Paul M. Aoki[1] · Rowena Luk[2] · Melissa Ho[3]

When mobile experience comes apart at the seams

Emerging markets infrastructure brings us back to “nomadic computing” in more ways than one

1 Introduction

There is a widespread notion that the “future” of mobile computing lies in emerging markets. This is partly meant as a business argument – for example, the markets for mobile handsets and wireless network equipment are not yet fully saturated, so market (unit) growth is “cheaper” there. However, it is often meant as a broader technology adoption argument as well. In countries with poor fixed infrastructure (of all kinds, including roads and electricity as well as telecommunications), it is natural that mobile platforms with onboard power storage and wireless connectivity would often be far more compelling to users than desktop platforms with neither.

The problem with the latter argument is that mobile computing platforms need infrastructure, too. Batteries and radios mean that the infrastructure doesn’t have to be deployed as densely, but the gaps in infrastructure relevant to mobile computing are often considerable. Indeed, they arguably go beyond where it is productive to apply Chalmers’ notion of “seamful design” [1]. Consider Ghana in West Africa. Ghana has a great deal of relatively low-cost power generation capacity, thanks to the Lake Volta hydroelectric system. Nevertheless, for a variety of reasons, rolling blackouts currently make power unavailable one day in five. Similarly, Ghana has a great deal of GSM coverage and relatively low prices for basic GSM service. However, for a variety of reasons, coverage is mostly concentrated in urban centers, and connections between networks (including between carriers as well as GPRS access to the Internet) are limited because gateway service is oversubscribed and expensive. Hence, access to infrastructure is not simply subject to occasional and irritating “seams”; the gaps between infrastructure availability may be quite wide (in terms of either time, in the case of power, or space, in the case of wireless). Often, then, the effect is more akin to 1980s ideas of nomadic computing – frequent and lengthy periods of disconnection from infrastructure, with frequent and irregular periods of connection.

What we would like to suggest is that such environments may lead us to think about “mobile experience” in a different way than the “always on…with seams” model to which mobile designers have largely become accustomed.

First, flexibility in what we normally consider “mobile experience” may be limited by disparities in the capabilities and expense of available platforms and networks.

Second, the relative capabilities and expense of the various computing and networking options can make hybrid application architectures and hybrid network architectures more compelling.

Third, a “gapful,” nomadic design approach may have implications for systems software infrastructure.

Only a few of the points made in this position paper are at all original, but all are relevant to user experience in emerging markets.

[1] Paul M. Aoki (aoki@acm.org), Intel Research, 2150 Shattuck Ave., Ste. 1300, Berkeley, CA, 94704-1347, USA.

[2] Rowena Luk (rowenaluk@gmail.com), Intel Research, 2150 Shattuck Ave., Ste. 1300, Berkeley, CA, 94704-1347, USA.

[3] Melissa Ho (mho@ischool.berkeley.edu), School of Information, University of California, Berkeley, CA, 94720-4600, USA.

2 Disparities

In terms of market penetration, mobile phones are clearly the most successful computing "platform" in emerging markets. While it is clearly useful to think of mobile phones apart from PCs, it makes less sense to think about mobile phones as a unitary platform in the same sense that one often thinks about Web browsers as a platform. Aside from the obvious software differences between phones, the disparities in basic capabilities (of hardware and software) and infrastructure (across space and time) are quite large.

We will sketch this disparity using some episodes drawn from fieldwork in Ghana over the 2006 holiday season. However, points along these lines have been made by many other researchers – in particular, anyone who follows Jan Chipchase's blog will find much here that is familiar.

Figure 1. Low-end phone (Accra, Ghana)

Figure 2. Camphones at fireworks show (Accra, Ghana)

2.1 Platform disparities

If one is deploying a vertical application (as is often the case in CS research), one is often in a position to deliver homogeneous hardware/software packages, or to specify that potential users must obtain specific hardware platforms to use the software. For example, microfinance applications based on CAM [4] have been successfully deployed in rural India even though CAM requires a camphone running a specific operating system.

It is also certainly true that, even in emerging markets, high-end phones are broadly available and have strong appeal in many demographics. We saw this in Ghana when we accompanied our guide to holiday services at a number of neo-Pentecostal megachurches in central Accra. (Ghana is famous for the public nature of its Christianity.) The end of the last service involved a fireworks show, and the blue glow of camphone screens quickly appeared all over as people tried to capture the event. This crowd was by no means the established elite – the elite do not generally attend the charismatic megachurches, preferring to attend more established churches – yet there was no shortage of RAZRs in this crowd.

However, it is also true that many phones used in emerging markets are either low-end or second-hand. The newly-manufactured low-end phones we bought in Ghana cost about US$30 (an amount, incidentally, that is several times the annual per capita spending on healthcare). However, these phones have only basic capabilities for voice calls, texting and addressbook management – no color screen, no camera, no complex functions and applications. Used phones sold on the street are typically older models, whether recycled from overseas markets or resold within the local market (or locally stolen).

It is therefore no surprise that mobile services in emerging markets focus so strongly on SMS, whether using the "command-line"-style of sending text messages manually or using "DOS menu"-style SIM Toolkit applications (as the MIT/Nokia Project EPROM is currently doing in Kenya). After all, doing otherwise would greatly reduce the number of potential users. This is particularly true for applications that are truly aimed at users with limited economic resources. To avoid excluding users, basic services (such as rural information services and mobile banking) must continue to work when only basic capabilities are available. Clearly, however, this limits how "rich" a mobile application can be.

2.2 Network disparities

The GSM Association's public network coverage goals are expressed in terms of the percentage of the world population covered by wireless service. It is important to note that a long-term goal of (say) 90% coverage of the world's population is very different from a goal of 90% world coverage, and this difference is relevant to design.

As can be seen from Table 1, the bulk of Ghanaians are already covered by wireless networks in a way that was never realized by fixed telephone networks. (This data is from a 2006 World Bank report – more recent but less "authoritative" data indicates that mobile subscriptions are now 2-3 times higher.) In Figure 4, we see a small shop in a shack in the foothills of southeastern Ghana that has been labeled with the proprietor's mobile number.

Figure 3. Rural landscape (northern Ghana)

Table 1. Telephone access (per 100 people), 2006.

Indicator	Ghana	USA
Telephone main lines	0.15	60.6
Mobile subscribers	12.9	68
Population covered by mobile telephony	69	99

Figure 4. Mobile number address (southeastern Ghana)

However, this high population coverage is largely a function of the country's population distribution. Traveling outside of the urban centers, one finds that wireless coverage is quite spotty. Indeed, in sparsely-populated rural areas (Figure 3), coverage comes and goes quickly as one drives through towns. Furthermore, as previously mentioned, sections of the infrastructure are underprovisioned and therefore unreliable.

When the ability to send text messages or make voice calls, let alone send packet data, can be intermittent, the very idea of networked applications comes into question.

3 Hybrid applications and networks

Because of these large disparities, there would seem to be little "room to move" for mobile experience design. Even though the individual users are themselves so clearly focused on the importance of the mobile, however, it may be worthwhile to stop thinking of mobile experience as being limited to what happens in the individuals' handsets.

The more conventional "hybrid" or "multi-channel" view of Web service designers seems like a relevant alternative. Just as social networking websites increase their stickiness by sending alerts and notifications through email and SMS, drawing users back into the PC-oriented Web site, designers for emerging markets should consider applications in the context of the entire ecosystem of platforms that are available to users and consider an individual's mobile in the context of that ecosystem. "Available" can be meant in terms of borrowing or reselling. For example, in parts of the world where schools, homes and workplaces have no PCs or Internet connectivity, Internet cafés are often available. The Grameenphone reseller model for access to basic mobile service is being copied worldwide. There is no clear reason to believe that access to higher-capability mobile platforms is fundamentally different from (more expensive) PCs or (less expensive) mobile service. What is necessary is for someone to be able to obtain the capital necessary to gain access to network services.

3.1 A suggestion: doubly-hybrid systems

As an example of what we mean by hybrid design, consider an asynchronous medical teleconsultation system that we are currently deploying in Ghana. (Figure 5 is representative of the trial deployment but is not literal – there are 7-8 candidate sites for the trial deployment, and the end-goal is to extend the system to the rural north.) For the most part, this is very straightforward Web-based system. Doctors use PCs to connect to a Web server; doctors who wish to request an expert consultation (e.g., general practitioners at rural hospitals) enter case details into the Web application, and doctors in a position to provide consultation (e.g., specialists in urban centers or abroad) can reply by adding comments. PC displays are useful in this context because wants to be able to browse the case lists, view data for individual cases (including photos and X-rays) easily, and enter extensive comments.

However, two factors distinguish it from the many Web-based telemedicine systems that are available:

First, hospital sites may be entirely disconnected from the Internet (many have no Internet connection) or periodically disconnected (by power or network outages). Therefore, doctors necessarily interact with local Web servers; doctors can enter new cases and modify existing cases at any site, and changes are propagated to all sites. A distributed, replicated database is used to propagate updates between the sites and a store-and-forward delay-tolerant networking (DTN) layer [3] is used to transport the update streams. DTN works whether the bits are transported over IP using low-cost microwave links [1], a WiFi-enabled vehicle (as in Daknet [5]) or Symbian-based smartphone acting as an intermittently-connected DTN node, or a "sneakernet" of dumb USB storage devices carrying formatted DTN "bundles" between DTN nodes. The application is designed in a conflict-free way so that updates can be applied as long as updates from a single site are delivered in-order.

Second, for all of the reasons we have discussed as well as others, frequent workplace use of PCs is typically not in the doctors' existing work practice. However, occasional use of email (usually webmail rather than organizational mail) and mobile phones is. Therefore, their local Web servers can use both email and SMS alerts to "ping" doctors when new cases arrive for their attention. For example, SMS may be received when a consultation is being requested from the doctor, or when a new reply to a consultation requested by the doctor arrives.

This is a hybrid system in two senses. First, the overall design is based on a hybrid application architecture that recognizes that users are "locally mobile" and not tied to desktop PCs. Hence, even though most of the system is non-mobile, mobile reminders play an essential role in bringing nomadic users back into the system. These rely only on the most common available capabilities (SMS). Second, the system relies on hybrid networks rather than an abstract IP or wireless "cloud." The use of DTN exploits physical mobility to bridge gaps in Internet connectivity; nomadic users themselves may carry data as they move between sites.

In both senses, however, the system is explicitly "gapful" – the designers' emphasis is less on a "good mobile experience" in the conventional sense and more on a (sustainably) mindful role for the user in the operation of the system.

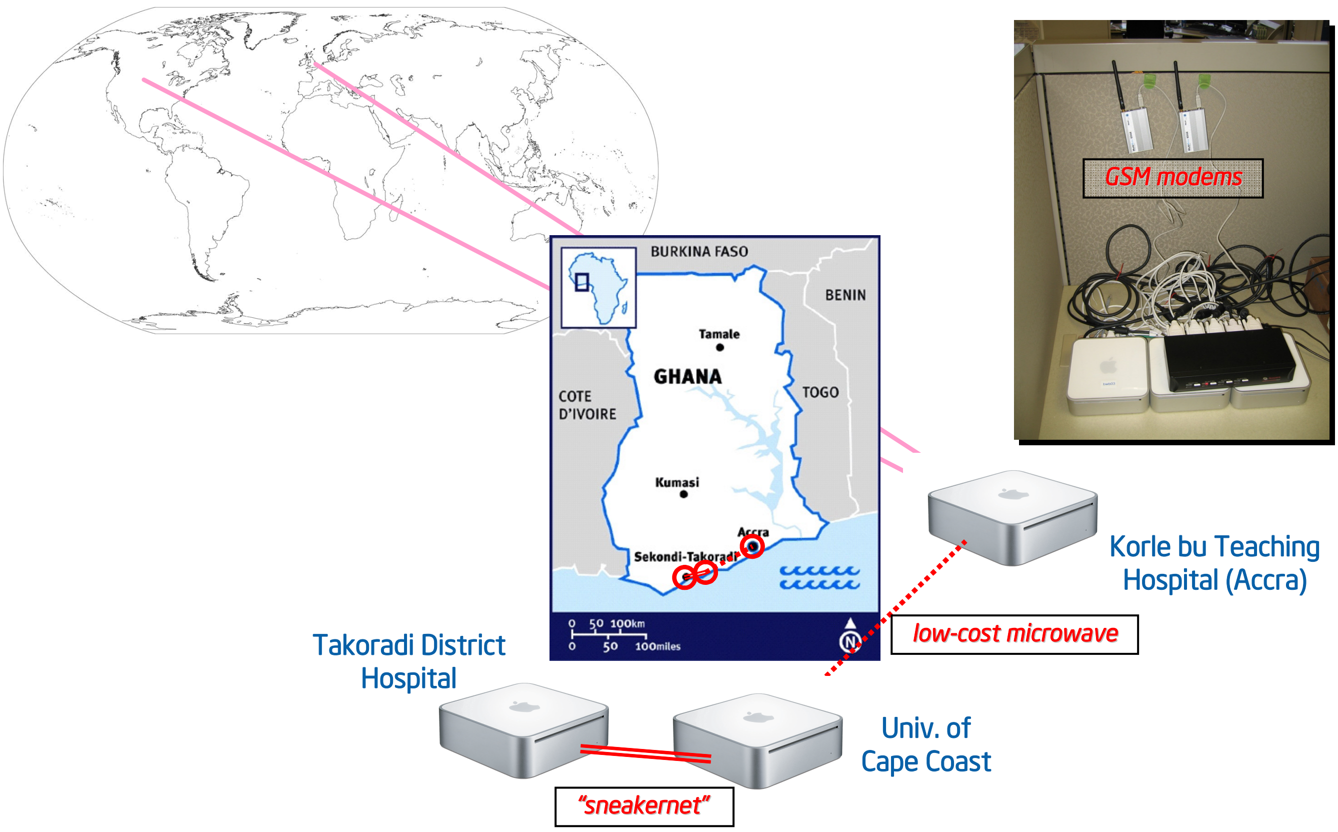

Figure 5. Asynchronous medical teleconsultation system using hybrid networks

3.2 Another suggestion: software for extreme asynchrony

As the hardware base gradually improves, the need to address the minimal hardware configuration (e.g., SMS) will decline and it will become more important to enable applications to work asynchronously. What is the right level of abstraction for software that should underlie asynchronous applications? DTN and related networking systems provide only a basic messaging/routing framework. Asynchronous middleware like Google Gears may be closer to what we want, but a generic application middleware that requires live connectivity and an embedded SQL engine may be too heavyweight – cost-sensitivity will continue to be an issue.

It would be useful to have a midpoint between a message transport system and a replicated database engine. One such midpoint is the class of message queuing (MQ) systems. MQ systems are mainly used in enterprise environments (notably in workflow and transaction processing) and are basically frameworks that enable packaged-up, transactional RPCs to be asynchronously processed by collections of servers. Hence, unlike existing "edge" facilities like Google Gears or MSMQ smartphone clients [6], or existing transport facilities like DTN, a MQ infrastructure would be a mid-weight way to enable phones to manage data and act as in-network processors.

Conclusion

We have suggested that appropriate mobile experience design for emerging markets is not simply matter of producing "minimal" services based on basic phones. This is certainly important and relevant (and will remain so for some time to come), but one can go further. First, we have argued that designers should imagine mobile phones and mobile services in the context of larger ecosystems of available computing and networking resources – it is this "mobile experience" that should be considered rather than that of the mobile interface in isolation. Second, in spite of the growth of mobile networks, we have argued that the need for asynchrony to bridge gaps will remain a fact of life and there are mid-weight networking models, software and services that we expect can be usefully extended to make asynchronous applications more flexible.

References

1. Brewer, E.A., Demmer, M.J., Du, B., Ho, M., Kam, M., Nedevschi, S., Pal, J. Patra, R.K., Surana, S., and Fall, K. The Case for Technology in Developing Regions. *IEEE Computer 38* (6), 25-38
2. Chalmers, M. and Galani, A. Seamful Interweaving: Heterogeneity in the Theory and Design of Interactive Systems. *Proc. ACM DIS 2004*, 243-252.
3. Fall, K.R., A Delay-Tolerant Network Architecture for Challenged Internets. *Proc. ACM SIGCOMM 2003*, 27-34.
4. Parikh, T.S. and Lazowska, E.D. Designing an Architecture for Delivering Mobile Information Services to the Rural Developing World. *Proc. WWW 2006,* 791-800.
5. Pentland, A.S., Fletcher, R., and Hasson, A. DakNet: Rethinking Connectivity in Developing Nations. *IEEE Computer 37* (1), 2004, 78-83.
6. Redkar, A., Walzer, C., Boyd, S., Costall, R., Rabold, K., and Redkar, T. *Pro MSMQ: Microsoft Message Queue Programming*. Apress, 2005.