. al., 2005).

The exponential growth of mobile phone networks in BoP markets is fueled by the need for communication in environments where there are few alternatives. The lack of traditional wired infrastructure creates an opportunity: much of the developing world is a "green field" where new computer and communications technologies can be deployed without being hampered by existing business models, infrastructures, or user expectations. For example, in many parts of Africa, wireless networks have leapfrogged the public switched telephone network in terms of installed base. In 2007, the African continent had 280 million total telephone subscribers, but 260 million of these were mobile cellular subscribers. Building a new wireless network is faster, easier, more reliable, and less expensive than putting in a whole new wired infrastructure.

Despite the growth of wireless networks, few developing countries yet have data communications channels sufficient to provide rural populations with access to the public Internet. In 2008, only 7% of India (Internet World Stats, 2008) and 5% of Africa (Appfrica, 2008) had access to the Internet.

The capabilities of mobile phones are also increasing rapidly. In the early 1990s, few phone users would have been aware of mobile text messaging, but by 2008, almost 3.5 trillion SMS messages were sent worldwide (Portio Research, 2009). The first deployments of camera phones occurred in 2001 and by 2004, 370 million mobile phones with digital cameras were sold (InfoTrend/CAP Ventures, 2004). In the late 1980s and early 1990s, cell phones were used for voice communication only and users typed on a numeric keypad. Today's smartphones have high resolution touch screen displays, miniature keyboards, and other flexible input methods. Worldwide smartphone sales increased 12.7 percent in the first quarter of 2009 (Gartner Group, 2009) and sales are anticipated to grow at more than a 30% compound annual growth rate over the next five years. Today more smartphones are sold globally than laptops (INSTAT, 2007).

Literacy

Literacy is typically defined as the ability to read and write, however there is an inherent lack of precision that results from the methods of assessment and thus official figures often over-estimate functional literacy. For example, the commonly cited statistics, such as those compiled by UNESCO (2009), are based on census and other self-report methods which are fundamentally inexact. Also, the definition of literacy can vary from 'the ability to write a simple sentence' to 'being able to freely communicate ideas in literate society.' Individuals as young as fifteen who may have been counted as literate because they were counted when they were attending primary or secondary school may, because of lack of language use, be functionally illiterate as adults (Seshagiri, Sagar & Joshi, 2007). According to UNESCO (2009), two-thirds of the world's 785 million illiterate adults are found in only eight countries (India, China, Bangladesh, Pakistan, Nigeria, Ethiopia, Indonesia, and Egypt). Low literacy rates are concentrated in South and West Asia, sub-Saharan Africa, and the Arab states (CIA, 2009), with percentages averaging in the 60s, though some countries like Mali and Niger report rates for 15 to 20 year olds of less than 30%. Men typically have higher rates of literacy than women in traditional societies (UNESCO, 2009).

While there are no generally accepted statistics on how much of the Internet is available in different languages, it is generally accepted that the dominant language on the Internet is English, making much of the Internet linguistically inaccessible to the large majority of the BoP

(EnglishEnglish.com, 2009). The large number of languages spoken in BoP countries is intertwined with literacy and access to written information. While countries such as India have two official languages (Hindi and English), there are an additional 22 "scheduled" languages, and approximately 400 other languages in use by significant numbers of the population (Ethnologue, 2006). Thus individuals who may be literate in their native language may nevertheless be functionally illiterate if information is available only in one of the official languages (Plauché and Nallasamy, 2007). A report by UNESCO indicates that economically developed countries may be marginalizing speakers of hundreds of local languages (UNESCO, 2008).

Designers of applications geared towards illiterate users have focused on non-text modalities in order to design more generally accessible applications in the countries with low rates of literacy. For example, speech is a widely used modality, even in kiosks (Morris, 2000). However, limitations on the generality of speech recognition technology in multi-lingual environments (Plauché and Nallasamy, 2007) demands the use of other modalities for a broad set of functions needed to complement spoken language interfaces. For example, Joshi, Welankar, Kanitkar and Sheikh (2008) developed and tested a phonebook they call Rangoli aimed at low literacy populations. Rather than entering phone numbers based on alphabetical order, users are able to use a combination of color, icon and spatial location. Similarly, Froehlich and colleagues (2009) proposed applying digital storytelling (for example, video or sequences of still photographs accompanied with spoken annotations) as ways of enabling low literacy individuals to participate in information creation and sharing.) As noted above, even for literate users in the population, the large number of languages in BoP countries makes it unlikely that the user's preferred local language will be used in the user interface. Thus the use of pictures and spoken language may allow many more people to have meaningful access to information.

Social and Cultural Context

Technologies deployed in developing regions must be sensitive to the social and cultural contexts in which they operate. To provide one example, in the southeastern Indian state of Kerala, fishermen now use mobile phones to get market price information before deciding where to sell their fish (Abraham, 2007). About 40% report an increase in income and 50% report fewer losses due to unsold or spoiled fish when they start calling for prices. Interestingly, however, few of these fishermen consistently go to markets with the highest prices; instead many choose ports where their "commission agent" has a presence. Because commission agents invest in the fisherman's business (e.g., financing the purchase of a fishing vessel), the fisherman feels a social obligation to bow to the agent's wishes, even when doing so may prevent him from maximizing his income.

Several field reports illustrate specific ways in which trust in one's social network and distrust in official sources of information influence the use of computing technology. For example, farmers in the southern state of Tamil Nadu use web-connected kiosks (telecenters) fielded by a local sugar factory to ask only "simple" (i.e., low-stakes) questions of a purported agricultural expert who is not known to them, saving "high-stakes" questions for successful farmers with whom they have some pre-existing relationship (Srinivasan, 2007). Gopakumar (2006) explains that local people play a critical intermediary role in the success of telecenters. For example, living in the same village led target users of the Akshaya telecenter to develop trust in the entrepreneurs

and intermediaries who ran the centers. By extension, they also developed trust in the abstract systems of medicine and government that were the ultimate sources of the information.

To summarize, these studies demonstrate the power that access to information can have in improving people's lives, but also how the impact of information is gated by social factors like trust, accountability, and social and institutional pressures.

The question we address in the remainder of the paper is: how can we address the impact of the factors we have discussed – constrained technology landscape, low literacy rates and a traditional social and cultural context – when designing systems appropriate for the billions of potential users at the base of the economic pyramid?

A Mobile Social Computing Application Framework

Picture Talk is a software application framework intended to support a wide range of social interactions that can be accomplished through asynchronous communication, including conversations with remote participants, question and answer exchanges, and peer production of localized content. This section begins by laying out the rationale that underlies Picture Talk by describing the scenarios and design sketches that marked the beginning of the design process. After presenting the initial vision, it goes on to describe a working prototype.

Rationale and design sketches

Because of the large numbers of local languages and widespread written illiteracy, speech seems like an obvious choice for supporting mediated interaction in many areas of the world. However, when speech is transposed into digital settings, many things change and a number of well-known problems arise. In the type of application we were envisioning, conversations would be asynchronous, carried out between people in different places speaking at different times. This means that Picture Talk conversations would lack some characteristics that are important for establishing and maintaining common ground (Clark & Brennan, 1991) -- "the knowledge that the participants have in common, and they are aware that they have it in common" (Olson & Olson, 2000, pp. 157). For example, it would mean that speakers would not be able to see one another, or share visual cues like glances, gestures and shrugs that in collocated speech enable interlocutors to control the conversation's flow, easily refer to objects, and verify that they are being understood (e.g., Yankelovich et al., 2004). It also potentially means that many more people can engage in a conversation, something that could be valuable but which also could exacerbate these problems.

The concept of Picture Talk arose out of consideration of these problems, and how they might be addressed in the context of a mobile phone-based communication system. The crux of the solution was to augment speech with three types of visual component: comment proxies, pictorial contexts and visual controls. Comment proxies are visual representations of digital speech that depict various types of meta-information, such as the identity of the speaker, the length of the comment, and the relationship of the comment to other comments (e.g., a reply); they also provide direct access to the comment they represent, thus mitigating the difficulty of navigating voice posts. Pictorial contexts are diagrams or photographs that provide a background

for a particular conversation; pictorial contexts serve both to represent the conversation as a whole, and allow comment proxies to take on additional meaning by virtue of their location with respect to the pictorial background. Finally, visual controls are a variety of visual user interface components for controlling the system, for example, a message play button.

Figure 1 shows three early design sketches of Picture Talk developed in the context of a scenario set in rural India. (By 'design sketches' we mean provisional concepts that are intended as conversation starters with stakeholders, rather than as depictions of well-considered solutions.) The first sketch, Rice Talk, envisions an asynchronous conversation among farmers about pests and diseases affecting their rice plants. It consists of (1) a white 'card' showing a diagram of a rice plant (the pictorial context); (2) a series of colored bars (the comment proxies) that represent spoken comments, showing their durations, which of them have been made by the same speaker, and the part of the plant to which the comments refer; and (3) a floating 'talk' button (the visual control). The second sketch shows a health-oriented conversation with red and

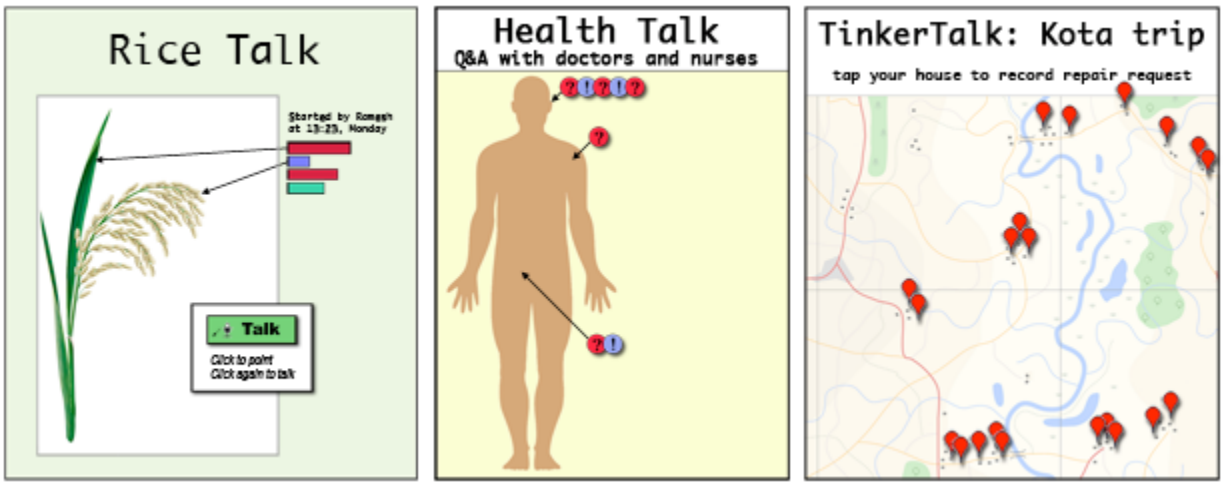

Figure 1: Three design sketches of the Picture Talk concept for applications set in rural India: (a) Rice Talk, for farmers to discuss problems with their Rice plants; (b) Health Talk, for villagers to discuss health problems; and (c) TinkerTalk, for people in a region to indicate that they need the services of a traveling tinker. Background images of rice plant © Ivan Kopylov | Dreamstime.com, of human body © Dannyphoto80 | Dreamstime.com, and of map © Robert Adrian Hillman | Dreamstime.com. Used with permission.

blue circles (the comment proxies) superimposed over a diagram of the human body (the pictorial context), the circles' positions indicating what aspect of the body or health they refer to and how they are related to other comments. The third sketch shows a conversation between a traveling tinker (i. e., a mender of pots) and potential customers, the pictorial context being a map of the region, and comment proxies (the red balloons) indicating where the speaker is located.

Besides communicating the basic idea behind Picture Talk – using pictures, and simple visual representations of voice comments to provide scaffolding for asynchronous speech-based communication – the sketches serve other purposes. First of all, they illustrate the flexibility of the basic concepts. The pictorial contexts, and similarly the comment proxies, can represent a large range of topics, and even when depicted as simple geometric shapes, they can represent a

considerable array of meta-information. Perhaps more importantly, the sketches are useful in raising a number of questions both within the design team, and with other audiences. How do the pictures get into the system? What sort of meta-information should comment proxies depict? Do different conversations benefit from the display of different comment meta-information? What sort of visual representations will be understandable by the envisioned user populations? How do users find their ways to particular conversations? As the aim of this paper is not to trace the trajectory of the design, it will not detail its evolution, but will instead move on to describe the user experience of the resulting working prototype.

Prototype Implementation

Our implementation of Picture Talk consists of a client application running on the Android™ G1 ™ mobile phone and a centralized data server running a Picture Talk Web service in the Ruby on Rails™ (RoR, 2009) Web application server environment.

When users launch the application, their phone number is used to retrieve their user profile from the Web service. If this is the first time the user has accessed the service, they are prompted to record their name and take a picture of themselves. The user is then presented with a menu that has four options: take a picture, view the gallery of the pictures taken by other users, view the profiles of other users, or update one's own profile.

Users can start a discussion by simply taking a picture and tapping anywhere on the photo. The system stores the picture in the gallery of shared pictures and records various metadata (e.g., who started the discussion, the time and date). Additional metadata could be stored, such as the location where the picture was taken, using the built-in Global Positioning System (GPS) receiver on the G1 phone. Subsequently, users can join an ongoing discussion by finding the picture in the gallery and tapping on it, and being lead to the discussion screen.

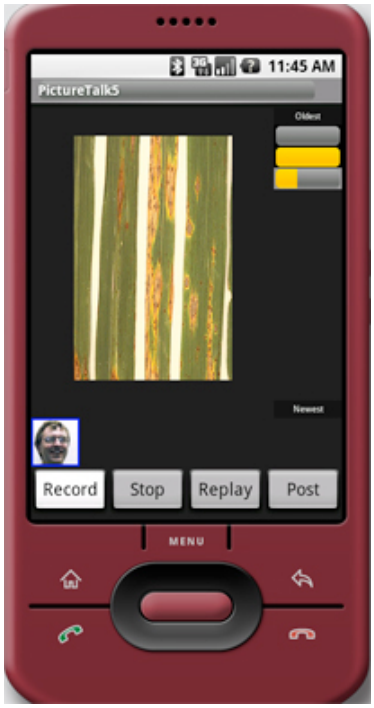

Figure 2: A Picture Talk discussion

The discussion screen (see Figure 2) has four elements: the context (a picture), comment proxies (graphics on the upper right depicting spoken comments about the picture), participant icons (a horizontal scrolling gallery of photos), and visual controls (buttons beneath the pictures of the participants to control audio recording and playback). In the spoken comments area, each graphic represents a single comment from a user. We are exploring various techniques for associating the speaker's photo with her comment.

We designed the audio controls to allow the user to compose and review a recording before posting it to the discussion for others to hear. A bar graphic is drawn under the audio buttons to reflect the length of the recording. While recording audio, the user can tap on the picture to point out something of interest in the picture, for example, the diseased part of a rice plant. The visual annotation will then be associated with the comment. When a

user posts a comment, the bar graphic is posted to the discussion area to the right of the picture. The bar graphic provides a visual "residue" of the comment recording (Hollan, Hutchins, & Kirsh, 2000) for subsequent users. The length of the bar reflects the length of the recording. Pressing anywhere on the bar graphic starts playing the recorded audio and displays any corresponding visual annotation on the picture. The same set of controls is used for both recording and playback, much like a music player. Users can pause the playback or replay the audio from the beginning. The bar graphics are listed chronologically from top to bottom in a scrolling window with the most recent always visible.

Posting a comment stores the comment, and any visual annotations, with the discussion so that subsequent users accessing the picture can access the comment's audio and visual elements. It also stores metadata (who made the comment, the date and time of their comment), posts the author's photo to the scrolling gallery of discussion participants, notifies other users in the discussion that there is a new comment (using SMS), and makes the respondent's profile visible to other discussion participants. The respondent's profile can help participants determine their trust in the information provided by the respondent. For example, a respondent may be a friend who is instantly recognizable from their photo or may be someone not known to the discussants but nonetheless reputable. The person starting the discussion is able to invite additional discussants by sending SMS messages. Individuals may block their ability to receive these notifications.

The photos of each user in the discussion are posted below the picture, in a scrolling picture gallery. Touching a user's photo leads to their user profile. The user profile has their contact information (telephone number) and a scrolling gallery of the pictures anchoring discussions they have started. Touching a picture leads to a discussion screen with the given picture as the context. In this way, users can quickly find and engage in discussions started by other participants. This could be useful, for example, if a user has come across a farmer who has posted useful information about rice fungi and wants to see what other advice the farmer may have provided on other topics.

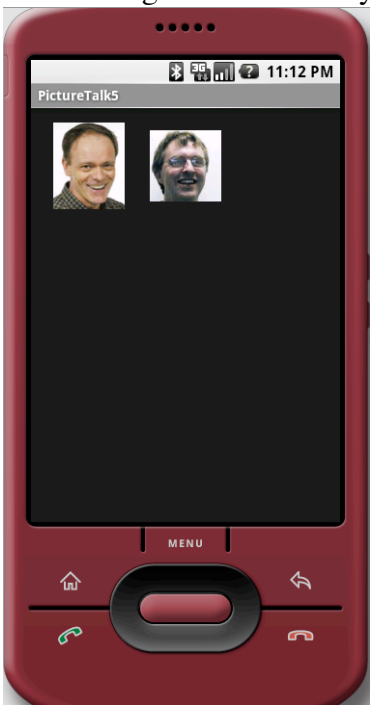

Figure 3: A visual menu of a Picture Talk user's social network

Picture Talk provides the option to view photos of one's co-discussants on a topic (see Figure 3). Organizing discussions by the people who start them has several benefits: First, users can learn about other people by browsing the discussions those users have started; second, other users can make use of the person's photo to cue memory of relevant discussion contexts; and finally, the total number of pictures may be large, but the number of discussions started by an individual will likely be more constrained and thus easier to remember.

Picture Talk is architected as a client-server application (see Figure 4). A Java application, running on the Android Linux-based operating system, is launched from the Android phone and accesses the Web Picture Talk data server which is a server

machine with a Web service running on Ruby on Rails. The Picture Talk data server provides a persistent data model for the application's objects (discussions, pictures, comments, audio clips, people, etc.). To minimize the data exchanged between clients and server (and hence conserve wireless bandwidth), the server assigns version numbers to the data objects so that both client and server know when data object updates are needed to synchronize the data model. The server uses Rails' active record support to store and access the data objects in a MySQL® database. Pictures and voice recordings are stored in files.

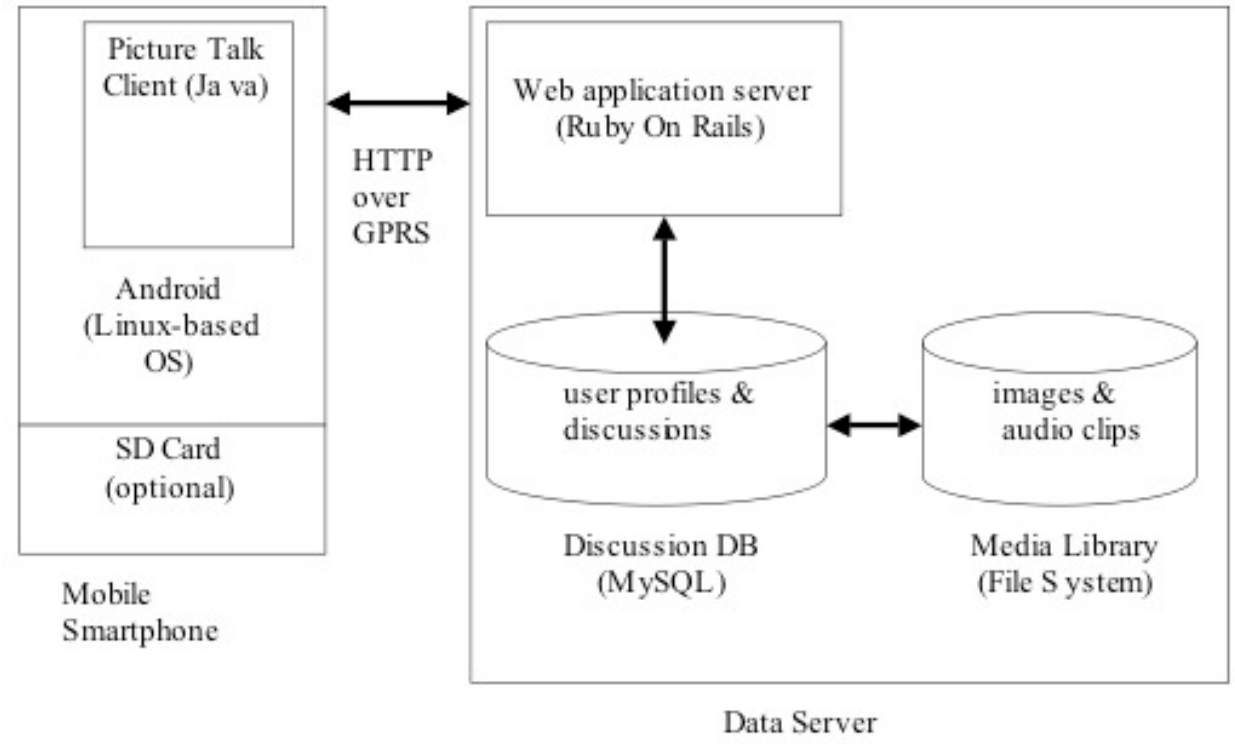

Figure 4: Picture Talk architecture

The Picture Talk client is installed as a third-party application and runs on Google's Android open source operating system (OS). Android has an emulator that we have used to prototype our application. Communication with the Ruby on Rails server happens over General Packet Radio Service (GPRS), a package-oriented data service with increasing penetration into the developing world. Several wireless carriers offer compatible phones for the Android platform. The G1 ™ phone has suitable hardware for running Picture Talk client: a 3.2-inch touch-screen display, wireless networking, a microphone, built-in speakers, a camera, and gigabytes of external storage. A number of smartphones provide similar functionality, but Picture Talk takes advantage of the Android OS's capability of accessing the phone's hardware, including detecting the presence of wireless network services, recording and playback of audio, controlling the built-in camera and storing pictures on the phone and in external storage.

When the Android client has access to a wireless network, it sends pictures captured with the phone's camera and audio captured with the phone's microphone to the Picture Talk server and automatically updates the currently displayed discussion. When the phone is disconnected, new

pictures from the camera are stored on its Secure Digital (SD) card, when available, or on the phone's local storage and users can still start discussions, make audio postings, listen to existing audio postings, and update personal information. When disconnected, data objects are stored and retrieved from a database local to the phone using Android's SQLite software library.

Kiosk and Voice-only Access

Given the current technology trajectory in developing nations, we expect to see increased adoption of smartphones (such as the G1 Android) in developing nations in the next three to five years. But in order to get early feedback on our designs, we are interested in deploying Picture Talk as widely as possible in the near term as well. Thus, we have developed a voice-only version of Picture Talk in order to make the application accessible to users of lower end phones. We have also created a web version, suitable for kiosk or telecenter use. The voice-only client allows people using basic mobile phones, commonly found in BoP environments, to listen to and record discussion comments, and even exchange pictures with the discussion server via MMS, if available. An additional server-side component, built using the open source Asterisk® Public Branch Exchange (PBX) telephony toolkit, provides voice and telephone keypad Interactive Voice Response (IVR) interfaces for low-end camera phones, and in turn uses the persistent data server (described above) to access discussion objects. Both of these clients access the same data as the Android client, but display that data in a suitable way for the platform at hand. For example, the rice plant anchoring the discussion in Figure 1 is sent using MMS. Subsequently, when another user wants to participate in the discussion, the Picture Talk server first sends the picture to their phone in another MMS message. Users listen to voice comments over a normal voice channel. We had to develop additional server-side functions to transform the audio and image objects into formats usable by and optimized for both wireless phones and desktop computers.

While the user experience for voice-only clients is necessarily more restrictive than with the smartphone or web browser clients, having the voice-only option makes Picture Talk discussions potentially available to a broad range of BoP users. Further research is needed to enable voice-only clients to more effectively find and navigate relevant content, share information, and tap into discussion databases that have heretofore been usable only from data-capable devices in the hands of literate users.

Future Research Directions

Many of Picture Talk's features represent general capabilities that could be applied in a variety of mobile applications. In this section, we look at several such features and discuss some additional challenges in developing applications for BoP markets and future research directions to address these challenges.

Identity

Like many social software applications, Picture Talk helps users share their appearance, contact information, and so on, to each other. However, in many developing regions of the world, it is common for mobile phones to be shared amongst members of a family or even an entire village. Adeya (2005) describes one African couple that shared one mobile phone: the wife used the

phone during the day for business and the husband at night for personal calls. To address this issue, some handset manufacturers have added support for multiple address books on one phone. In some cases the very notion of ownership may be quite different from the idea of "personal property" common in developed nations. Mobile social computing applications in many BoP contexts will need to allow users to identify themselves to the system explicitly and in innovative ways (e.g., by selecting their picture or identifying a vocal sample).

Participation, Inclusion, and Viral Spread

Picture Talk promotes participation and inclusion in three ways. First, anyone can start and manage a Picture Talk discussion. Second, anyone can discover and engage in discussions started by other users. Finally, as stated previously, Picture Talk provides the ability to send a text message to anyone inviting him or her to join in a Picture Talk discussion. This ability to notify and invite people to participate, even people who are not currently registered users of the system, supports the possibility of "viral" growth of the Picture Talk user population. In developing regions where the idea of using technology to access information beyond one's social network may unfamiliar, viral spread can be a key bridging mechanism. If a friend recommends a Picture Talk discussion to a potential new user, they may be more likely to engage, find something of value, and become a "consumer" of information than if they needed to find the information themselves. And becoming a consumer can in turn lead to producing information – in the case of Picture Talk, starting a discussion oneself.

Blended Synchrony

Picture Talk implements a concept we call 'blended synchrony' (Erickson et al., 2006), meaning that the same application supports (near) synchronous and asynchronous interaction among participants. Picture Talk discussions persist over time, with remarks separated by seconds, minutes, days, or even months. Some discussions will feel quite immediate and rapid-fire, whereas others may be slower paced, or might be more like announcements than a true conversation. It just depends on the pattern of participation. Blended synchrony is useful in environments where communication needs to be close to real-time in some cases but can be asynchronous in others. Cultural and societal as well as pragmatic factors may come into play in deciding when and how to communicate with Picture Talk or any ICT application (Hudson, Christensen, Kellogg, & Erickson, 2002).

Navigational Affordances

As a conversational system, Picture Talk's audio postings could quickly grow to an unmanageable size as many users access the system. The problems with navigating a large amount of voice content are well known (Muller & Daniels, 1990). Time-varying multimedia do not offer the same navigational affordances as visual interfaces (Muller, Farrell, Cebulka & Smith, 1992). In Picture Talk, we have mitigated this problem by anchoring aural information to metadata that is made explicit through photos and graphics. For example, the author is depicted by the author's photo and the duration of the recording is shown using a graphic. Ultimately we would like users to be able to easily switch between visual or voice menus organized by authors, topics, time periods, locations, photos, tags (or other kinds of descriptive labels), and so on. A more complete solution will no doubt ultimately be needed.

Information Sharing

A number of researchers are looking at how to enable people in the developing world to share information using mobile technologies. For example, Steele and Tisselli (2006) describe three systems that enable BoP users to share information for mutual benefit. In the first, citizens documented cases of inaccessible spaces. In another case, messengers using motorcycles documented travel hazards. In a third case, nomadic Pygmies in the Congo Basin were provided with portable GPS-enabled PDAs with an iconic interface. They walked to various places and labeled trees and forest areas as food supplies, burial grounds, and so on, to prevent deforestation. In these cases, a map was used as the central visual device. Picture Talk could be extended to provide special support for maps or other types of special-purpose graphics, as we explored in the Tinker Talk design sketch.

Synchronization and Offline Use

Several applications in India and parts of Africa have been designed for mobile users without Internet access who periodically travel to areas with Internet access. For example, Prahalad (2004) reports on how ITC, one of India's largest private companies, developed a community of e-farmers with direct access to global prices, weather forecasts, farming techniques, et cetera, through centrally located Internet kiosks. Our Picture Talk implementation synchronizes content when an Internet connection is available. This provides a more flexible solution than a kiosk, where a single computer must be shared and users are unable to produce and consume information offline. However, more research is needed to understand when and where users might require connectivity and how this impacts the user experience of asynchronous conversation.

Conclusion

We are at an exciting point in the history of mobile computing. For the first time, the billions of people in some of the world's poorest countries have the promise of participating in the information revolution through mobile computing and communications devices. If successful, this could bring about positive social, political, and economic change in regions struggling with illiteracy, disease, poverty, natural disasters, oppression, and other challenges. Enabling ordinary citizens to become both producers and consumers of information could facilitate viral spread of critical information during crises, encourage broad exchange of ideas, connect experts with those needing help, strengthen social networks, and enable people at the base of the economic pyramid to become full participants in society and world economic markets.

We introduced Picture Talk, a software application we designed for use in environments with low literacy, limited Internet connectivity, and little familiarity with information services. Because basic mobile phones are the most common devices used by BoP populations, we have implemented Picture Talk on mobile phones. We are now investigating ways of providing access to some Picture Talk features on less expensive mobile phones using voice and text messaging. The limitations of using these devices to access rich structured content by users with limited literacy skills exposes human-computer interaction challenges that are key to enabling broad access to information by people in BoP populations.

Acknowledgements

We thank Ketki Dhanesha1 for sharing ethnographic studies of Indian villages and Nitendra Rajput, Arun Kumar, Amit Nanavati and all of the members of the IBM India Research Lab's Spoken Web team for helpful information about mobile phone use in India. We also thank John Ponzo for help with mobile phone platforms and Gail Hepworth and Steve Koeblen for helping us get started on projects in Africa.

References

Abraham, R. (2007). Mobile phones and economic development: Evidence from the fishing industry in India. *MIT Press Journal*, 4(1), 5-17.

Adeya, C. N. (2005). *Wireless technologies and development in Africa*. Unpublished report. Retrieved June 15,2009 at http://arnic.info/workshop05/Adeya_WirelessDev_Sep05.pdf

AfriGadget (2009) *Harnessing Personal Movement for Power in Rural Africa*. Retrieved on June 15, 2009 at http://www.afrigadget.com/2009/02/12/harnessing-personal-movement-for-power-in-rural-africa/

Appfrica (2008) *The current state of Internet penetration in Africa*. Retrieved on June 15, 2009 at http://appfrica.net/blog/archives/248

Danis, C. Bailey, M., Christensen, J., Ellis, J., Erickson, T., Farrell, R., and Kellogg, W.A. (2009) Social Computing Applications for the Next Billion Users. In *Designing Future Mobile Software for Underserved Users Workshop* at CSCW 2008.

Central Intelligence Agency (2009) The World Factbook. Referenced June 15, 2009 at https://www.cia.gov/library/publications/the-world-factbook/fields/2103.html.

Clark, H.H. & Brennan S.E. (1991) Grounding in Communication. In L. Resnick, J. Levine & S. Teasley (Eds.), *Perspectives on Socially Shared Cognition* (127-149). Hyattsville, MD: American Psychological Association.

EE Times (2008) *India's wireless network base will soon be the world's second largest*. Article by K.C. Krishnadas on March 24, 2008. Retrieved on June 15, 2009 at http://www.eetimes.com/news/latest/showArticle.jhtml?articleID=206905386

EnglishEnglish.com (2003) *What percentage of the internet is in English?* Retrieved on June 15, 2009 at http://www.englishenglish.com/english_facts_8.htm

Erickson, T., Kellogg, W. A., Laff, M., Sussman, J., Wolf, T. V., Halverson, C. A., & Edwards, D. (2006) A persistent chat space for work groups: the design, evaluation and deployment of loops. In *Proceedings of the 6th Conference on Designing Interactive Systems 06* (pp. 331-340) New York, ACM Press.

Ethnologue (2006) *Ethnologue, Languages of the World*. Referenced June 15, 2009 at http://www.ethnologue.com.

Frohlich, D. M., Rachovides, D., Riga, K., Bhat, R., Frank, M., Edirisinghe, E., Wickramanayaka, D., Jones, M., and Harwood, W. (2009) StoryBank: mobile digital storytelling in a development context. In *Proceedings of CHI 2009*. ACM, New York, NY, 1761-1770.

Gartner Group (2009) *Gartner Says Worldwide Mobile Phone Sales Declined 8.6 Per Cent and Smartphones Grew 12.7 Per Cent in First Quarter of 2009*. Press released dated May 20, 2009. Retrieved on June 15, 2009 from http://www.gartner.com/it/page.jsp?id=985912.

Gopakumar, K. (2006). E-governance services through telecentres: Role of human intermediary and issues of trust. In *Information Technologies and Development* , 4(1), 19-35.

Herring, S. C., Scheidt, L. A., Kouper, I., & Wright, E. (2006). A longitudinal content analysis of weblogs: 2003-2004. In M. Tremayne (Ed.), *Blogging, Citizenship, and the Future of Media* (pp. 3-20). London: Routledge.

Hollan, J., Hutchins, E., & Kirsh, D. (2000) Distributed cognition: toward a new foundation for human-computer interaction research. In *ACM Transactions on Computer-Human Interactions*. 7(2), 174-196.

Hudson, J.H., Christensen, J., Kellogg, W.A., & Erickson, T. (2002) I'd be overwhelmed, but it's just one more thing to do. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. (pp. 97-104) New York, NY: ACM Press.

Infotrend/CAP Ventures, (2004) W*orldwide Camera Phone and Photo Messaging Forecast: 2004-2009.* Retrieved on June 15, 2009 at http://store.infotrendsresearch.com/PhotoGallery.asp?ProductCode=MobileImagingStudy 10106London: Kluwer.

INSTAT (2007) *Size and Growth of Smartphone Market Will Exceed Laptop Market for Next Five Years*. Retrieved on June 15, 2009 from http://www.instat.com/press.asp?ID=2148&sku=IN0703823WH.

ITU (2009) *New ITU ICT Development Index compares 154 countries*. http://www.itu.int/newsroom/press_releases/2009/07.html (press release dated March 2, 2009).

Internet World Stats (2008) *Internet usage in Asia*. Retrieved on June 15, 2009 at http://www.internetworldstats.com/stats3.htm

Joshi, A., Welankar, N., Bl, N., Kanitkar, K. and Sheikh, R. (2008). Rangoli: A Visual Phonebook for Low-literate Users. *MobileHCI 2008*, September 2–5, 2008, Amsterdam, the Netherlands.

Kumar, A., Rajput, N., Agarwal, S., Chakraborty, D., & Nanavati, A.A. (2008). Organizing the unorganized: Employing IT to empower the under-privileged. In *Proceedings of the International World-Wide Web Conference* (pp. 935-944), Beijing, China: ACM Press.

Lampe, C., Ellison, N., & Steinfield, C. (2006) A face(book) in the crowd: social searching vs. social browsing. In *Proceedings of the Conference on Computer-supported Cooperative Work*, (pp. 167-170) New York, NY: ACM Press.

Morris, T. (2009) *Multimedia Systems*. New York, NY: Springer, p. 67

Muller, M. & Daniels, J. (1990) Toward a definition of voice documents. In Conference on Supporting Group Work. In *Proceedings of the ACM SIGOIS and IEEE CS TC-OA Conference on Office Information Systems*, (pp. 174 – 183), New York, NY: ACM Press.

Muller, M. J., Farrell, R., Cebulka , K. D. & Smith, J. G. (1992) Issues in the usability of time-varying multimedia, In M. M. Blattner and R. B. Dannenberg, Eds. *Multimedia interface design*, (pp. 7-38), New York, NY: ACM Press.

Olson, G. M. & Olson, J. S. (2000) Distance matters. *Human-Computer Interaction*. 15(2), 139-178.

OLPC (2009) *One Laptop Per Child*. Retrieved on June 15, 2009 at http://laptop.org

OSISA (2009) *Electricity for Africa?* Retrieved on June 15, 2009 at http://www.osisa.org/node/4164

Plauché, M., & Nallasamy, U. (2008). Speech interfaces for equitable access to information technology. Information Technologies and International Development, The MIT Press, 4 (1), 69-86.

Portio Research (2009) *Mobile Messaging Futures 2009-2013*. Retrieved on June 15, 2009 at http://wwww.portioresearch.com/MMF09-13.html

Prahalad, C.K. (2004) The Fortune at the Bottom of the Pyramid: Eradicating Poverty through Profits Pearson Education Inc., Wharton School Publishing, Upper Saddle River, New Jersey.

Rheingold, H. (1993). *The virtual community: Homesteading on the electronic frontier.* Addison Wiley.

RoR (2009) *Ruby on Rails*. Referenced on June 15, 2009 at http://rubyonrails.org

Samuel, J. et al. (2005) Mobile communications in South Africa, Tanzania and Egypt: Results from Community and Business Surveys. In *Africa: The Impact of Mobile Phones, The Vodafone Policy Paper Series*, No. 2, March 2005, pp. 44-52

Seshagiri, S., Sagar, A. and Joshi, D. (2007). Connecting the 'Bottom of the Pyramid' – An Exploratory Case Study of India's Rural Communication Environment. *WWW 2007, May 8-12, 2007,* Banff, Alberta, Canada.

Srinivasan, J. (2007). The role of trustworthiness in information service usage: The case of Parry information kiosks in Tamil Nadu, India. *Proceedings of the International Conference on Information and Communication Technologies for Development (ICTD).* (pp. 345-352) New York, NY: ACM Press.

Steels, L. & Tisselli, E. (2008), Social Tagging in Community Memories, in USA Proceedings of the 2008 AAAI Spring Symposium: Social Information Processing. Stanford University, ed., Menlo Park, CA: AAAI Press.

The Pew Research Center for the People and the Press. *The Pew Global Attitudes Project*, October 4, 2007, available from http://www.pewglobal.org.

Ushahidi.com (2009) *Crowdsourcing Crisis Information (FOSS)*. Retrieved June 15, 2009 from http://www.ushahidi.com/

UNESCO (2008) *UNESCO WebWorld News | Point of View*. Retrieved June 15, 2009 from http://www.unesco.org/webworld/points_of_views/tawfik_2.shtml

UNESCO (2009) *UNESCO Institute for Statistics*. Retrieved June 15, 2009 at http://www.uls.unesco.org

Yankelovich. W., W., Roberts, P., Wessler, M., Kaplan, J., and Provino, J. (2004) Meeting Central: Making distributed meetings more effective. In *Proceedings of the Conference on Computer-supported Cooperative Work, (pp*. 419-428) New York, NY: ACM Press.

Footnotes

1. Android is a trademark of Google, Inc.
2. G1 is a trademark of T-Mobile USA, Inc.
3. Asterisk is a registered trademark of Digium, Inc.
4. MySQL is a registered trademark of MySQL AB in the United States, the European Union and other countries.