Components of SMS Based Data Collection and Service Delivery

Matthew Berg, and Vijay Modi

***Abstract*—An overview of the components, approaches and techniques used to build mobile phone accessible applications that use SMS text messages as a conduit for data collection and service delivery. SMS based applications represent a paradigm shift allowing innovative new approaches to monitoring and data collection fundamentally changing the way we can approach the delivery of critical health, economic and social services in resource poor settings. SMS has the potential to fill significant connectivity and service gaps, particularly for the world's poor, until data networks and phones that can support them become more ubiquitous.**

***Index Terms*—SMS based applications, mobile phone data collection, mobile phone service delivery**

I. Introduction

THIS paper presents and overview of the components, approaches and techniques used to build mobile phone accessible applications that use SMS as a conduit to data collection and service delivery.

The innovation of computer originated and terminated text messaging has allowed for the creation of rich SMS based applications that both send and receive data, in effect, extending the power of computer algorithms, networked connectivity and the resources of the Internet to any basic mobile phone. This represents a paradigm shift allowing innovative new approaches to monitoring and data collection fundamentally changing the way we can approach the delivery of critical public and social services. Specifically, this paper will address how the body of a text-message, with a 160-character limit, can effectively be used to pass data and issue commands to perform specific tasks on a remote computer. In the process, the authors aim to define a common terminology for the components, techniques and approaches used to do this.

This paper will not seek to address or compare the wide variety of software platforms such as Kannel, Gnokki, FrontlineSMS and RapidSMS that have been developed to facilitate the computer to SMS link.

II. SMS

Short Messaging Service or SMS was developed in 1984 by the GSM cooperation. Its 140-160 character limit was born out of the necessity to constrain the message size to less then 128 bytes, the packet size that would allow messages to be transported at a nominal cost along the same signaling pathways that controlled network traffic [1]. Since the first SMS message was sent from one mobile phone to another in 1993, SMS has grown to be the most widely used data application in the world with over 3 billion active users (compared to email with 1.4 billion); sending over 12 billion SMS text messages every day [2].

Text messages in languages that use Latin characters, like English for French, are limited to a maximum of 160 8-bit characters while text messages in languages, like Arabic or Chinese, which use non-Latin characters are limited to 70 characters due to required 16-bit Unicode encoding.

Compared to IP based GPRS data, SMS is a very expensive way of transmitting data on a per-bit basis. An SMS message represents only 0.1367 kilobytes[1] of data. Assuming that the price of an SMS is approximately equal to the price of 1 MB of GPRS data[2] the cost of transmitting data over SMS is approximately 7,500% more – over $345 per megabyte. Given this, SMS should only be considered as a means of transmitting serialized data between machines in conditions when networks are not available and data volume transfer needs are low.

From a human communication standpoint, 160 characters is not that limiting. Freidhelm Hillebrand, one of the inventors of SMS, discovered after careful analysis that most postcards and telex messages on average are actually less than 150 characters in length [3]. The popular messaging service Twitter also limits its messages to 160 characters. A 160 character text message, as this paper will discuss, provides adequate space to send and receive both qualitative and quantitative data via simple interfaces with which a human is able to understand and interact.

III. SMS for Data Collection and Service Delivery

SMS has successfully been used for a wide variety of data collection and service delivery applications including market pricing, job board, and mobile banking services in business and in mobile health, behavior change public health messaging, treatment adherence for chronic disease, medical results delivery, disease monitoring and drug stock inventory management.

SMS based applications offer a number of advantages that are particularly relevant in the context of service delivery in resource-poor settings.

A. Access

Since nearly all basic handsets support SMS and an increasingly large share of people in the world have mobile

M. Berg the Director of ICT for the Millennium Villages Project, Columbia University, New York, NY 10027. Phone: +1 646 463–1273, e-mail: mberg@ei.columbia.edu

Vijay Modi is a professor of Mechanical Engineering at Columbia University and the Director of Infrastructure for the Millennium Villages Project. e-mail: modi@columbia.edu

[1] 140 octets = 140*8 bits = 1120 bits/8 = 140 bytes/1024 = 0.1367 kilobytes

[2] $0.046 USD per SMS and $0.053 USD per MB, Zain Kenya 2010

phones, the number of people – even the very poor – who could potentially access SMS based services is dramatic. If users do not have to be provided with phones, SMS based applications may have a large user-base without needing to incur the costs of purchasing phones. One of the main and most challenging barriers to accessing SMS services, however, will remain literacy.

SMS messages are queued by the network, which means that a user will not lose messages when their phone is turned off or they are out of coverage. This is particularly important in areas of the world where mobile phone charging and sporadic network coverage remain real problems. SMS delivery is "best effort" meaning that the network will make a best attempt to deliver a message but not guarantee receipt. Issues like network congestion, for example, usually result in delays of delivery and not lost messages.

Phone based clients, typically J2ME applications written in Java, are useful in that they can help conform data entry and provide support for client-side logic. Mobile phone based clients, however, require a more expensive feature phone to run and require the user to download the application to their phone. SMS based applications require no client and are thus immediately accessible on any phone. More advanced SMS based applications that take specific formatted texts as input typically will require some training to use.

SMS based services do not require data coverage to operate and are thus available anywhere there is mobile phone coverage. This is important as many remote, resource poor settings still lack data coverage. Even when there is network data coverage, many mobile phone operators make the process of accessing data services too difficult for users by requiring them to activate their SIM and then enter configuration settings into their phones.

B. Managing Credit

One of the most challenging aspects of implementing mobile phone based application projects is controlling costs and managing credit. A key advantage of SMS is that it provides the ability to setup a toll-free SMS number that is free to text. This defrays the cost for the end user who does not need credit to participate. Providing mobile phone credit can be logistically difficult and hard to ensure that credit is being used for its intended purpose. Toll-free SMS numbers also help ensure participation in public campaigns where you have a large, unknown base of end-users.

Setting up a toll-free SMS number can be an involved and costly process since it often requires working directly with a local mobile phone operator. Getting a short-code, useful for large programs, usually involves an application process with a local government communication agency and then working with multiple operators to support an aggregated number. Fortunately, in many countries it is possible to find companies, whom acting as middleme'n, facilitate the setup of toll-free numbers and short codes.

While GPRS provides a cheaper way to transfer data, there is currently not an equivalent of a toll-free number for data services. The ability to zero-rate bill data for a phone accessing a certain IP address (the application server) would be a useful service that would allow for this.

C. Mobility

Mobile phone based services offer the advantage of mobility since the end-user can access them at the location or time that they are most useful. Data collection can be done *in situ* providing a continuous measure of data. Knowing the time and frequency in which data is captured, is useful in itself, creating paradata that can be used to detect things like fraud or to monitor an end-user's activity levels.

IV. Components, Approaches, Techniques

Structured text messages provide the input that makes SMS based data collection and service delivery possible. Similar to a command line interface, the body of the text message is parsed and using a delimiter, usually a space "_", is broken into a number of tokens which the interpreter converts into a series of commands and parameters used to perform specific tasks on the remote machine.

A. SMS Reports

A *SMS Report* represents the data captured in an SMS form sent by a reporter's text message. The SMS report combines the data from the form with metadata from the form submission including when it was submitted, by whom and even at times where.

B. Reporters

A *reporter* is an end-user who uses SMS reports to send data to the system. Reporters are usually identified by the number of the mobile phone they use for texting.

Many applications will require the ability to register reporters to control which functions they have access to. To do this most SMS applications support a registration function, SMS or computer based.

```
>> REGISTER CHW tiby Mamadou Coulibaly
<< REPORTER REGISTERED! CHW Mamadou Coulibaly
   @mcoulibaly Location: Tiby. M: 421-4248
```

In the example above, a community health worker (CHW) Mamadou Coulibaly, is registered as a reporter and assigned the alias @mcoulibaly which can be used for user messaging. The reporters profile is often used to set the default metadata for the reports a user sends in. For example, the location of the CHW is often used to define the default location for their patients registered.

Applications such as ones that involve crowd sourcing, will need to accept SMS reports from unregistered or anonymous reporters. While a system will not know the identify of an *unregistered user*, it will still accept messages from that user and potentially build a profile on that user based on his message history. For applications with great data sensitivity that require anonymity, systems can be designed to support *anonymous reporting* where rules are explicitly set not to store any data that could later link the sender to any report that they send to the system. This

usually involves omitting the phone number as incoming messages are logged.

C. Keyword Response

A *keyword response* is when the body of an incoming SMS contains a predefined keyword, that when found, triggers an action, usually a specific text response which can be either static or dynamic. A simple keyword response application could, for example, return the price of local crops whenever the keyword FOOD is texted in. Another common example is the use of the keyword JOIN to subscribe a phone-number to an opt-in SMS based news service.

Keyword responses trigger actions. They do no support parameters making their use with data collection limited. Keywords responses can be used in series to generate a menu like interaction, similar to that of USSD. For example, an incoming message with the keyword HEALTH could trigger the reply:

```
For more information on HIV/Aids please reply
with HIV or BABY for information related to
newborn health.
```

D. SMS Forms

An SMS form typically is space delimited and consists of a *form keyword* followed by a predefined number of required parameters and then an undefined number of optional parameters. When not prefixed with a delimiter like "+", the form keyword is defined as the first word in the text message. In such cases, it is only possible to have a single form in an SMS report.

```
MUAC AB4D 105 Y F V
```

In this example, the SMS form uses a mid-upper arm circumference (MUAC) measurement and an oedema check to diagnose if a child is acutely malnourished. MUAC is the form keyword and the second parameter Patient ID is the child for whom the acute malnutrition test is being done. These are followed by the MUAC measurement and oedema status [both required]. Lastly, there is optional space for any risk symptoms, in this case F and V for fever and vomiting, that the child may be presenting.

Form keyword selection requires striking a balance between brevity and the ability of the keyword to act as a mnemonic device. Single or double character keywords, without a delimiter prefix, are not ideal as they compete in the same namespace with parameters and can be easily confused with them. Parameters can be any string object used to represent words, dates or numbers. In some cases, parameters may have their own delimiter like a dash "-" to break up a parameter into sub-variables which is useful when the parameter is optional. For example, the following SMS could be used by a clinic to report 52 cases of malaria and 5 cases of tuberculosis.

```
DISEASE MA-52 TB-5
```

Since all parameters in this form our optional, if there was no incidence of tuberculosis the TB-5 could simply be omitted. The end optional parameter can also be used for free-text to record a qualitative note related to the report.

A common question relating to SMS forms is how many parameters can a single SMS form contain? While an SMS message can support up to 160 characters, it is not advisable for a form to exceed 5-6 parameters. This is particularly the case when using basic phones with limited single screen viewing space. Multiform SMS reports could potentially provide better support for additional parameters but this is yet to be verified.

E. Form Processing

There are a number of things that can be done to account for basic human errors that occur frequently during data entry. SMS forms, for example, should never be case sensitive. The SMS can also be preprocessed with some string cleaning functions that do things like remove double spaces, swap letters for numbers, and remove periods when between numbers.

Once the form has been parsed into its components, additional logic can be added for each parameter type to intelligently support things like various date formats and automatic age calculations. When selecting delimiter or prefix characters, it is important to chose characters like "+" or "-" that are easy to enter on the phone and do not require an "insert symbol" command. Additionally, when entering the text on the phone it is best to always instruct reporters to use the multi-tap method and not T9 mode.

F. Form Logic and Responses

Form logic takes the inputs from the processed form and converts it into specific commands to be executed by the computer. Examples of this include updating a database record, performing a calculation or triggering an alert. For SMS based applications, it is important that for every SMS message sent there is a corresponding response to indicate that the SMS was received and successfully processed, or to indicate if an error occurred. Since data entry errors, like a misspelled name, can occur in a properly formatted message, responses should be as descriptive as possible to confirm specifically to the reporter what was sent.

```
>> new diallo fatimata f 151209 mary tiby
<< Patient #AB4D. DIALLO, Fatimata (F/6m)
   Mother: Mary. Village: Tiby. Registered!
```

In this example, the NEW calls a patient registration where the name, gender, date of birth, mother's name and location is recorded. The response message for the patient registration confirms not only the patient registration but returns a generated patient ID, and the full patient details including the age of the child based on the date of birth entered. In cases where the date of birth is unknown, the function could also accept the age in months (14m) or years (4y) and then estimate the birth date. If a mistake was made, granular error messages that show explicitly where the error occurred, either on the data entry or logic side, enable the sender to learn how to correct their mistakes.

G. Multiform SMS Reports

Adding a prefix to the keyword form makes it possible to support multiple forms in a single SMS report.

```
AB4D +MUAC 140 N +MAL Y +MED ACT +REF B
```

In this case, there are four SMS forms that are being applied to the patient with Patient ID AB4D . While the patient has a healthy MUAC (acute malnutrition <125), the malaria form (+MAL) shows that the child is positive for malaria has received ACT from the medicine form (+MED) and therefore needs a basic (B) referral (+REF).

Multiform SMS reports offer a number of advantages. First, prefixing each form with a plus symbol "+" makes it easier to visually parse the form which is helpful when entering the SMS form into a phone. Second, it is possible to break larger forms into smaller more functionally specific forms with fewer parameters facilitate memorization – even taking on a sentence-like mnemonic structure. Third, since SMS forms can be sent in any order and are by nature optional, it allows great flexibility when creating paper-based versions of the forms. It is possible, for example, to design and program a wide variety of functions represented in forms in advance to be used later by a non-technical team. The logic chosen to incorporate a selection of SMS forms into paper forms will structure how the data is reported.

Multiform SMS reports allow great flexibility to the reporters using the system. By simply adding +NOTE, a reporter is able to append an optional note to any information sent in. Reporters can also use parameter-less forms (ex. +TWIN) to tag an object in an SMS report. The *tag*, like any form, can be used to trigger an action or append metadata to a form object.

Form keyword prefixes also allow SMS form(s) to apply to multiple objects. For example:

```
ID1 ID2 ID3 +MEASLES
```

can be used to record that three patients have been vaccinated.

H. Multiple SMS Reports

A *multiple SMS report* combines multiple individual SMS reports together into a single report object with an associated confirmation ID. An example of this is a weekly report to capture key aggregate disease and treatment data from a clinic. Since the number of indicators would exceed the capacity of a single SMS report, *multiple SMS reports* can to be used to capture the data and ideally would be designed into a paper form.

As each report of a multiple SMS report is sent in, the confirmation message should include a status ratio indicating the completion status of the report. For example:

```
Thank you! Drug Stock report received (2/4)
```

Once all required reports are sent in, the reporter would then receive an additional message confirming the report, consisting of multiple text messages, for that week was complete, including a confirmation ID.

I. SMS Based Data Correction

There are several techniques for data correction that can be used to help handle SMS based data entry errors. For some SMS reports, like the MUAC report presented earlier, it is possible to correct the earlier submission by simply resending the data for that report within that same day (or predefined time-frame). This will over-write the preexisting report and update it with the new status.

In the case of a patient registration, where a unique patient ID is assigned, resending the SMS would risk creating a duplicate record. In this case, a CANCEL command could be sent to negate the last SMS report received by that reporter to the system. This technique is also useful for canceling all the forms in a multiform SMS report. For multiple SMS reports or some multiform SMS reports, it can be useful to have the system return in all confirmations a transaction ID for canceling. For example:

```
CANCEL 4534
```

In all cases, it is important that a confirmation is sent providing clear details of what happened.

J. SMS Surveys

SMS surveys is when a reporter is contacted by the system and is prompted to answer a single or series of questions that the user can then reply to usually with a yes/no, number or qualitative text answer. In some cases, SMS surveys can be triggered by the user but most often are system initiated. This is often the case in public health messaging campaigns.

With SMS surveys maintaining state is very important and is usually done by setting a time expiration per question, tied to the mobile phone number of the reporter. In other cases, the design of the questions can actually help maintain state by alternating yes or no questions with ones that require a numeric or qualitative answer. Support for basic natural language processing that can handle multiple aliases for potential responses, including common misspellings, (N, No, nope, non, noo) is also important.

K. SMS Diaries

Similar to SMS polling, an *SMS Diary* is when a reporter is prompted on a set frequency (daily, weekly) to answer predefined questions. In a health context, SMS diaries are often used to help patients improve self-management of chronic disease through *in situ* self-reporting of their treatment and health condition status[4]. This is based on the premise, known as the Hawthorne effect [5], that people change their behavior and perform better simply by being observed.

L. Alerts

An *alert* is a message sent to by the system to a user triggered by an SMS report. For example, when a CHW diagnoses a child with acute malnutrition via an SMS report, an alert can be automatically sent both to the doctor at the childs referral clinic and to the local nutritionist.

M. SMS Reminders

A *SMS reminder* is a timed-based reminder for a reporter to perform an action (execution of which can be captured in an SMS report). In health, SMS reminders can be used for a patient to take a treatment or CHW to follow up on a patient.

N. SMS Friendly Unique IDs

Many applications, particularly in health, require the ability to assign unique identifiers to people or objects. With patient identifiers, a unique ID is required. A persons name is not always unique and a patient ID provides an important mechanism for protecting the privacy of an individual and therefore must not contain any content that can be used to be link back to the patient.

With SMS it is particularly important for the ID to be concise. Therefore the use of alphanumeric characters instead of a purely numeric system is ideal. One method particularly well suited for SMS is that of a base-30 ID system using the following 30 alphanumeric characters:

0123456789ACDEFGHJKLMNPRTUVWXY

Which represents 0-9 and the full alphabet with the exceptions of B, I, O, Q, S and Z because they resemble numbers. A 4 digit base-30 ID provides $30^4 = 810,000$ unique IDs; with a 5 digit base-30 ID system allowing over 24,300,000 IDs. A fifth check digit character is generated to verify the validity of the ID using the Luhn function, a checksum formula used for credit card validation.

O. User Messaging

SMS applications typically support three forms of user messaging. *Direct messaging* allows users in the system to message each other using via their alias denoted with an @ symbol.

```
Eve   >> @adam How about them apples!
Adam  << @eve> How about them apples!
```

Group messaging works the same way but the user alias is replaced with that of a group's. With group messaging, it can be important to limit access to sending to the group to certain key members. *Message broadcasting* or *blasting* done through an SMS application interface, provides the ability to send an SMS message to a group of reporters, groups of reporters or a list of individually selected reporters in a system.

P. Recording Transactions with SMS Handshakes

SMS Handshakes provide an elegant means to record confirmed actions, usually a transaction, between two reporters.

```
Store >> TRANSFER @pharmacist act-100 asp-20
Pharm << @Store has transferred you 100 pkts
         of ACT and 20 pkts of aspirin.
         Please confirm by replying: YES
```

Once the pharmacist replies yes, the storekeeper will get a confirmation message of the transaction. Alternatively, the application can be designed to not require a confirmation but send a CANCEL with the transaction ID if the transaction is not confirmed.

Q. Importance of Paper

Paper still plays an important role in SMS based application systems. For more complex multiform SMS and multiple SMS reports, paper forms provide an important tool to help the reporter record the data before transmitting it by SMS. For simpler SMS forms, cheat sheets on card stock or laminated paper serve as helpful reminders when transposing data directly to SMS.

To help ensure value and sustain use of the system, it is critical for information to flow back to the reporters. While SMS can be used in many ways to achieve this, system-generated paper based reports can provide a rich and effective interface for synthesizing data back to the end user.

R. SMS Security

The SMS has a number of security and privacy issues affiliated with it. Messages that are sent or received, unless frequently cleared, are stored in the history (Outbox/Inbox) of the phone. While SMS data transfer over a network is reliably secure, the SMS messages are stored in plaintext in the operator logs. While many operators have a practice of clearing their logs, this does present a potential risk when transmitting sensitive patient data and can also be exploited by governments wishing to monitor their people.

V. Conclusion

A simple text message can be used in a structured way to enable any basic mobile phone to be used for robust data collection and service delivery. SMS based applications allow old problems to be addressed in new ways by streamlining team communication, improving monitoring and providing information and decision support in remote and resource constrained conditions. SMS has the potential to fill significant connectivity and service gaps, particularly for the world's poor, until data networks and phones that can support them become more ubiquitous.

References

[1] "SMS." Wikipedia: The Free Encyclopedia. Wikimedia Foundation. n.d. Web. 25 Mar. 2010.

[2] Ahonen, Toni. "Faster than a locomotive yes SMS text messaging twice the users of Email." 25 March 2010. http://communities-dominate.blogs.com/brands/2010/03/faster-than-a-locomotive-leap-over-giant-buildings-yes-sms-text-messaging-twice-the-users-of-email-t.html

[3] Milian, Mark. "Why text messages are limited to 160 characters" LA Times, 3 May 2009. http://latimesblogs.latimes.com/technology/2009/05/invented-text-messaging.html

[4] Cleland, Jennifer (2007). A qualitative study of the attitudes of patients and staff to the use of mobile phone technology for recording and gathering asthma data. *Journal of Telemedicine and Telecare*. 13, 85-89. http://www.jtt.rsmjournals.com/cgi/content/abstract/13/2/85

[5] McClusky, Mark. 'The Nike Experiment: How the Shoe Giant Unleashed the Power of Personal Metrics." Wired Magazine. 22 June 2008. http://www.wired.com/medtech/health/magazine/17-07/lbnp_nike?currentPage=all