Designing Healthcare Systems in the Developing World

The role of computer science systems research

Sonesh Surana
sonesh@cs.berkeley.edu
UC Berkeley

Melissa Ho
mho@sims.berkeley.edu
UC Berkeley

Rabin Patra
rkpatra@cs.berkeley.edu
UC Berkeley

Lakshminarayanan Subramanian
lakshmi@cs.nyu.edu
New York University

ABSTRACT

Developing regions account for a significant fraction of the world's health problems and are in a dire need for an improved healthcare delivery system. Achieving this goal is a very challenging task given the limited financial constraints and the scale of the problem. In this paper, we argue that computer science systems research will have a significant role towards addressing some of the pressing healthcare problems in developing regions. To motivate the need for systems research, we describe a spectrum of application areas where computer systems can aid in improving systemwide healthcare delivery and articulate the corresponding set of research challenges associated with each area. We also describe some of the important deployment challenges in developing countries that one would face while translating systems research solutions into real-world healthcare systems.

1. INTRODUCTION

Providing adequate health-care in developing regions where most of the world's population resides is one of the biggest challenges facing us today. Developing regions bear most of the global burden of disease, accounting for 95% of the 40 million HIV cases worldwide, 90% of TB cases and deaths, and 96% of the malaria cases [1]. Child mortality in the least developing countries is 278 per 1000 live births compared to 6 per 1000 live births in industrialized nations [2].

While the number of deaths is an alarming number, only a small set of health conditions are the underlying root causes for a large fraction of these deaths. Ninety percent of under-five child deaths in developing regions result from maternal and perinatal mortality, vaccine-preventable disease, acute respiratory infection, diarrhea, malnutrition, malaria, TB, and HIV/AIDS [3]. Most of these conditions have effective interventions and are preventable. Indeed, the developed world has almost 100% coverage of childhood vaccinations, skilled attendance at birth, and generally broad access to treatment [4]. Despite the availability of such interventions, limited financial resources combined with the overwhelming scale of the problem impede efficient delivery of health interventions in developing regions.

The question remains as to what we as computer scientists can do to address, if not at least partially alleviate this problem. In this position paper, we argue that *appropriate* computer science systems research[1] can make significant inroads into this problem. The primary goal of this paper is to highlight a spectrum of healthcare scenarios that arise in developing countries and showcase where systems research can make a difference. To the best of our knowledge, this paper is one of the first concerted efforts to identify and articulate systems research problems that arise in improving healthcare delivery in developing regions.

To drive home the need for system research in this area, we next address two questions in greater detail: (a) where computer systems can help in improving healthcare delivery in developing regions; and (b) how systems research in this context is different from traditional systems research.

1.1 Where can computer systems help?

The current status of healthcare delivery in many non-urban regions in developing countries has been poor; in fact, rural places in sub-Saharan Africa lack several basic healthcare services including rural clinics and medical expertise. Given such extreme conditions of the current state of affairs, it is natural to assume that computer systems have little role to play in healthcare delivery in such regions.

Contrary to this conventional wisdom, we argue in this paper that computer systems *can* have a substantial role to play in improving the current state of affairs. The primary arguments as to why this is the case are four-fold. First, despite the shortage of doctors and nurses, rural regions in developing regions have a body of informally trained health workers who perform most of the healthcare tasks. This is a workforce that can be tapped for efficient healthcare delivery by augmenting their capabilities through diagnostic and clinical decision support tools. Training health workers remotely can benefit from *network connectivity*; recent advances in long-distance wireless technologies can potentially be a building block for a low-cost infrastructure. Network connectivity also enables remote diagnosis to rural regions without having the need for sick patients to travel to close-by cities. Second, despite a concerted effort to supply vaccines and drugs to these countries at low costs, these resources often never reach their intended recipients. The current medicine distribution system in these countries is grossly inefficient; here again, computer systems can help in better *supply chain management* of health resources. Third,

[1]In this paper, we use the term "systems" to refer to computer systems.

WWW2006, May 22–26, 2006, Edinburgh, UK.
.

there is ongoing work in developing low-cost medical devices, an example of which is a point-of-care multiple disease diagnostic device for developing regions [5]. In this context, integrating these medical technologies with computer systems is essential to make them usable and accessible to the computer-illiterate but resourceful health workers. Finally, based on a recent health-study, by the year 2020, 7 out of 10 deaths in developing regions will be due to non-communicable diseases (a.k.a. "lifestyle" diseases) with chronic symptoms that will require more involved interaction and tracking of patients for effective prevention and control [6].

1.2 The need for systems research

While computer systems do hold the promise of improving healthcare in developing countries, the role of computer science has been largely limited to IT systems implementation through *technology transfer* from developed countries. One successful project has been the TEHIP [7] project in Tanzania, which analyzed medical data from two district to determine the correct set of medicines to be ordered to prevent child-mortality. This project reduced child mortality by 40%.

Still, the biggest problem with technology transfer based solutions is that while they are successful in small-scale pilot projects, they *fail to scale to large-scale deployments i.e.,* a replication of the same solution on a larger scale has been a failure. Products built on developed world assumptions of stable power, reliable network connectivity, sophisticated users, and higher levels of affordability tend to fail in the relatively chaotic operating conditions of developing regions.

In this paper, we argue the need for systems research towards the development of appropriate systems solutions for healthcare delivery that can achieve large-scale deployment in developing countries. We now describe five distinguishing factors between the type of computer systems problems that arise in the healthcare domain in developing countries in comparison with traditional systems research:

Need for low-cost network connectivity: Several rural regions in do not have any form of network connectivity largely because existing networking technologies are not economically viable [8]. In the healthcare space, the presence of network connectivity can aid in a variety of areas: remote diagnosis, tracking medicine distribution, mobile hospitals, and disease tracking. In Section 4.1, we describe some new advances in long-distance wireless technologies which can potentially alleviate this problem. Perennial connectivity is not always essential; one can make significant progress on healthcare diagnosis with intermittent network connectivity as we elaborate in Section 4.1.1.

Distributed information systems are hard to build: Building distributed information systems to support the collection and distribution of medical data from different sources is a very challenging problem. Traditional distributed systems solutions break down in the face of chaotic operating conditions in developing countries with unreliable power, lack of good network connectivity, and usability challenges. Additionally, the need for low cost solutions makes the task of building robust distributed information systems even more challenging. Our vision is to create a computing and communications substrate over which health applications can function even in unexpected operating conditions.

System awareness of new medical technologies: Traditional systems research has largely been application agnostic. However, in developing countries, we require health workers to be able to use new medical technologies. This would require the design and development of a *simple* and *accessible* user interface to act an intermediary for potentially complex underlying medical technologies.

Design for developing regions must consider cost: Traditional systems research is often not driven by economics. However, given that many of the rural poor in developing regions make less than $4/day, economics is one of the key factors that determines the adoptability of new technologies. Apart from low-cost connectivity solutions, computer systems can play an important role in reducing the cost of healthcare including: (a) building an efficient supply chain management for medicine distribution, (b) providing the ability to do remote diagnosis (significant savings in travel costs), (c) remote medical education of health workers, (d) disease tracking and epidemiology surveillance.

Social considerations are important: Every rural region is associated with its set of local social cultural practices. A technology that is not tailored to these social considerations is doomed for failure due to lack of adoption. These social considerations have implications on how we design computer systems. For example, the literacy rate in many of these regions is very low; hence, a text-based data-entry based user interface may be inappropriate for such regions.

The rest of the paper is organized in the following manner: In Section 2, we provide some context about a few ground realities and preconceived notions with respect to working in developing regions. In Section 3, we present a set of healthcare applications that has the potential to improve system wide healthcare delivery. Then, we describe the corresponding systems research problems required to enable these applications in the face of ground realities in these regions. Next, we discuss deployment issues in Section 5, and we conclude with a discussion in Section 6.

2. REALITIES AND MISCONCEPTIONS

In order for systems research to make any tangible impact in developing countries, it is essential to understand the ground realities involved in addressing problems of development successfully. Without experience of working in such a vastly different culture context, it is also natural that we as systems researchers (in developed countries) may have certain misconceptions about research in developing countries. In this section, based on our limited experiences, we pinpoint some of the important ground realities and misconceptions that we have come across.

2.1 Ground Realities

Developing regions face a number of challenges that limit the real impact health care services can offer. The majority of these challenges emerge from the interplay between constraints related to human resources, knowledge base, physical infrastructure, communications infrastructure, and financial capital.

Human Resources: There is an acute shortage of health care workers in developing regions. For example, Rwanda has 1 physician for 50,000 people while Italy has 6 per 1000 [9]. In addition, the distribution of health care workers is skewed with only a small fraction present in rural areas where there is a greater prevalence of diseases. Plus, as in the Philippines, where many doctors re-train themselves as nurses for lucrative markets abroad [10], many health workers leave abroad for better economic opportunities. Thus, the existing staff of health workers in rural areas is lesser experience and overburdened.

Knowledge Base: In many developing regions, availability of accurate and appropriate medical information is a serious problem. Firstly, lack of a proper infrastructure prevents the popularization of even well-known medical practices and cures for common health conditions. As a result, health workers and even doctors in rural areas might be either unaware of or not properly trained in basic procedures. Secondly, access to healthcare reference information is not widespread, particularly because the price of medical journals is beyond the financial means of the libraries. Most healthcare information, even if available, is "western-centric" (donated material) and may not always be relevant, suggesting tests or medicines that are not available locally. In addition, traditional medical knowledge often not organized in any repository that is easily accessible.

Physical Infrastructure: Health sectors in developing regions generally operate in the presence of poorly developed essential infrastructure such as roads, clean water, and reliable electricity. Under these circumstances it is not easy to transport crucial medical supplies and personnel to appropriate areas in a timely manner. Without clean water it is very challenging to maintain clean health facilities. Unreliable power not only causes equipment downtime but also damages equipment in the long run through periodic blackouts or prolonged brownouts.

Communications Infrastructure: Very often connectivity for voice (let alone data) is non-existent, unreliable or very expensive. In general, the infrastructure is a mix of varying degrees of postal, TV, radio, telephony (fixed and mobile), and Internet coverage. The prospects for any kind of supply chain or logistics management, remote health worker collaboration and training, or provision of telemedicine services over such limited communications scenarios are quite limited.

Financial Capital: People in developing regions live on a few dollars a day. The existence of such a low purchasing power make it hard to cover the costs of treatment. Moreover, financial resources are distributed improperly in developing regions; very little trickles down to rural health care facilities from provincial hospital due to mismanagement, corruption, or general lack of distribution infrastructure

2.2 Potential misconceptions

We expect that, like ourselves initially, many researchers from the developed world undertaking research in developing regions may have some preconceived notions of sufficient approaches without understanding some ground realities. These notions are not entirely untrue but not entirely true either, so we attempt to bring some balance in viewpoints which may have useful technology implications.

We cannot do much: With such widespread constraints set in the backdrop, it seems that there is very little systems research can do to alleviate the situation. However, we argue that the time now is ripe for such technology forays owing to opportunities in shared computing, availability of low cost networking hardware, and access to microfinance which makes creation of sustainable business models around technology a possibility [11]. As illustrated in Section 3, there are several scenarios in which computer systems can play an important role in improving the health delivery system in these regions. Notably, systems can contribute heavily in the areas of: (a) augmenting the capabilities of a health worker by providing clinical support tools, automated diagnosis support *etc.*, (b) using low-cost networking technologies to provide the ability to perform interactive telemedicine, disease tracking, remote storage and access of health information.We describe several health application scenarios in more detail in Section 3.

It's just tech transfer: Transferring first world technology directly to developing regions setting is never the answer in general. Solving the problem in developing countries in a cost-effective manner is much harder than mere technology transfer. First, world information systems projects in developing regions mostly fail (totally or partially) because of a set of "design gaps" [12] in between the idealized design of the system and the challenging realities in which the system must operate. For instance, It's hard to design for transparent transfer of data over a network in which the links can down because of power outages or forced shutdowns for cost savings. Also, it's hard to design solutions that assume no prior expertise with computers. Remote management and troubleshooting of hardware and software needs to be bulletproof since frequently the users, who have no prior experience with operating systems, will not be able to identify the problem. It is crucial to co-design and co-deploy the system with the end users during the testing phase before clinical deployment.

Patients don't know what they're suffering from: It is a common misconception that patients are unaware of their medical condition. Most of the diseases are well-known and, in many cases, the patient is aware of his/her ailment. Quite often, clinics are short of basic medical supplies (like oxygen cylinders, bandages, sterilized gloves *etc*) and thereby cannot deliver effective treatment. The problem is that we, as technologists, tend to focus more on diagnosis than on treatment; the patients however care more about treatment than diagnosis. This is not to say that we need to reduce our focus on diagnosis, but that we need to raise our efforts in making treatment accessible. The biggest hindrance to effective treatment is the lack of health resources in a locality due to poor supply chain management of medicine. Any low-cost technological solution to address the inefficiencies of health treatment including improving the supply chain of medicine can have a substantial impact in improving overall healthcare delivery.

Only doctors will use the system: In most developing regions, there in am imbalance in the number of doctors between urban and rural areas. Indeed, doctors form just a small fraction of the total health care work force. Most of the front-line health workers comprise a great variety of people with varying degrees of expertise and experience. These workers may have limited training such as community health workers, long experience such as nurses practicing for 30 years, limited experience such as first year medical students, or alternative medicine expertise (traditional healers). These workers are typically the front-line response to public health problems. Any technology solution must take into account their differing needs, expertise, and operating conditions. To a rural health worker, the usage of any computing device should appear transparent whether they are out on the field out of communication coverage or in a clinical setting within communication range.

3. APPLICATION AREAS

In this section, we present some of the several application areas where systems research can make an impact on improving healthcare delivery.

3.1 Remote Diagnosis and Consultation

The single most scarce health resource in developing regions is trained personnel [13]. According to the WHO statistics, Rwanda has 1.87 doctors and 20.97 nurses for every 100,000 people. Assuming a perfectly even distribution of population across the entire country, this amounts to less than two doctors for every 300 square kilometers [14]. Since, in fact, rural areas are significantly less dense than that, it is often difficult for residents of rural villages to access health care, because the nearest health facilities with available personnel are too far away, requiring a day or more of travel. Likewise, it is not cost effective for doctors to travel to patients in rural areas, because the travel time involved would be better used on treatment and diagnosis of local patients.

Telemedicine and remote diagnosis can partially alleviate this problem. Rural healthcare is offered in the form of health clinics run by health workers who are not well trained in medical practice. Telemedicine offers the capability for a rural health worker to remotely consult medical experts in nearby cities without having the need to travel there. Today, for many diagnostic and lab-based tests, remote diagnosis is the norm. X-rays and blood samples are taken and sent to a specialized lab that will run the relevant tests and send a report back to the doctor. In developing regions, the same model works, with the caveat that the test samples may have to travel longer distances to be processed, which would require portable refrigeration, or that lab facilities may not even be available in the local vicinity. In the Solomon Islands, they have set up a pre-processing laboratory, in which the blood samples are treated and prepared. Photos of the slides are then posted to a shared content management site, iPath, hosted in Sweden, where specialist doctors are able to make a diagnosis based on the high-resolution pictures [15, 16].

Telemedicine and remote diagnosis can enable a more efficient use of the time of doctors and health workers, with the hope of having the same level of effectiveness in medical diagnosis and treatment. The goal is to provide basic medical diagnosis training to health workers and use remote diagnosis to consult with medical experts for more complicated medical conditions. This model of remote diagnosis can benefit from the current advances on low-cost networking technologies [17, 18] which can enable health-workers to communicate with doctors remotely. An extended version of this model is *real-time telemedicine* where a patient in a rural health clinic can have a live consultation with a doctor remotely where the medical records of patients (*e.g.,* medical images) are transmitted over the network in real-time for diagnosis purposes.

Much research needs to be done in order to make existing communications technologies appropriate for real-time telemedicine. This translates to low cost alternatives to traditional access technologies like fiber. There is also a need for research on protocols, applications, and compression techniques that make the most effective use of available bandwidth. Another problem in this context is reliable power - frequent power outages can result in situations in which desired communications endpoints are not simultaneously available. In these cases, we require research on store-and-forward mechanisms which can enable the doctor to assess diagnostic information at their convenience, independently of the data acquisition during the patient visit.

3.2 Epidemiological Data Collection

Epidemiological work involves extensive data collection and field analysis, ultimately using complex statistical techniques to determine patterns and predictors for the spread of disease. Health practitioners use this data to create informed plans for medicine distribution and implementation of region and disease appropriate interventions towards the control of outbreaks. The field of computational epidemiology focuses on the application of computer science (such as Bayes networks [19]) and geographical information science tools towards the analysis of disease prevalence data. However, the space remains open for the design of a systems architecture for support of massively distributed data collection and aggregation.

Data must be collected from potentially remote or hard-to-reach locations, and eventually returned to a preferably highly available central repository for analysis. One could envision a field worker, rather than returning to the distant local capital each time, returning to a village kiosk to upload data to reliable storage. The data from a number of field workers is then transferred from the village via satellite or phone (overnight to avoid peak rates) to the central repository in the nearest city. This data collection architecture must be able to support a variety of connectivity models, from completely disconnected operation, to a cost-conscious scheduled transaction, while simultaneously supporting the data storage and synchronization requirements resulting from multiple people collecting epidemiological data. In addition, since it cannot be expected that these health workers will simultaneously be network administrators, the underlying operation of this system must be seamless and invisible to both the data collectors and the people using the data. Finally, there needs to be a registry or set of registries that support efficient statistical analysis of the data stored in them [20].

3.3 Supply Chain and Logistics Management

One of the contradictions of supply chain management is that it costs more to deliver to rural areas, where people have less income. Distribution costs increase with distance - so as communities become less dense there is more per capita overhead. However, Janini [21], a non-profit specializing in family planning and the delivery of the associated products and services, has demonstrated that once the supply chain problem has been addressed, rural communities can benefit greatly from the availability of the services, and are also willing to pay enough to make their franchises financially sustainable.

At the same time, general shortage of medical supplies leads to corruption and mismanagement. Urban areas hoard medications because of general shortage, so rural centers never receive requested amounts of supplies. As a result, medications are often procured through bribes, or arrive only after they have expired. More alarming, however, are the WHO's estimates that up to 25% of the medicines in developing nations are counterfeit or substandard, further exacerbating the health problems [22].

Projects like Janini [21], E-Choupal [23], and Unilever's Shaktis [24] have been largely successful in providing supply chain and logistics services, eliminating much of the corruption through the implementation of strict auditing of financial and inventory controls. However, there is always more work to be done in improving these mechanisms, as well as service and cost efficiency.

Distributed data systems can be used for synchronization of current inventory status in various (potentially disconnected) locations. There is an additional problem of synchronization without concurrent connectivity; however the problem can be simplified with the assumption of a single supply source, or a set of well-connected supply sources. Another aspect of this problem is the development of affordable and easily maintainable communications infrastructure, perhaps leveraging low cost mechanisms like SMS. Assuming that the costs can be brought down sufficiently, there is also space for the use of GPS data or RFIDs in tracking the flow of inventory. As these systems advance in sophistication and scale, it will also be necessary to develop appropriate privacy and security mechanisms for protection of transaction integrity. There may also be an opportunity for the use of machine learning and Bayes networks for anticipating shortages and pre-supplying the health clinics.

3.4 Cost-effective Clinical Equipment

Rural health clinics can benefit a lot from low-cost medical equipment which can aid in health diagnosis and are simple to use. Rural health care clinics necessarily serve a much smaller population; hence, the cost of equipment in these clinics is a critical issue. Cost manifests itself not only in the original capital investment, but also in the maintenance and repair of the devices, the lifetime, and in the cost of training health workers to use the equipment.

Currently, the most robust and easy-to-use equipment are generally the traditional mechanical tools (i.e. fetal-scope, stethoscope, ear-scope, orthoscope, and patella hammer). Even so, there are opportunities for electronic devices to play a role, provided that they are durable and can be made inexpensive (perhaps leveraging economies of scale). Electronic devices with automatic readers (e.g. blood pressure monitors) require less expertise and training than the traditional blood pressure cuff. Automatic entry of the data into the associated medical record via wireless could reduce transcription errors, improving the quality of statistics. However this incurs significant extra cost - which may not be justifiable for the perceived benefit. A lower-cost alternative would be a combination camera-phone with character recognition or 2D barcodes [25].

In addition, rural clinics also need rapid-response diagnostic test kits, such as pregnancy [21] and malarial tests, both for determination of immediate appropriate treatment, and for accurate epidemiological data collection (See Section 3.2). If made sufficiently low cost, these devices can play a significant role in reducing the costs of health care, by providing the health workers with the information tools to do better differential diagnosis, resulting in less waste on ineffectual treatment.

Hospitals in developing regions also have design requirements that differ from those typically pursued in developed countries. X-ray, radiology, and ultrasound equipment remain too expensive, in part because traditional research focuses on increasing resolution and dimensionality. The genuine need is for equipment that is "good enough" in terms of image quality, but less resource intensive, more portable, and less sensitive to environmental conditions.

3.5 Continuity of Care

More often than not, a single clinical visit is insufficient for treating a patient's health problem. Many health conditions, especially chronic conditions like AIDS, require a consistent regimen of treatment over multiple visits. Basic preventative care also requires persistent monitoring of basic health parameters like heart rate, blood pressure, cholesterol, weight, and sugar levels.

However, one of the prerequisites for managing the care of a patient across visits is often missing in developing regions: the medical record. In countries where many individuals lack unique identifiers, from national identity numbers to addresses [2], it becomes difficult to maintain reliable records on a given patient. Furthermore, frequent migration [26] and "clinic-hopping" mean that patients must be tracked across multiple locations. Indeed, this is often deliberate - due to the stigma of AIDS, patients will avoid going to the same clinic twice, so they are not recognized. This leads to problems in which the same tests and the same ineffective medications are tried on a given patient; a waste of both time and scarce resources. Thus, there is a need for effective ways to manage care records and to synchronize data from varied locations, in addition to the design of accessible user interfaces for data entry and recollection.

Rural health infrastructure typically consists of several health clinics, with varying resources in terms of power, information connectivity, and general reachability. They generally share a hospital in an urban area, which provides the health workers and supplies for each clinic; as a result, the health system is also constrained in terms of cost. At the same time, it is necessary for each of these clinics to have access to medical records in a timely manner.

Thus the technical challenge is to design a portable low cost and low power device that uses minimal, preferably pre-existing, infrastructure (e.g. mobile wireless) to synchronize care records. To minimize the training required, addressing the issue of unavailability of health workers, the user interface must be accessible in terms of both data entry and recollection. To account for the "clinic hopping", a potential solution to be investigated would be a machine-readable paper audit trail [25], or some other inexpensive and durable mechanism like smart cards for patients to own their personal records. As with the other scenarios, the new technologies introduced to achieve this intervention must be changes that can be readily assimilated into existing society, both socially and economically.

4. SYSTEMS CHALLENGES

In the previous section, we looked at some important healthcare areas and briefly touched upon the technological challenges that are presented in each of these areas. It is

[2]Urban slum dwellers and rural villagers will often use semi-permanent landmarks to describe where they live.

clear that developing and deploying a healthcare infrastructure in developing regions presents very different issues than the ones faced in the developed world. In this section, we take a detailed look at the areas of computer systems that are relevant for healthcare in developing regions and identify the different challenges and possible future directions for systems research.

The key areas for computer systems research are low cost communication that includes both networks that provide continuous connectivity and the ones that provide only intermittent store-and-forward data transfer, distributed data storage, and coding and compression algorithms for video, audio and images.

4.1 Low-Cost Communication Infrastructure

Communication infrastructure, essential for any comprehensive health services system, is especially important for developing regions where the current physical and communication infrastructure is very inadequate.

While fiber or satellite communications can provide good connectivity as well as high bandwidth connections, more than 70% of the capital cost of this infrastructure would be in the a communication infrastructure is in the access network, not in the backbone. In the rural areas of developing countries, the low density of users results in low combined purchasing power, such that their high capital cost cannot be justified for the current low penetration. Likewise, cellular technologies are designed for high population densities, and thus the recent growth has been mostly limited to urban areas.

Traditionally, satellite-based systems (e.g. VSAT) were the first choice for covering remote areas. In practice, however, their use is unsustainable; initial and recurring costs are too high and the bandwidth limitations are too small to afford sharing of these costs over larger user bases. A new alternative is WiMax, the IEEE 802.16 standard for wireless metropolitan-area networks [27]. The technology is designed to accommodate large number of users at ranges more than 50km and makes efficient use of the available spectrum. Currently available WiMax equipment is still very expensive; it is not clear if it will enjoy sufficient market success to bring costs within the range of affordability for developing regions. As a result, there is a need to look at novel technologies or innovative adaptations of current technologies and develop new models on deployment that combine continuous and intermittent connectivity.

4.1.1 Intermittent Networking

We have previously seen that a number of health related applications do not need continuous, real-time, synchronous communication. Services like remote radiology and pathology where medical images have to be transferred from remote clinics to experts in hospitals do not need interactive communication at all. Similarly, transfer of medical records and health information collection for epidemiological research can be done in a store and forward manner even with a network that offers only intermittent connectivity.

As a result, in many cases the cost and power infrastructure required to support a continuously available network might not be justifiable. Many networks that are naturally intermittent include wireless networks that suffer from power outages, low-earth orbit (LEO) satellite services and physical data mules (like buses, boats) that can ferry messages from one location to another in a store-and-forward manner. This form of infrastructure that does not guarantee complete reliability and continuous connectivity can be created at a significantly reduced costs [28, 29].

However, traditional network protocols for routing and forwarding packets do not work very well for networks with frequent partitions and large, variable delays. As a result we need research on new approaches to network routing, naming, and protocol design [30, 31].

In addition, an asynchronous network system can potentially be more economically efficient by taking advantage of multiple transports (e.g. wireless, cellphone) depending on their cost and availability. The network can also smooth out traffic by delaying traffic from peak hours to increase the overall network usage efficiency. This will allow us to design the network for average daily usage instead of the peak usage. The challenge now is to integrate techniques like delaying of data, priority sending of data and selection of the appropriate transport layer into the routing decisions.

4.1.2 Wireless Networks

In cases with a little more capital and infrastructure, and where local geography permits it, wireless technologies show the best potential as a communications platform for the delivery of health services to rural areas. While none of the established solutions are suitable in their current form, technologies like 802.11 (or WiFi) and CDMA-450 offer some promise. However, there are significant research challenges in trying to adapt them to the unique constraints in developing regions.

802.11: The 802.11 standard (up to 54Mbps) for local area wireless networks is widespread enough today that chipset prices have come down to US$10. However, the 802.11 technologies have mostly been isolated to office environments, and only more recently for local-area mesh networks for urban areas.

The low cost hardware could also be cross-leveraged for rural connectivity, using long-distance point-to-point wireless links to connect villages into wide area wireless access networks. Yet, while long links using 802.11 high-gain directional antennas spanning impressive distances (125 miles [32]) have been demonstrated, they still suffer from substantial performance degradation resulting from of inherent shortcomings in the protocol design. Fundamentally the radio and protocol design, originally oriented for broadcast environments, is not appropriate for long links with non-trivial propagation delays.

The 802.11 radios are half-duplex, and therefore cannot listen for collisions while transmitting. The protocol thereby relies on acknowledgments to determine success of delivery and occurrence of collisions. All nodes sense the carrier to get a free channel before starting transmission of any packet. To avoid collisions with hidden nodes, nodes also back-off for a random interval before every transmission.

However, these mechanisms fail to work for long distance links; the carrier sense mechanism is unable to avoid collisions because long propagation delays prevent wireless nodes from hearing other nodes' transmissions in time. This causes substantial packet loss for long links. The carrier sense mechanism also leads to suboptimal use of bandwidth for relays as it restricts simultaneous transmit on multiple links from the same wireless node [17, 18]. The standard acknowledgement mechanism for reliability for 802.11 is not sufficient for link reliability as it puts a hard limit on the maximum time that nodes wait for a packet acknowledgement.

All these factors lead to significant degradation in bandwidth and jitter performance for Internet protocols like TCP and even UDP. A number of techniques ranging from bulk acknowledgments and TDMA (Time-division Multiple Access) at the medium access layer, and optimizations at the TCP layer can be used to overcome these problems.

Other open research problems in the area are design of cost-effective beam steering antennas and wireless channel allocation and scheduling for large-scale multihop networks.

CDMA 450: CDMA-450 is a variation of the cellular technology developed by Qualcomm that is widely deployed as CDMA-2000 . It runs at 450MHz, thus using a lower frequency with better propagation characteristics enabling longer communication range.

Using directional antennas and fixed terminals, the range can be extended to more than 50km. With 1xEV-DO, the maximum download and upload speeds at the physical layer are 2.4Mbps and 155Kbps respectively. Field experiments have demonstrated that download speeds upto 1.3Mbps can be achieved at 50km.

However the use of CDMA450 for long distance links faces a number of challenges. At 450 MHz, the large size of the antennas is problem when receive diversity (use of two antennas oriented differently to achieve better receive performance) is desired. As a result, the handsets are either too bulky or feature poor receive sensitivity. Current base stations also have high power consumption, a significant consideration for developing regions.

4.2 Distributed Data System

As we have seen in discussion of different healthcare areas, one of the common requirements for many applications is the creation of robust heath information systems in developing regions. There is substantial evidence that health information systems can lead to a number of benefits. While Electronic Medial Record (EMR) systems will help in tracking and managing drug usage for conditions like MDT-TB and HIV and possibly other chronic diseases, data from such information systems can be used for prompt response to healthcare emergencies and epidemiological research as well. A health information system would also aid in monitoring and supervising the status of remote health centers as well as tracking and delivery of drugs and other supplies. In addition, as mentioned in Section 2.1, we also need to provide appropriate information about medical practices and procedures to health-workers and doctors in the field - a distributed communication infrastructure can be used to send out relevant content, updates for epidemics etc. and training material to remote health centers.

The first challenge in designing a health information systems is to choose appropriate data models for each type of health data. The data domain can range from electronic medical records for individual patients, logistics data for supply chain management, and statistical data for health information collection. The goal should be to ensure that the chosen data model can support all current and future expected operations, is extensible and amenable both for exchange of data and techniques for validation of data. There has been lot of work done in the developed world on standardization of data models for both storage and exchange between healthcare providers (HL7 [33], SNOMED).

Most previous healthcare projects in developing regions build information systems residing in single place and usually for a single application (see [34] for a survey). However, a distributed data store design is much more appropriate for large-scale systems in these scenarios. Network constraints like unreliability and intermittency imply that we cannot instantly send updates or additions to any single data-store. If we want records to be accessible from multiple locations like rural health clinics and hospitals even with a network that is frequently disconnected, the only way to design a scalable system is to have a distributed data store. At the same time, a central hierarchical model where data is finally synchronized at a central master data-store is suitable for a variety of reasons. The model would be good fit for epidemiological and statistical analysis, offer better security and privacy guarantees and also simplify management and administration issues.

However building a distributed data store that takes into account unpredictable network behavior (including intermittent and unreliable connectivity) and limited bandwidth resources is very challenging. The data store needs to be able to propagate updates in the network to achieve eventual synchronization of data objects to their central master copy. The challenge is to reduce both the network usage and number of interactions by using techniques like differential updates and automatic conflict resolution. There has been some previous work in this area to build systems that provided update propagation for disconnected networks. Coda [35] required manual conflict resolution and only supported the client-server model for file synchronization, and Bayou [36] was not designed for systems that follow a central hierarchical model.

Healthcare also presents us with some unique characteristics for data access and updates. For example, the fact that medical records are by nature append only can be used to simplify conflict management. Also the patterns of migration of patients can be used for optimizations like proactive pushing of medical records and updates to relevant health centers.

4.3 Remote Video/Audio and Imaging

As we have seen previously, a key requirement of many remote health services is transfer of video/audio and images. While video and audio need continuous and possibly high bandwidth connectivity, images can be transferred even using a intermittent store and forward network.

4.3.1 *Image Coding and Transfer*

Health services like remote diagnosis require transfer of high resolution images from rural eye-care centers to hospitals where experts are located [37, 16]. Other applications include sharing of high resolution X-ray and MRI images between testing labs and hospitals.

However the challenges for coding, compressing, storing and transferring images are different depending upon the intended application and the network resource constraints. For online viewing of large-size high-resolution X-ray and MRI images on a local terminal from a remote data store, multi-resolution encoding of the image might be needed. Then we can reduce bandwidth consumption by selectively starting with a low resolution image, so that the doctors at the local terminal can quickly receive enough information to choose what part of the image is needed in greater resolution. This multi-resolution coding can be done using wavelet transforms that encode the image into multiple layers of increasing resolution [38, 39].

For applications like teleradiology and telepathology, we only need store-and-forward transfer of images over the available communication network. For very large images, we might have to use a store-and-forward network that seamlessly integrates physical media like optical discs and USB keys.

4.3.2 Video/Audio

Services like realtime consultation with remote doctors, or distance education need transfer of good quality video and audio. However, since wireless links experience high loss, real-time video and audio across them are subject to high jitter and high delay. The choice of the correct loss recovery mechanism depends on several factors: the availability of feedback channel, the delay of feedback channel and the required interactivity (conferencing vs streaming). If end-to-end delay is not a problem, the most efficient strategy is to combine hop-to-hop retransmissions with good compression algorithms. Otherwise, forward error correction can reduce loss with a less significant impact on delay.

Typical networks could also be very heterogeneous, with endpoint connectivity ranging from very low to reasonably high bandwidth. One way to deal with this challenge is to use multi-layer coding of video. There is previous work on receiver layered multicast [40] where receivers can choose what video layers they want to receive.

We also have an important advantage in that we can design an application aware networking infrastructure. This gives us the freedom to put intelligent mechanisms in the core of the network. For example, we can use the routers in the network to support multimedia services. Real-time encoding of the video can be done either at the source or inside the network on the fly depending on the demand and the current set of the receivers. Other possibilities for in-network optimizations include using the routers for opportunistic and multipath routing.

5. DEPLOYMENT ISSUES

In this section, we outline some of the basic challenges one faces while deploying new technologies in rural regions in developing countries. Associated with these challenges are also specific design tradeoffs that one may have to make depending on the local healthcare settings in a region. Finally, we try to define a framework for evaluating the social cost and benefits of deploying new healthcare technologies.

5.1 Co-existence and evolvability

The deployment of any new healthcare technology is dependent on two factors: *co-existence* and *evolvability*. The existing health conditions, health infrastructure and the set of health resources available varies significantly across different rural regions in developing countries. Correspondingly, the current set of health practices vary distinctly across different geographic regions. Any new healthcare technology should coexist with the current infrastructure and current health practices (e.g. it is important to be able to fall-back on traditional paper-based processes) in order to maximum the chances of success in terms of user adaptability and survival in the face of very diverse failure modes. In addition, the adoption of a technology benefits from *gradual evolution* where the goal of the first deployment should focus on meeting the coexistence criterion; later versions can be gradual improvements over the previous versions of the technology.

For a systems researcher developing a new technology, coexistence and gradual evolution are non-trivial constraints to meet especially since the ground realities are variable across regions. Based on our limited experiences working with Aravind Eye Hospitals in South India, we believe that actual deployments in clinical settings is absolutely essential for systems research in this field for validation of healthcare benefits. This is unlike traditional systems research, where there are no human subjects and so experimental setups in lab under controlled conditions are sufficient. Thus, from the outset we have to be mindful of these resource constraints and ensure that the systems architecture can coexist and evolve based on the health practices in a region.

5.2 Dimagi: A thought experiment

We demonstrate the benefits of *coexistence* and *evolvability* by conducting a thought experiment to robustly scale up Dimagi's [26] (see Section 3.5) pilot implementation in 3 clinics of a continuity of care program to prevent and control the spread of HIV/AIDS in Lusaka, Zambia. Patients are given a $1.25 smart card that stores their HIV related history. The clinics are equipped with smart readers and PCs that maintain a patient database. Whenever a patient visits a particular clinic, the medical history of the patient is accessed by the physician. Treatment decisions are based on this history and the card is written to reflect the updates. Patients are free to visit any of the 3 clinics. Periodically, data is physically transfered among all 3 clinics for synchronization. Barring loss of smart cards, accurate patient data is available whenever the patients present themselves at clinics. Physicians with access to this data can provide appropriate Anti-retro viral therapy (by rapidly intervening in case the virus develops drug resistance). This can prevent migratory patients from slipping into second line of treatment regimen which costs $2,000 annually versus $200 for first line of treatment resulting in significant costs savings per patient. Clearly, such a system can have tremendous impact provided such a system can scale to millions of patients.

The scaling problem: Based on our limited experience in developing regions, we anticipate some problems during scale-up. Participants in a pilot study are generally chosen for factors that maximize the success of the initial implementation to demonstrate efficacy. However, during the scale-up, several clinics may be resource limited with no access to PCs or smart card readers or technically qualified people device failures when they go down for various reasons (*e.g.,* drive crashes, dust, heat). In our own experience, We have observed a lot of computers sitting idly in rural areas, owing to the non-availability of local repair expertise.

A potential solution: Most of these rural clinics in Zambia have GSM coverage. One idea could be to equip these clinics with cell phones that run on intermittent power and have access to a central patient repository over the GSM network. When patients arrive and the smartcard readers or PCs are down, patient history can still be retrieved over

SMS. We would need a reliable transport layer over SMS to transfer data over intermittent cellular network availability to accomplish this. If treatment changes are made, then the delta updates can be encoded as a bar-code or glyph and printed on paper and given back to the patient. When the patient makes the next visit to another clinic, the cellphone camera can be used to read the bar-code or glyph to get the latest update. Thus, we can increase robustness by coexisting with current cellphone infrastructure and paper-based technologies.

While this represents one potential solution, there are many design possibilities and in the face of financial constraints we may have to make choices on selecting a subset of technical components (discussed in Section 4) to build the system. The choices depend on various factors such as cost, power, skill of human resources, each of which comes with a specific set of tradeoffs which we discuss in the next section.

5.3 Choices and Tradeoffs

Power: Since the power infrastructure is often unreliable due to outages, we may have to use alternate means as a backup option. Other potential options include solar power, wind energy (especially for towers that support wireless antennas) and battery backup. For some cases like the toolkit carried by a mobile health-worker, we might have to rely on battery power. A key concern then is to choose components that are more power efficient.

Size and mobility: A key design parameter for different health technology enabled toolkits is portability. Previous work has focused on designing a mobile toolkit fitting in a suitcase that can be easily carried by a doctor to field visits [41]. The Nethope kit [42] is also a suitcase sized data communications devices equipped with a small VSAT. Larger semi-mobile versions include the clinic in a shipping container [43], which acts as both a health clinic and a cybercafe once installed.

Connectivity: The degree of connectivity is determined by factors like cost, power and space availability. A high-bandwidth wireless link that can support high-resolution video or large number of audio streams is possible only in the case of a fixed installation where we can install wireless antennas on tall towers. However with a mobile toolkit, feasilble options are cellphones or VHF/HF radio connections. In cases where even this is not possible, we can rely on store and forward style data transfer where data is synchronized only when either the doctor travel to the data center or when we physically transfer some form of storage media. At the same time, the available connectivity in turn determines the set of applications that can be supported and therby the health services that can be provided. For example, realtime video consultation is possible only if we have high-bandwidth connectivity, whereas medical records can be synchronized even using an intermittent network.

Computing Devices: The choice of the computing platform has to be determined by cost and power availability and the computing power required. While a PC might be required at a rural health clinic where we want to connect a video camera for telemedicine, a mobile toolkit that is used only for data entry might use a PDA or a cellphone.

Human Resources: The skill levels of the healthcare workforce (doctors, nurses and health-workers) can vary significantly. The level of support and training needed for adoption of a new technology may also be dependent on the skill level of the health worker.

5.4 Evaluation of Effectiveness

To evaluate the effectiveness of health technology deployment, we can use methods like social cost benefit analysis [44]. Social cost benefit analysis evaluates the total social health and non-health benefits and costs of a particular health intervention. However, a common problem with such analysis is that it makes implicit valuations of health care benefits and human lives.

The economic cost of a health system has to take into account the capital cost of all the different system components: the rural health clinics, the equipment carried by the health-workers, the support required at the hospitals and the networking infrastructure. The recurring costs include operating costs for the network and the hardware and regular training of all the users of the system i.e. the health-workers and the doctors.

Physical accessibility is one criteria that can be used to measure the effectiveness of a particular technology deployment. As we have seen previously, a key objective of telemedicine and remote diagnosis is to increase the reach of healthcare. To estimate the increase in physical accessibility, we need to know the distribution of the population and detailed geographic data and the locations of the rural clinics.

6. CONCLUSIONS

Applying computer systems research to solving healthcare problems in the developing world has been perceived to be very challenging due to the need to be cost-effective and the need to operate robustly under demanding infrastructural constraints. Both these conditions are generally not experienced by systems researchers in developed countries. Under these circumstances, it has been unclear as to what kind of systems research can even be done to enable relevant healthcare services in developing regions.

In this position paper, we describe the role of systems research in five basic healthcare application areas in developing regions: (a) telemedicine and remote diagnosis; (b) epidemiology data collection and disease tracking; (c) supply chain management of healthcare resources; (d) cost-effective clinical equipment; (e) continuity of care. In this context, we also outline the set of new systems problems that arise in the research areas of low-cost networking, application support for intermittent connectivity, distributed data storage, and video and image coding. The overall hope of this paper is to outline the set of systems research problems in the space of healthcare in developing countries and encourage researchers to tackle them.

7. REFERENCES

[1] The Global Fund to Fight AIDS, Tuberculosis, and Malaria. `http://www.theglobalfund.org/en/`.

[2] UNICEF. The State of the World's Children 2005 - Childhood Under Threat. `http://www.unicef.org/sowc05/english/fullreport.html`.

[3] Christopher J. L. Murray and Alan D. Lopez, editor. *Global Burden of Disease*. Harvard University Press, August 1996.

[4] Prabhat Jha et. al. Improving the Health of the Global Poor. In *Science*, volume 295, march 2002.

[5] Grand Challenges in Global Health. Grand Challenge 14: Developing Useful Diagnostics Tools. http://www.gcgh.org/subcontent.aspx?SecID=391.
[6] Abdesslam Boutayeb and Saber Boutayeb. The Burden of Non Communicable Diseases in Developing Countries. 4, January 2005.
[7] Don de Savigny, Harun Kasale, Conrad Mbuya, and Graham Reid. *Fixing Health Systems: Linking Research, Development, Systems and Partnerships*. International Development Research Center, 2004.
[8] Shridhar Mubaraq Mishra, John Hwang, Dick Filippini, Tom Du, Reza Moazzami, and Lakshminarayanan Subramanian. Economic Analysis of Networking Technologies for Rural Developing Regions. In *1st Workshop on Internet and Network Economics*, December 2005.
[9] Joint Learning Initiative. *Human Resources for Health: Overcoming the Crisis*. Harvard University Press, September 2005.
[10] Florian A. Alburo and Danilo I. Abella. Skilled Labour Migration from Developing Countries: Study on the Philippines. In *International Migration Papers*, Geneva. International Labour Office.
[11] Eric Brewer, Michael Demmer, Bowei Du, Melissa Ho, Matthew Kam, Sergiu Nedevschi, Joyojeet Pal, Rabin Patra, Sonesh Surana, and Kevin Fall. The Case for Technology in Developing Regions. In *IEEE Computer*, June 2005.
[12] Richard Heeks. Failure, Success and Improvisation of Information Systems Projects in Developing Countries. In *Development Informatics*, Manchester, UK, Janurary 2002.
[13] Fitzhugh Mullan, M.D. The Metrics of Physician Brain Drain. *New England Journal of Medicine*, 353, October 2005.
[14] WHO — World Health Statistics Database 2005, 2005.
[15] University of Basel. iPath: Open Source Telemedicine Platform. http://ipath.ch.
[16] Solomon Islands Telemedicine Network. http://telemed.ipath.ch/solomons/.
[17] Bhaskaran Raman and Kameswari Chebrolu. Revisiting MAC Design for an 802.11-based Mesh Network. In *Hot Topics in Networks (HotNets-III)*, 2004.
[18] Bhaskaran Raman and Kameswari Chebrolu. Design and Evaluation of a new MAC Protocol for Long-Distance 802.11 Mesh Networks. *Proceedings of MobiCom*, 2005.
[19] K. Abbas, A. Mikler, A. Ramezani, and S. Menezes. Computational epidemiology: Bayesian disease surveillance. In *Proc. of the International Conference on Bioinformatics and its Applications (ICBA '04)*, Fort Lauderdale, FL, December 2004.
[20] Tanuja Rastogi and Allan Hildesheim and Rashmi Sinha. Opportunities for Cancer Epidemiology in Developing Countries. *Nature Reviews Cancer*, 5(11), November 2005. http://www.medscape.com/viewarticle/493232.
[21] Janini Progress Report 2004-2005: New Pathways in Public-Private Partnerships. http://www.janani.org/pdf/progressreport.pdf.
[22] Counterfeit Medicines: Some Frequently Asked Questions - May 2005, May 2005. http://www.wpro.who.int/media_centre/fact_sheets/fs_20050506.htm.
[23] ITC - e-Choupal. http://www.itcportal.com/sets/echoupal_frameset.htm.
[24] Hindustan Lever Limited: Shakti Project. http://www.hllshakti.com/.
[25] Tapan S. Parikh. Using mobile phones for secure, distributed document processing in the developing world. *IEEE Pervasive Computing*, 4(2):74–81, 2005.
[26] Dimagi. Continuity of Care in Zambia. http:\\www.dimagi.com/case-studies-zambia.php.
[27] Working Group for WLAN Standards. IEEE 802.16 WirelessMAN Standard for Wireless Metropolitan Areas Networks. http://grouper.ieee.org/groups/802/16/.
[28] Alex Pentland, Richard Fletcher, and Amir Hasson. DakNet: Rethinking Connectivity in Developing Nations. *IEEE Computer*, 37(1):78–83, January 2004.
[29] Wizzy Digital Courier. http://wizzy.org.za.
[30] The DTN Research Group. Delay Tolerant Networking. http://www.dtnrg.org.
[31] Sushant Jain, Kevin Fall, and Rabin Patra. Routing in a Delay Tolerant Network. In *Proceedings of the ACM SIGCOMM 2004 Conference*, 2004.
[32] Team iFiber Redwire. World Record for Unamplified WiFi at 125 miles. *Defcon WiFi Shootout*, 2005. http://www.wifiworldrecord.com/.
[33] HL7 Standards Body. HL7:Standards for electronic interchange of information among health care computer systems. http://www.hl7.org/.
[34] Hamish Fraser, Paul Biondich, Deshen Moodley, Sharon Choi, Burke W Wamlin, and Peter Szolovits. Implementing Electronic Medical Record Systems in Developing Countries. *Informatics in Primary Care, British Computer Society*, 2005.
[35] M Satyanarayanan, J J Kistler, P Kumar, M E Okasaki, E H Siegel, and D C Steere. Coda: A Highly Available File System for a Distributed Workstation Environment. *IEEE Transactions on Computers*, 1990.
[36] A. J. Demers, K. Petersen, M. J. Spreitzer, D. B. Terry, M. M. Theimer, and B. B. Welch. The Bayou Architecture: Support for Data Sharing among Mobile Users. *Proceedings of the Workshop on Mobile Computing Systems and Applications*, 1994.
[37] P.K. Nirmalan, A.L. Robin, R. Krishadas, R. Ramakrishnan, R.D. Thulasiraj, and J. Tielsch. Utilisation of Eye Care Services in Rural South India: The Aravind Comprehensive Eye Survey. *British Journal of Ophthalmology*, 88:1237–1241, 2004.
[38] John Huffman, Paul Chang, Carlos Betancourt, and John C. Liu. iSyntax: A Flexible Representation for Image Data. 1999. http://www.stentor.com/pacs/iSyntax.php.
[39] JPEG Committee. JPEG 2000 Standard. http://www.jpeg.org/jpeg2000/.
[40] Steven McCanne, Van Jacobson, and Martin Vetterli. Receiver-driven layered multicast. In *ACM SIGCOMM*, 1996.
[41] MIT Media Labs. Portable Tele-Diagnostic Devices. http://www.media.mit.edu/resenv/telemed/.
[42] NetHope Toolkit. http://www.nethope.org.
[43] MIT Media Labs. Little Intelligent Communities project. http://www.media.mit.edu/unwired/.
[44] R Eastwood and M Lipton. The Impact of Changes in Human Fertility on Poverty. *Journal of Development Studies*, 2000.