

Using Mobile Phones and Open Source Tools to Empower Social Workers in Tanzania

M. Beatrice Dias, Daniel Nuffer, Anthony Velazquez, Ermine A. Teves, Hatem Alismail, Sarah Belousov, M. Freddie Dias, Rotimi Abimbola, Bradley Hall, M. Bernardine Dias

Abstract—Although para-social workers carry the primary responsibility in providing essential services to the growing population of orphans and vulnerable children in Tanzania, they are often not paid for this work. Moreover, these para-social workers are unable to access governmental resources due to the lack of an efficient means of reporting their needs to relevant government officials in a timely manner. In this paper we describe a text message (SMS) based solution that harnesses the prevalence of mobile phones coupled with several Open Source tools to empower these para-social workers. Specifically, we build a more efficient mechanism for reporting summary data on orphans and vulnerable children to relevant government officials in a cost-effective and efficient manner. This paper reports on our needs assessment process, reviews the related work, describes the implementation and testing of our prototype solution, and concludes with a discussion of relevant future work.

Index Terms— Mobile (Cell) Phones, Text Messaging (SMS), Open Source, Social Workers

I. INTRODUCTION

THE goal of this work is to explore the role technology can play in alleviating some of the burden placed on social workers and para-social workers who are attempting to serve the millions of orphans and vulnerable children (OVC) across Tanzania [1]. Specifically, we explore the use of SMS (Short Message Service or text messaging) to improve the cost, regularity, and timeliness of reporting data to the relevant

Manuscript received July 31, 2010. The research reported in this publication was supported in part by discretionary gifts to the TechBridgeWorld research group at Carnegie Mellon University, by the Qatar Foundation for Education, Science, and Community Development, by the support of National Priorities Research Program grant #30-6-7-91 from the Qatar National Research Fund, by Carnegie Mellon University in Qatar, by the University Computing Centre, Ltd located at the University of Dar es Salaam in Tanzania, and by the Holleran Scholars Fund. The opinions expressed are those of the authors and do not necessarily represent the views of any of our sponsors.

M. Beatrice Dias is with the Department of Engineering and Public Policy at Carnegie Mellon University (CMU), Pittsburgh, Pennsylvania 15213 USA (beadias@cmu.edu). Daniel Nuffer was with the School of Computer Science at CMU. He is now with Jane Street Capital in New York (nuffer@gmail.com). Anthony Velázquez was with the School of Computer Science at CMU (anthony.velazquez@gmail.com). Ermine A. Teves, Hatem Alismail, Sarah Belousov, M. Freddie Dias and M. Bernardine Dias are with the Robotics Institute at CMU (eteves@cs.cmu.edu, halismail@cs.cmu.edu, sarahtbw@ri.cmu.edu, mfdias@ri.cmu.edu, mbdias@ri.cmu.edu). Rotimi Abimbola was with the Department of Social and Decision Sciences at CMU (rotimi.abimbola@gmail.com). Bradley Hall was with the Department of Mechanical Engineering at CMU (hallbt@gmail.com).

government database so that resource distribution can be better informed and services provided to OVC can be enhanced.

A. Background

As of 2007, an estimated 33 million people across the globe were infected with HIV [2]. Prevalence of HIV among adults in Tanzania is estimated to be 6.2%; that is an order of magnitude greater than prevalence in the United States (0.6%).¹ HIV renders people more susceptible to infection, so exposure to these diseases increases the likelihood that HIV-positive Tanzanians will develop AIDS. In 2007 alone, Tanzania reported approximately 96,000 deaths due to HIV/AIDS. Furthermore, Tanzanians are at a high risk of contracting other infectious diseases such as Hepatitis A, Typhoid Fever, and Bacterial Diarrhea. Adult mortality in Tanzania has left approximately one million children (aged 0 to 17) parentless and in a vulnerable position.² The HIV/AIDS Twinning Center estimates this number is even higher and reports that “nearly 2.5 million Tanzanian children have been orphaned or made vulnerable by HIV/AIDS, and the number is expected to reach four million by 2010 if effective prevention and treatment measures are not implemented.”³

Poverty is a factor that further cripples those afflicted by HIV/AIDS. According to the World Bank, 36% of the Tanzanian population falls below the national poverty line.⁴ Economic hardships exacerbate the HIV/AIDS epidemic by limiting the healthcare resources accessible to patients. The social stigma concerning HIV/AIDS is an added burden borne by Tanzanians who contract the disease, and this stigma often extends to the relatives and associates of the HIV infected Tanzanians. Hence, fear of being ostracized deters many Tanzanians from getting tested for HIV and contributes to the continued spread of the disease.

Due in part to this spreading epidemic, Tanzania faces a severe shortage of trained social workers, particularly to provide services to affected individuals. Some of the factors escalating the need for social workers in Tanzania are the increasing number of HIV/AIDS adult deaths, the concurrently rising number of OVC, and the breakdown in

¹ <https://www.cia.gov/library/publications/the-world-factbook>

²

http://www.unaids.org/en/CountryResponses/Countries/united_republic_of_tanzania.asp

³ http://twinningagainstaids.org/pr-073107-social_work_training.html

⁴ http://devdata.worldbank.org/AAG/tza_aag.pdf

family support structures due to greater migration from rural areas to cities.⁵ Although need is abundant, human and financial resources available to social welfare groups attempting to address these challenges is limited.

B. Para-Social Workers

Tanzania needs at least 8,000 additional social workers, according to the country's Institute of Social Work (ISW), to meet the current demand for services. To address this challenge, several groups in Tanzania are involved with training thousands of para-social workers (PSWs) to provide services to HIV/AIDS victims and their families; especially to OVC. Many of these PSWs are already community development officers or representatives of community-based organizations.⁵

An estimated two-thirds of Tanzania's 127 districts are left with no social welfare support due to a lack of trained welfare workers.⁵ Thus, in 2007, the HIV/AIDS Twinning Center joined with ISW, the University of Illinois at Chicago's Jane Addams College of Social Work, and the Midwest AIDS Training and Education Center to launch their Para-Social Worker Training Program. This program was "designed to arm community-based caregivers with critical skills needed to help improve the lives of orphans and vulnerable children."⁶ At the time of this work, through their efforts, these organizations had trained 516 PSWs, 40 district social workers, and 55 master trainers in Tanzania.⁶ "The newly trained PSWs are identifying new children and families in need and connecting them with critical support and assistance."⁶ Home visits are conducted to gather information on the children and their current situation. However, the supply of social workers still falls short of the number needed to reach all OVC around the country.

C. Supporting Efforts

In addition to the ongoing work by PSWs, social welfare organizations such as HUYAWA, the Institute of Social Work, and WAMATA provide OVC and their families with school supplies, counseling services, and advice on a range of life skills such as caring for family, managing with a small budget, as well as family planning and HIV/AIDS prevention.⁵ Furthermore, there is a national effort to monitor and evaluate HIV/AIDS data in Tanzania by way of the Tanzania Output Monitoring System for HIV&AIDS (TOMSHA). This national system is dedicated to collecting information on the impact of the disease in Tanzania, which includes data on OVC. The Department of Social Welfare (DSW), a governmental organization, maintains a database specifically on OVC. The goal for this database is to enable leaders to make more informed decisions on how to combat the spread of HIV/AIDS in the country, and to allocate limited resources more effectively. However, PSWs usually receive no compensation for their work, and therefore do not generally

have incentives to collect or submit high quality data to the DSW database.

In summary, while PSWs carry the primary responsibility in providing essential services to the growing population of OVC in Tanzania, they are often not paid for this work and are unable to access governmental resources due to the lack of an effective means of reporting needs and relevant information to the national database on OVC in a timely manner. We worked with several relevant groups in Tanzania to understand the needs of this community and based on these assessed needs, we proposed a SMS-based solution to address the reporting challenges faced by the PSWs. This paper reports our needs assessment process, reviews the related work, describes the implementation and pilot testing of our solution, and concludes with a discussion of future work that could further enhance this solution.

II. NEEDS ASSESSMENT

Our needs assessment process informed us of the challenges faced by the social worker community in Tanzania, their desired solutions, and the existing infrastructure that has been created by associated organizations to support them. To gather this information we identified two community partners who could assist us with our needs assessment process and also provide us with valuable feedback on our developed solutions. Our first partner was the Institute of Social Work (ISW), an academic institution that works with several partners to train PSWs as well as the trainers of new groups of PSWs through programs throughout Tanzania. Our second community partner was the Department of Social Welfare (DSW), a government office, which currently maintains a national database of OVC.



Fig. 1. Researchers working with social workers and PSWs. Photos used with permission.

⁵ <http://news.bbc.co.uk/2/hi/africa/7239047.stm>

⁶

Initial meetings were held with several local organizations that conduct or oversee social work. ISW houses a program for supporting OVC, which started in October 2006 and is sponsored by USAID. At the inception of the program, they trained community members who already work with OVC in order to empower and motivate them to continue to provide services to the OVC. After a social worker symposium, a curriculum was developed for training supervisors who then go on to train PSWs. ISW partnered with the Jane Addams College of Social Work at the University of Illinois at Chicago in the United States to deploy the “Train the Trainer” approach. Through this partnership, ISW has been able to train several hundred PSWs under the new curriculum. The PSWs are all trained in their respective districts with an emphasis on methods for identifying OVC and a standard process for collecting data and providing services.

Several groups helped us to better understand the data collection process for the national database. At the WAMATA non-profit organization that was the first HIV/AIDS support organization of its kind, the staff members explained that a major challenge is keeping records and tracking the OVC. The OVC Data Management Specialist at DSW stated the major issues that their institution faces with the data collection for OVC was incompleteness of and inaccuracies in the data. At the time of this work out of a total of 132 recognized districts, 81 reported to the national database while 51 did not submit any information. Those that reported often did not send complete information. Additionally, information verification occurs through a manual checking process conducted twice a year. If the national database received information electronically and more frequently than every three months, the entire process of data collection and verification could be streamlined.

A. PSW Interviews and Observations

We interviewed five PSWs all from the Dar es Salaam region, which contains three districts – Kinondoni, Ilala, and Temeke. Many of these workers are involved with the village or ward committees that provide data on OVC to the national database, and would therefore be the end users of the proposed technology solution. We also conducted a site visit to Tandika village, which is located in the Temeke District. Our interactions with the community in the Tandika village provided us with a firsthand view of the problems that the village members (including the PSWs and the OVC) face on a daily basis.

The interviews with the PSWs revealed that lack of resources and incentives were major obstacles to comprehensive data collection. The PSWs were able to give us a grassroots perspective as to why the data does not get to the various levels of government as fast as it is needed. PSWs focus on meeting the day-to-day needs of the OVC and other community members and often do not have the time or resources to submit paper reports. They also find that sending in reports to the district level governing body, yields little in terms of support or resources. Instead the village committee must find the resources to provide OVC with essential items such as food, health care, school supplies, and clothing. Apart from this, printing, photocopying, and travel expenses required for preparing and submitting reports are costly,

especially if there is no assurance that the district will compensate them for these costs. Furthermore, there is no official at the district level who advocates for providing funds and resources to social welfare-related work. Finally, although the PSWs receive some training, they often do not receive support beyond that initial instruction. Many of them are not paid for the social work services they provide. They perform social work on a voluntary basis and maintain other jobs to earn a living. The interviewees were schoolteachers, cooks, electricians, program coordinators at non-governmental organizations, and chairmen of their village committees.

B. Data Collection Challenges

At our initial meeting with partners, we found out that data transfer from the ward level to the district level is very slow. DSW in Tanzania collects data from village, ward, and district committees. Currently, this data collection occurs via paper forms, and as such, it can take a long time—on the order of months—for the forms to reach their final destination, the national database.

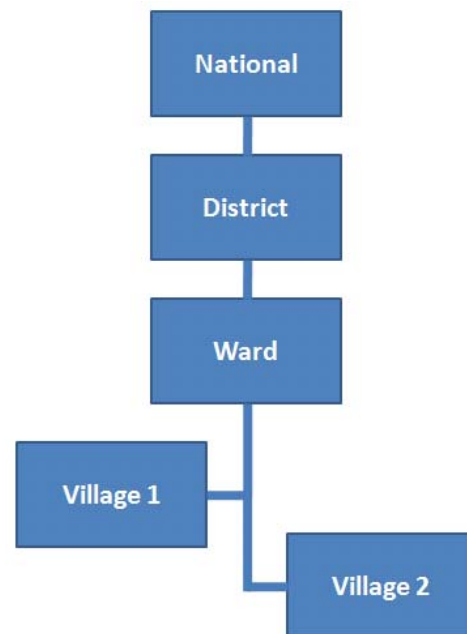


Fig. 2. Diagram of Data Flow.

Specifically, reporting is a multi-step process involving paper forms sent to the district where they are digitized and forwarded to the national database (Fig. 3).

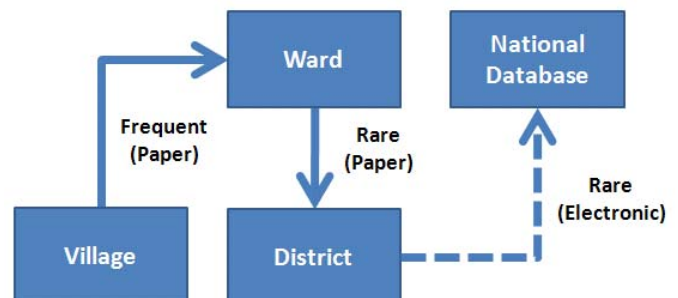


Fig. 3. Flow of Reports to National Database.

PSWs predominantly operate at the village level. Their efforts are concentrated on serving villages and the reports they construct are brought to a ward. A ward is a collection of approximately three villages, where the social workers can meet and interact, as visits to the ward happen very frequently, if not daily. At the ward level, the paper reports are collected and then are taken to the district level. This process is rare and expensive. The bus ride into the district can take hours and is often uncomfortable and costly. At the district level, the forms are digitized and submitted to the national database once every three months. However, the many difficulties of ward-to-district transportation can cause the forms to be delayed anywhere from a couple of months to a year, in addition to some forms being lost entirely. This in turn leads to the national database containing out-of-date and incomplete information.

There are two main issues with this data transfer: 1) lag time to transport the paperwork and 2) misplaced or lost documents during this transfer. Approximately 60% of districts send in reports to the national database, and only a handful of those reports are complete. The omissions and flaws in the reporting process hinder the flow of communication. Government officials and donors use this information to make critical funding and policy decisions. Therefore, it is important to obtain accurate data so that adequate resources are provided to support services for OVC.

C. Proposed Solution

Based on the findings of our needs assessment we proposed a technology solution that could address the reporting needs of the local social worker community. This solution was iteratively developed utilizing feedback from our community partners and led to the production of our SMS-based solution. The scope of this project was to test this prototype by ascertaining its potential to be deployed, adopted, scaled up, and sustained in Tanzania. To this end, we obtained relevant user feedback and also proposed a deployment strategy.

Prior to developing our solution, based on the needs identified with our partners, we surveyed related technology solutions to inform our solution design.

III. RELATED WORK

The problem of exchanging information without advanced infrastructure such as Internet access, stable mobile phone connections, or landline telephone service is a problem occurring throughout a variety of different ICTD (Information and Communication Technologies for Development) projects. Due to the prevalence of this problem, a variety of research groups and other organizations are working to improve information exchange in rural areas.

A. Voice-Based Solutions

Several research efforts have explored voice-based information retrieval systems relevant to developing community settings. One example is HealthLine, which is a speech-based system that allows users to access information from a database via a phone call.⁷ There are several

advantages to a speech-based solution. First, such systems do not require high levels of literacy, and only call for a low level of technical skill. Moreover, all data processing and custom software is server-side, eliminating the need for any installation on users' phones. However, this work does not address data entry, which is an essential component of our project [3]. Our work is targeted towards reporting rather than retrieving information. This can be a relatively expensive process via phone call and hence is not feasible for our target users.

Similarly, the TRACnet system in Rwanda was designed to collect, store, retrieve, and disseminate a variety of information related to HIV/AIDS care and treatment [4]. The Ministry of Health and the Treatment Research and AIDS Centre (TRAC) joined forces to deploy TRACnet to increase the efficiency of Rwanda's HIV/AIDS program management and enhance the quality of patient care. TRACnet has been deployed nationwide to connect every health facility providing HIV/AIDS treatment and related services. The primary mode of information exchange in TRACnet is a bilingual English and French telephone and web interface. The backend of TRACnet is a central information repository. The system has reportedly transformed a largely paper-based one way information flow that took several weeks, into a bi-directional data exchange completed in seconds. Hence, it allows decision-makers and supervisors to quickly analyze and respond to program information. Uganda's Rakai Center and others are utilizing similar systems to enhance information management and treatment provided to HIV/AIDS patients.⁸ However, none of these systems have explored methods for reporting information in settings where phone calls are cost-prohibitive. In working with PSWs in Tanzania who have very limited income and who were unable to afford phone calls or internet access for regular reporting of OVC data, we had to explore alternative solutions.

Patnaik, Brunskill, and Thies presented a review of accuracy rates of different data collection mechanisms using mobile phones [5]. Their study compared three different methods, utilizing a Java application to enter data into a form, sending a coded SMS message, and verbally reporting information to a human operator via a phone call. In comparing these three methods, they tracked mistakes that workers made as well as the duration of the interaction. The results showed that the SMS and Java application methods had a comparable error rate (under 5%). The voice solution, however, produced errors less than 0.5% of the time. The duration of an interaction using SMS and the Java application was also close, while interactions on the phone took almost one minute longer. In exploring alternatives to voice-based solutions which were cost-prohibitive for our targeted users, this study was useful in understanding the accuracy rates of alternate communication methods.

B. SMS Solutions

FrontlineSMS is a project that was developed to encourage stronger communication for non-governmental organizations and their workers. The system provides a suite of tools that only requires a computer and a mobile phone. It can then

⁷ <http://www.cs.cmu.edu/~healthline>

⁸ <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2880005/>

arrange text messages such that they can be sent either to individuals or to large groups, and conversations can be organized.⁹ It has been used in a variety of projects, including in a hospital in Malawi where it was able to track patients, respond to requests for care, and answer health worker's queries for dosage information.¹⁰ However, FrontlineSMS is not Open Source and is better suited for connecting and tracking field work. The goal of our project was to automate the data entry, and provide confirmation that the reported data reached the database. FrontlineSMS does not allow users to export and automate the process of data and therefore was not as useful for our target application.

Texas A&M University implemented the LINKS project in 2004 in Kenya and Ethiopia. LINKS, which stands for Livestock Information Network and Knowledge System, allows users to send coded SMS messages to a central server [6]. These SMS messages contain data such as livestock type, age, and condition and the server will reply with a SMS message containing the prices a farmer could receive for such an animal in different nearby markets. This enables farmers to maximize their profits and thereby improve their livelihood. The LINKS study emphasized that using coded text messages is a feasible way of interacting with a system in the context of ICTD work. This system was the most promising solution for the constraints of our target application, and hence the LINKS protocol informed several aspects of our solution design.

C. Other Relevant Tools

The OpenRosa consortium is a group of developers working to create Open Source protocols for data collection on mobile devices. Through projects such as JavaROSA, OpenRosa's standards have been used to develop mobile phone applications in developing communities [7].¹¹ Projects using the JavaROSA platform can be run on most Java-enabled phones, including the Nokia 3110c and 6085, which are readily available in low-income regions [8]. However, despite being available, Java compatibility is neither universal nor cheap. Since mobile phone customization is relatively uncommon in Tanzania, this may lead to a number of issues, especially concerning usability, as the PSWs did not seem to know how to use Java applications. In contrast, the PSWs were very comfortable using SMS.

Epihandy is a suite of tools constructed with OpenRosa for data aggregation on mobile phones. It includes a wealth of already developed tools both user- and server-side to run on smart and Java-enabled phones that would minimize the need to re-implement basic data collection functionality. Epihandy includes tools to design user interfaces that make data entry more user friendly [9]¹² but also requires a consistent connection for data collection [10]. Additionally, the storage of data on memory cards raises the cost to ensure that data is not lost before it is transmitted, and privacy becomes a concern in the event that a mobile phone is lost or stolen.

CAMBrower is a solution that targets form-based data entry. Utilizing the cameras that are becoming more prevalent in mobile phones, CAMBrower's developers hope to make

data entry more efficient. To enter data, users have a copy of a paper form listing all the fields and a corresponding bar code for each field. To enter data into a particular field, they take a picture of the bar code and CAMBrower is able to decode the image to ascertain the type of data the user is entering, thus ending the need to navigate through menus [11]. CAMBrower offers an alternative to cumbersome navigation menus or text coding that is valuable for low-literacy users. However, for our work, the users were literate and given the complexity of implementing CAMBrower, we decided that the cost outweighed the benefits. Additionally, mobile phone cameras are not universal and hence an application that requires a camera would limit the number of users that could be served.

None of the related work surveyed addressed the need for effective processing of reports at the ward level as well as the need for the application to be cheap and easily deployed on a variety of mobile devices. Thus we decided to develop an application specifically for enabling PSWs in Tanzania to report vital information on OVC more easily and quickly.

IV. IMPLEMENTATION DETAILS

Our needs assessment process revealed that the PSWs need a less complicated and less costly method for submitting reports, if they are to submit data more frequently and consistently. All of the PSWs we interacted with own basic mobile phones (mostly Nokia brand) and use their phones on a regular basis; typically every day. In contrast, none of them had easy access to a computer. As a result, we decided to focus the project on developing a mobile phone application that would help social workers and PSWs transfer relevant summary information more regularly into the DSW database. In Tanzania SMS is much more affordable than phone calls. The cost of a mobile phone call per minute can be as much as five or six times the cost of sending one text message.¹³ Therefore, a SMS-based solution would be a less cost prohibitive mode for transmitting data. Although a SMS-based technology solution is limited by the character allowance in a text message, our community partners identified the following key pieces of information that can be submitted in aggregate form via a SMS:

- Number of OVC
- Gender and age breakdown of OVC
- Location of OVC
- Needs of OVC and services provided to them (to match and determine which needs are still unmet)

A. Information Flow

Informed by the assessed needs of the PSWs, the primary goal of our project is to digitize information at the ward level as illustrated in Fig. 4.

⁹ <http://frontlinesms.com>

¹⁰ <http://www.jopsa.org>

¹¹ <http://www.dimagi.com/content/javarosa.html>

¹² <http://www.epihandy.com/index.php/EpiHandyMobile>

¹³ <http://www.zantel.co.tz/tariffs.html>

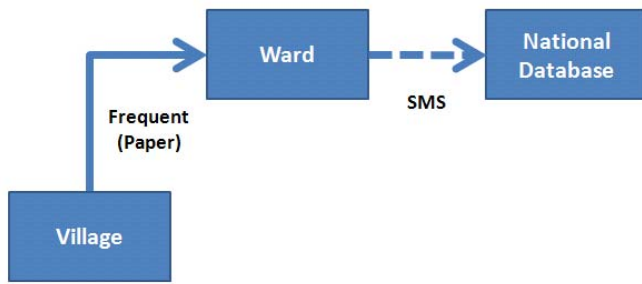


Fig. 4. Proposed Flow of Reports (dotted line shows digitized information while solid lines show paper-based or orally communicated information)

We targeted our efforts towards streamlining the process of data transfer out of the ward directly to the national database. In doing so, the lengthy and arduous task of transporting a large number of backlogged paper forms to the district level at uncertain intervals could be minimized and the national database could be updated directly from the ward level. Due to the limitations in the volume of data that can be transferred over SMS, it is infeasible to digitize all data at the ward level using the proposed approach. Hence, this system does not entirely eradicate the need for paper forms. However, it does allow the national database to be updated with key information at more frequent intervals.

B. SMS Solution Tradeoffs

SMS is a low-cost, low-bandwidth method of transmitting data that is supported by most mobile phones. In particular, we found a SMS-based solution most appropriate because many users already own mobile phones, and the cost of an individual SMS is very low (somewhere between US\$.03 and US\$.05 in Tanzania at the time of this work). Additionally, owing to the fact that SMS is considerably less expensive than standard voice communication, many Tanzanian mobile phone users are already familiar with SMS.

We explored several other options before ultimately settling on a SMS-based solution. A voice-based system, where users dial in to a call center and speak to a live operator, who then enters the relevant data, is highly robust. This method avoids the possibility of mistyping data, and allows for instant clarification of information and collection of a greater volume of data. Additionally, a voice-based option was shown to have lower error rates than SMS and mobile phone applications [5]. However, the high cost of “airtime” compared to SMS, as well as the additional cost imposed and time required to train and employ phone operators, dissuaded us from pursuing this solution path.

We also investigated the possibility of using a mobile phone application to help users format their responses properly; our hope was that this would decrease error rates by making data entry more intuitive. However, previous ICTD studies have shown that error rates were not significantly different between data entered with the aid of a Java application and hand-formatted SMS [5]. Each of the application formats we considered posed significant barriers to wide-spread deployment: Java applications tend to only run on higher-end mobile phones, and SIM card applications require approval from network operators (which is virtually impossible to get).

Therefore, we designed a system that uses hand-formatted SMS messages to transmit data.

C. System Components

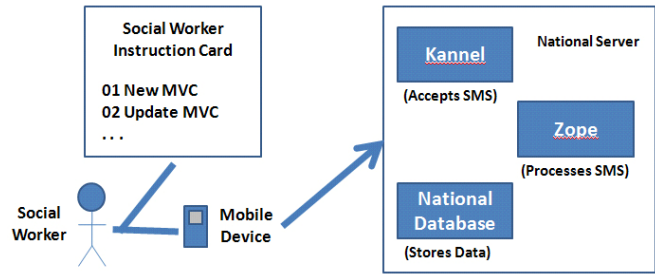


Fig. 5. System Component Overview

We employed a key code style format SMS message similar to the one used by LINKS [6]. As illustrated in Fig. 5, social workers are given a small card, which contains all of the codes to be used in the SMS message. This SMS message is then sent to a central server, which takes the information and process it into the national database.

We decided to implement the server using several Open Source technologies: Kannel¹⁴, Python, and Zope. Kannel is a SMS gateway, which serves as the connection between the mobile phone networks and the server, connecting one or more GSM modems to Zope. Zope is a web framework written primarily in Python. Our decision to use Zope was motivated by a need to design a simple and intuitive administrative interface as well as work together with Kannel and external databases. For database connectivity, we used SQL Alchemy¹⁵, an object relational manager for Python, which easily allows us to plug into various back-ends. Presently, our development version of the server runs using SQLite. However, for larger scale deployment a better choice would be a production quality database, such as MySQL or PostgreSQL.

We designed the system to receive text messages through the SMS gateway, Kannel. By connecting a phone or GSM modem to a computer, Kannel receives SMS messages from the phone, processes the information into an XML file, and sends it to a Zope application. Zope is a content management system that has been in development since 1995. Using an array of Zope modules, an application was developed that stores the received SMS messages in SQL by parsing the XML file for the required fields.¹⁶ If the Zope application finds the message to be improperly formatted or having anomalous values, Zope returns an error to Kannel, which replies the user with the problem Zope encountered. Following this process, the user can remedy the error and resend the information, which can then be submitted to the national database. In our solution design, users are given instructions on how to format a special SMS containing data they have collected in their work with OVC. Shortly after a user sends this SMS to the supplied number—a phone number that would be owned by the server operator—the server

¹⁴ <http://kannel.org>

¹⁵ <http://www.sqlalchemy.org/>

¹⁶ <http://zope.org>

receives, parses, and automatically enters it into a database. The server is also able to determine, from the SMS, which user sent the data and the region where he or she works. Server administrators can export and analyze the submitted data as they choose.

User-side

Data flow begins with the users—in our case the PSW or social worker. Users are instructed via cue cards to submit data by sending a specially formatted SMS. This technique of using cue cards has been shown to be effective in a previous ICTD study [5]. Still, despite its many advantages, SMS has its drawbacks. A SMS is limited to 160 characters, and the lack of a proper keyboard on most phones makes typing large amounts of data cumbersome, and error-prone. Thus, it is reasonable to expect users to be able to accurately enter no more than a small amount of numerical data—a total of approximately 10 to 20 characters. Based on discussions with personnel at our partner organizations, we determined that the most useful data would be summary data, such as the number of OVC identified in a village, their gender breakdown, their age range, as well as which services are currently best provided and which services are most needed (enumerated using an existing coding system). Such data is expected to allow an administrator to identify meaningful trends in social work across the nation and enhance the input of data to the national database.

Each SMS begins with a personal identification number (PIN) that uniquely identifies the user. Combined with user registrations in the database, the PIN also allows the server to identify where the data is coming from, as each PIN is tied to a specific village. Users can either register via SMS or can be manually entered into the database by an administrator.

Server-side

Within minutes, if not seconds, of a user submitting a SMS, the server receives the message. This is accomplished using Kannel in conjunction with a GSM modem. Since the format specified by the cue cards is machine-readable, as soon as the server receives a SMS, it is parsed and entered into the database. In the event that the server receives an incorrectly formatted message, or one with clearly incorrect data (such as a report of a thousand OVC in a given village), instead of entering incorrect data into the database, the system flags the message for review by a human operator. In particular, messages that cannot be properly parsed are flagged, and depending on the data, we could implement range-validation as well, which will only allow specific, reasonable ranges of values for different fields. This system combines the benefits of the robustness of a human operator when needed with those of digitization.

V. PROTOTYPE TESTING

After developing a prototype technology and cue card system, we conducted an initial evaluation to gauge the usability and feasibility of the proposed technology. This was done in two stages:

- **Stage 1:** Determining deployment potential of the mobile phone application by demonstrating the solution prototype to partner organizations to obtain their feedback.
- **Stage 2:** Assessing feasibility of cue card instructions with a village OVC committee to understand the degree to which they can follow the instructions to send SMS in the required format.

A. *Feedback on Technology Demonstration*

Our technology demonstration resulted in the following feedback about the technology:

- Highlighting errors in data records for human intervention is very useful.
- The limitation on the amount of data that can be sent via SMS is a concern. Currently, they use paper forms to send a lot more text-intensive data such as comments on particular children and other problems in the community. It is not feasible to send such data in one SMS message.
- Training people to send data using this technology could be challenging because many people in villages and wards have limited exposure to technology and some do not even own or use mobile phones.
- Lack of mobile phone network coverage in certain areas of the country will be an issue for a SMS-based solution.
- Since there are up to 12,000 villages in Tanzania, we need to make sure there are a corresponding number of unique PIN codes to assign to data senders at different locations.

This feedback was useful in understanding some of the obstacles we would face when deploying a larger-scale phase of the project. Since the feedback about the technology itself (in terms of its capabilities and design) was positive, the demonstrations did not lead to further modifications of the solution at this stage. However, it did help shape the instructional cue cards as well as the deployment strategy for this project.

B. *PSW Feedback on Instructional Cue Card System*

The instructional cue cards designed for data senders were tested on five OVC committee members (one woman and four men) from Tandika village. For this testing, we provided each of them with a cue card, which contained the following instructions, translated into Kiswahili (see Fig. 6):

We asked the PSWs to imagine they had 11 OVC in total, of which six were female and five were male. Equipped with this information and their mobile phones, we asked them to try to follow the instructions and transmit the data provided. The mobile phone on the receiving end of the SMS data was one of the researchers' personal phones. This allowed us to verify the accuracy with which the PSWs formatted and submitted the given data. There were several useful observations that resulted from this trial:

- Participants had a difficult time understanding step #2 of the cue card instructions. They asked for some guidance for this, and were subsequently able to send the SMS in the required format.
- None of them knew what “numerical input mode” meant, so we needed to show them how to access it on their phones so they could follow the instructions.
- It was clear that three of the five OVC committee members were quite proficient in sending SMS with their

phone, and mastered the instructions very quickly. The other two committee members had a difficult time getting past instruction #2.

- One of the members had little to no experience with SMS and required guidance through the entire process.
- We received correctly formatted messages from three of the members. A fourth member formatted it correctly, but didn't wish to send the SMS. The fifth member did not successfully format or send a message.
- We noted that the committee members worked together to assist each other with the task, so with more time they are likely to help each other be able to send correctly formatted SMS.
- The participants said that the instructions were not difficult to follow, except for step #2. They also mentioned that they think such an application will be useful to them in their work and that they would use it if it were available.

1. Begin a new SMS message
 - Make sure the message is blank
2. Switch to Numeric Input Mode
 - On many phones:
 - Press the menu button
 - Select "insert number"
3. Enter your PIN
 - Your PIN is: 1668
4. Enter a space, or a *
5. Enter the number of OVC in your village
6. Enter a space, or a *
7. Enter the number of male OVC in your village
8. Enter a space, or a *
9. Enter the number of female OVC in your village
10. Review the message
 - It should be formatted similarly to:
1668 10 4 6
- OR -
1668*10*4*6
11. Send the message to 077XXXXXXX

Fig. 6. Cue Card.

C. Key Challenges and Limitations

There are a number of challenges inherent to the problem of transmitting OVC data in Tanzania to a national database. Chief among these is costs. Once the server is set up, the cost per data submission will be just the cost of one SMS. At the time of this publication the cost of sending a SMS was roughly equivalent to the cost of making a photocopy, which is currently necessary for the PSWs to submit the paper-based forms.

Although we believe mobile phones can be an effective tool in a wide range of data transmission applications, including this one, there are a number of limitations to the platform

worth noting. More remote areas in Tanzania, where data collection is already challenging, may have little or no mobile phone reception. Thus, our system is designed to complement rather than replace the existing system of paper form transmission. However, the spread of mobile connectivity is likely to address this challenge in the near future. Furthermore, mobile phones are not designed as general-purpose data entry devices, and so they run into usability issues. Entering data via the keypad, particularly alphanumeric data, can be difficult and error prone.

D. Deployment Considerations

During the last week of our field research in Tanzania, we worked with our local partners to come to an agreement on what the next steps of this project would be. We agreed to continue communication with DSW through their data management specialist. Furthermore, we will work with DSW to come to an agreement on ownership of the application and a plan for future deployment of the project.

We are also interested in helping ISW to explore other uses of the application that can benefit Tanzania such as designing and implementing a system to help monitor the PSW training program and collect statistics on the spread of diseases.

PSWs were given the opportunity to learn about the technology and provide feedback on the instructional cue cards. Apart from this, we committed to keeping them apprised on the progress of the project via ISW. The sustainability of this project relies heavily on a local party taking further ownership of the application and hosting it on their server. The ideal candidate for this role would be DSW. Once ownership of the technology is undertaken by a local partner, we can work with that group to plan further testing as well as a long-term study to deploy, collect data, and evaluate the effectiveness of this technology solution.

VI. CONCLUSIONS AND FUTURE WORK

In this work we show that a SMS-based data transmission system can be feasibly implemented within the social worker community in Tanzania. However, several steps need to be taken to build the human resource and government advocacy infrastructure that would be necessary in order to successfully deploy this system. The SMS-based technology will help supplement the current paper form method for data collection within this community while improving the efficiency and timeliness of transmitted data. Additionally, the affordability of SMS compared with alternative avenues for information communication, render this approach more feasible to and adoptable by the end user; the PSWs at the ward or village level.

In future work, we also plan to enhance our system given further feedback, and begin deploying the application in a relatively small trial, potentially in the Dar es Salaam region. To accomplish this deployment, we would need further cooperation from DSW which is officially in charge of social welfare programs within Tanzania. During this trial, we will collect further feedback on the application and iteratively improve and expand the technology. Additionally, we intend to focus on establishing local ownership of the project and

infrastructure for broader-scale deployment. We anticipate that the application could easily grow to cover the entire social welfare network of Tanzania.

In addition to further testing and implementing this solution, there are a number of interesting technical areas where our work could be extended. Two interrelated ideas we would like to highlight are server-to-user communication and time-sensitive surveys. Server-to-user communication would allow server administrators to contact users via the same mobile phones they use to submit data. There are some issues with this, including the bandwidth limits of a GSM modem, the possibility that such communication is unwanted by its recipients, and of course additional costs. That said, by initiating two-way communication, valuable new possibilities emerge. One such possibility is a method for collecting time-sensitive surveys. A survey could be distributed via SMS and then collected via SMS, with a response time that would theoretically be in minutes. Again, there are challenges involved in implementing this, particularly regarding usability, but it could prove a versatile and useful tool. Some possible uses could include tracking the spread of diseases or in disaster monitoring.

Another area we hope to explore is generalization. The project generally solves the problem of submitting data, especially location-tied data to a database. This has many possible applications beyond the scope of collecting data for social welfare in Tanzania.

ACKNOWLEDGEMENTS

The TechBridgeWorld research group is grateful to the many people and organizations that made this project possible. We would especially like to thank our main community partner, the University Computing Centre (UCC), who made it possible for us to conduct technology research projects in Tanzania, including our contacts Professor Beda Mutagahywa, Eric Beda Mutagahywa, Callista Duwe, Anthony Gesase, and Dorothy Mbonike. We would also like to acknowledge Vice Chancellor Professor Rwekaza Sympho Mukandala, Saudin Jacob, and Dr. Tito Mwinuka from the University of Dar es Salaam for their assistance, as well as University of Dar es Salaam students for their interest in the project. We are extremely grateful to the Institute of Social Work (ISW), and the Department for Social Welfare (DSW) for their help with facilitating needs assessment and feedback for this project. We are grateful to Dr. Theresa Kaijage, Bernard Sefu, Leah Omari, and Para-Social Workers from ISW along with Sesil Charles Latemba and Prisca Mgomberi from the DSW for their help. We also thank members of the Tandika village community for their involvement with the project.

Special thanks are due to several advisors to TechBridgeWorld's iSTEP 2009 internship program which provided an avenue for the research conducted with this project: David Dausey, Karumuna Kaijage, Roni Rosenfeld, and Jahanzeb Sherwani.

The research reported in this publication was supported in part by discretionary gifts to the TechBridgeWorld research group at Carnegie Mellon University, by the Qatar Foundation

for Education, Science, and Community Development, by the support of National Priorities Research Program grant #30-6-7-91 from the Qatar National Research Fund, by Carnegie Mellon University in Qatar, by the University Computing Centre, Ltd located at the University of Dar es Salaam in Tanzania, and by the Holleran Scholars Fund. The opinions expressed are those of the authors and do not necessarily represent the views of any of our sponsors.

REFERENCES

- [1] R. Abimbola, H. Alismail, S. Belousov, B. Dias, M. F. Dias, M. B. Dias, I. Fanaswala, B. Hall, D. Nuffer, E. Teves, J. Thurston, and A. Velazquez, "iSTEP Tanzania 2009: Inaugural Experience," Technical Report CMU-RI-TR-09-33, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, 2009.
- [2] The Joint UN Program on HIV/AIDS (UNAIDS), 2008 Report on the Global AIDS Epidemic. [Online] Available: <http://www.unaids.org/en/KnowledgeCentre/HIVData/GlobalReport/2008/>.
- [3] J. Sherwani, S. Palijo, S. Mirza, T. Ahmed, N. Ali, and R. Rosenfeld, "Speech vs. Touch-tone: Telephony Interfaces for Information Access by Low Literate User," in Information and Communications Technologies and Development, Doha, Qatar, 2009.
- [4] The United Nations, "Innovation for Sustainable Development: Local Case Studies from Africa," the UN., May 2008.
- [5] S. Patnaik, E. Brunskill, and W. Thies, "Evaluating the Accuracy of Data Collection on Mobile Phones: A Study of Forms, SMS, and Voice," in Information and Communications Technologies and Development, Doha, Qatar, 2009.
- [6] R. J. Kaitho, J.W. Stuth, and A.A. Jama, "Application of Information and Communication Technology in Livestock Marketing in the Pastoral Areas of Eastern Africa," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Part XXX, Vol. 34, 2004.
- [7] J. Klungsoyr, P. Wakholi, B. Macleod, A. Escudero-Pascual, and N. Lesh, "Open-ROSA, JavaROSA, GloballyMobile—collaborations around open standards for mobile applications," in International Conference on M4D Mobile Communication Technology for Development, Karlstad University, Sweden, 2008.
- [8] Y. Anokwa, C. Hartung, A. Lerer, B. DeRenzi, and G. Borriello, "A new generation of open source data collection tools," in Information and Communications Technologies and Development, Doha, Qatar, 2009.
- [9] C. Tumwebaze, F. Nkuyahaga, "Epihandy Mobile – A Mobile Data Collection Tool," in International Conference on M4D Mobile Communication Technology for Development, Karlstad University, Sweden, 2008.
- [10] B. DeRenzi, Y. Anokwa, T. Parikh, and G. Borriello, "Reliable data collection in highly disconnected environments using mobile phones," in workshop on Networked Systems For Developing Regions, Kyoto, Japan, 2007.
- [11] T. Parikh, E. Lazowska, "Designing an Architecture for Delivering Mobile Information Services to the Rural Developing World," in International World Wide Web Conference, Edinburgh, Scotland, 2006.