ATMosphere: A System for ATM Microdeposit Services in Rural Contexts

Michael Paik
New York University
mpaik@cs.nyu.edu

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

Abstract—**This paper describes strategies to lower the cost of providing Automated Teller Machine microdeposit services in rural contexts. Microdeposits represent a growing market in the developing world, but the cost of running a conventional ATM network is prohibitive due to the capital investment required to deploy networks and terminals.**

Our novel contributions are to use the Short Message Service (SMS) over high-penetration GSM cellular networks in conjunction with a system using location awareness to intelligently distribute available balances among machines. This allows us to provide high levels of service while simultaneously reducing risk to the financial institution and lowering per-transaction cost.

Using a simulation of ATM usage patterns and distributions, our primary results under our model are: (1) transaction cost per user per year can be optimized to less than USD 0.18 given an SMS loss rate of approximately 10% while (2) customer withdrawal success rate can be maintained at approximately 98% with (3) a maximum of 5% of funds on deposit available in cash in ATMs at any given time.

These results make wide deployment of rural ATM services by financial institutions feasible and economically viable in the near term using existing commodity technology.

Index Terms—**Finance, Financial data processing, Simulation.**

1 INTRODUCTION

READY ATM access at the village level has been an open problem in the microfinance field, as some means for making deposits and withdrawals of small amounts of capital is necessary for the inception and growth of microfinance markets and personal accumulation of wealth for the rural poor. However, for-profit institutional financial entities have found the idea of providing such access to populations living outside of urban areas with both high teledensity and population density unattractive because the amount of capital represented by rural microdepositors is only institutionally significant once aggregated, and the cost of providing access to these customers has historically been far greater than any revenue gained by doing so. This is particularly true with the high capital investment required for traditional brick-and mortar retail banks, which also incur the cost of trained staff. More recently, systems such as M-PESA[27], operated by Safaricom in Kenya have made person-to-person payment using mobile phones viable. However these, while providing a convenient method to send funds from the city to more rural areas, do not provide significant capacity for savings as they are intended for person-to-person money transfer rather than the accumulation or retention of wealth.

This paper discusses findings from a simulation of ATM transaction behavior using real-world population and geographical data from Arua, Moyo, Yumbe, and Adjumani Districts in Northwestern Uganda, bordering Southern Sudan and the Democratic Republic of the Congo, both of which have been in political upheaval and/or civil war during the past decade. This area was selected because 1) the population represents the poorest of the poor, being made up primarily of refugees from conflicts in bordering nations who have fled to an otherwise sparsely populated region[11] and 2) GSM mobile coverage of the area is relatively dense[23] as seen in Figure 2. As such, this milieu is a real-world example of settlement patterns among the most disenfranchised. The results presented here focus on satisfying a high level of consumer withdrawal demand without imposing artificially low withdrawal limits. We present a novel scheme for balance partitioning and message flow control which makes use of location awareness of ATMs alongside other optimizations to provide a solution that is performant both in terms of risk to the institution and service to the end customer.

2 THE PROBLEM

2.1 The Economic Challenge

While the precise cost breakdown of deploying an ATM in a rural environment will vary from context to context, a significant part of the marginal cost overhead is in the creation and use of the network to support communication between the managing financial institution and the ATM. There is little in the way of existing network infrastructure (including wired telephone service) and the cost involved in laying even a single link given high costs of materials, transportation, and scarcity of skilled labor, construction equipment, and telecoms equipment is often prohibitive. The traditional wired model (used e.g. in the United States) is infeasible due to the capital investment required to lay network cables to sparsely populated regions. This is also a

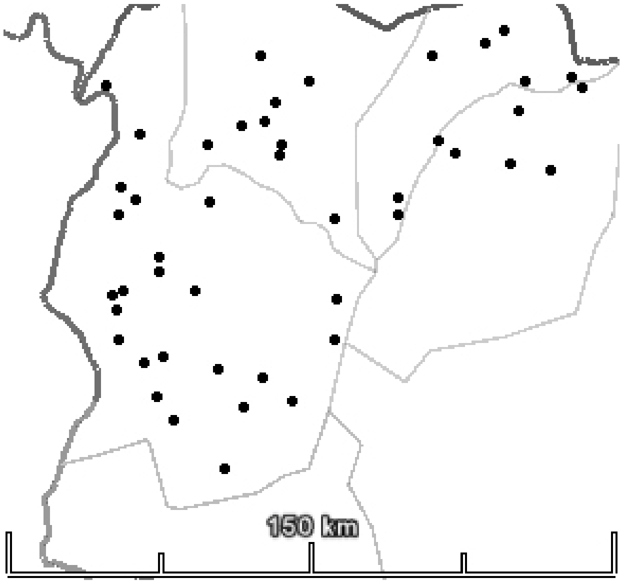

Fig. 1. Population centers in northwestern Uganda

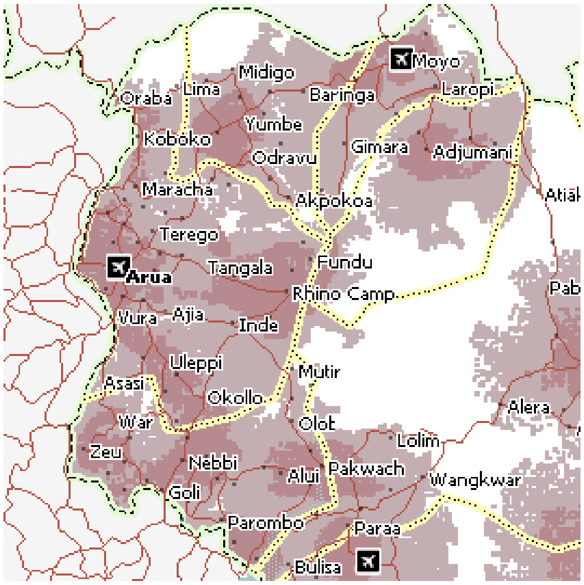

Fig. 2. GSM network coverage in northwestern Uganda (darker areas indicate better coverage)

primary factor in the proliferation of self-contained cellular telephone towers which uplink via satellite or fixed wireless transmission. Other than the existing wired model, there have been efforts to connect ATMs to financial institutions using long-distance wireless links[6] which also provide telephony and internet services, but the case for capital investment of this sort for areas in which cellular telephone service already exists is more difficult to make as some degree of connectivity is already present. Any solution which proposes to address this problem must take into account lack of network capacity as well as keeping cost of deployment and operation low enough to be practicable for national banks in the region to afford.

We use the wired ATM metaphor rather than simpler kiosks which read smartcards with encoded balances or paper passbooks because these methods are vulnerable to counterfeiting, theft, loss, and damage and have poor provisions for centralized auditing.

2.2 *The Scope of our Approach*

The approach presented in this paper is aimed squarely at reducing cost and making maximum use of existing capacity in the form of the installed base of GSM network cells. As the network's reliability, latency and bandwidth are not subject to any guarantees, we create a protocol which is designed to be resilient against extended outage conditions and slow message delivery while simultaneously maintaining an extremely low message budget and high quality of service.

2.3 *The Technical Challenge*

The use of the installed GSM base is a major benefit to the design of a solution to this problem, but there are underlying difficulties, not least of which is cost. A naïve approach to the problem might be to send a series of messages back and forth between the ATM and the financial institution – one to request a withdrawal, one to issue an identity authorization challenge, another to supply the response, yet another to issue an approval, one to acknowledge the approval, etc. etc. While the absolute cost of these messages is not extraordinarily high, the cost does not scale well as the user base increases. If the institution pays for the messaging, the cost eats into any financial gain the client may realize from interest on his savings if he is expected to pay, and in either case this approach does not scale well as loss rates may be expected to increase if message traffic increases substantially while network capacity remains fixed.

3 DESIGN

Our initial objective is merely to provide some form of ATM service to rural customers with small amounts of personal savings. Our simulated testbed contains 46 population centers of sizes varying between 2001 and 70872, totaling 637952 in total. We simplify this by considering households as accountholders, using the average household size across Uganda of 4.7, yielding 136541 households if we consider each center separately and round down.

Having ruled out wired infrastructure, our next observation is that long-distance[10] wireless is feasible as a physical network layer, but has the disadvantage of requiring potentially extensive infrastructure deployment depending on the geography of the region, distances between nodes, and the network topology necessary to scale with growth of the size of the network. Newer technologies such as WiMAX[30] may allow connectivity between adjacent villages with commodity hardware, but how much effective bandwidth a rural WiMAX mesh installation will have is still an open question, and there is still the question of deployment lead time and network maintenance costs. GSM cellular technology is widespread in the developing world, and reaches into some of the most sparsely populated regions. While the General Packet Radio Service (GPRS) is part of GSM Release 97 and newer (colloquially referred to as 2.5G and 3G), it is not a viable option for data transport as much of the installed base in rural areas is comprised of cheaper 2G cellular hardware. Due to these factors, we choose to use SMS.

Once connectivity is established, a natural design principle for our system is that we want to reduce network traffic as much as possible, as this represents a marginal cost to our system, and cost is precisely what we wish to reduce. In opposition to this is a desire to maintain overall responsiveness of the system as perceived by the customer. This perception is directly related to satisfaction rate of withdrawal requests, and tied only weakly to the immediacy of funds availability from deposits, as existing microdeposit schemes often take up to a month[13] to register updated balances. A final critical design principle is that, given the intermittent network communication enforced by our traffic reduction, we need to intelligently distribute balances in such a way as to prevent any user from being able to game the system by visiting multiple ATMs within a given update window and withdraw more than his account balance.

3.1 Simplifying Assumptions

In this paper we make several simplifying assumptions. Firstly, we assume that the only parties in communication for these rural ATM systems are the machines themselves and some server-side system controlled by the host institution. This neglects considerations of legal regulations which may enforce some neutrality or transparency upon financial institutions. We also assume that because of this simple two-party communication when considered in the realm of a single ATM, that communication scales linearly with transactions ($O(n)$) for a given ATM. In addition we observe that there is a limit to the number of transactions per day based on human factors. Assuming that an ATM is available for 12 hours per day from 8AM to 8PM, and a transaction takes approximately one minute to execute, we estimate a rough upper bound of 720 transactions per day per ATM.

We further consider all physical security considerations outside the scope of our problem; we assume that any machine is located in some safe environment where the likelihood of tampering is low, and also that any tampering of the machine will be made evident using standard techniques currently used in other devices, such as credit card swipe readers. A possible solution to this problem is to use a similar model to the M-PESA system and locate ATM kiosks within vendors' shops, allowing the vendors to charge a small fee for their use.

We assume that any losses incurred through physical attacks on the machine are covered by the financial institution, as with machines in urban locations. In the same vein we neglect considerations about powering the machine (solar is an option for low-powered embedded computers), how deposits will be marked, validated, and conveyed to the host institution, and the technical details of the tokens used to identify the users of the machine other than that these tokens contain some cryptographically secure writable medium, one such device being a smart card (ICC).

Finally, we assume that we may neglect certain transaction metadata, such as the exact timestamp on each transaction, or that several transactions occurred within a given timeframe, preferring rather to reduce the communication budget required to send information about a given transaction by reducing timestamp granularity to an hour or day-level timescale based on the aggregation window of transactions and potentially aggregating multiple transactions which occurred close together in time into one larger transaction (e.g. five transactions that happened at the same ATM on the same account within 15 minutes of each other may be represented as one transaction which is the sum of those individual transactions, at an approximate time.)

3.2 SMS

The Short Message Service (SMS) was defined in 1985 as part of the draft GSM specification[22] and as part of the GSM standard, shares GSM's ubiquity. As GSM is the cellular technology most used in the developing world, SMS as a data protocol has a significant installed base that can be leveraged with little cost. The SMS format can carry 140 bytes per message, at a cost of approximately 0.05 US Dollars per message in the region our testbed represents[25]. It is important to note, however, that SMS does not guarantee delivery, nor does it guarantee in-order delivery for those messages that are delivered, though it is a store-and-forward system. As such, our protocol has to be robust against sporadic dropped packets and out-of-order delivery. Additionally, the extremely small payload of the messages requires us to address the issues of packet fragmentation and careful data management. GSM does in theory support SMS concatenation of up to 255 messages, but in practice most systems do not support any more than 10, providing a theoretical aggregate payload of 1400 bytes. However, SMS concatenation has a per-message payload cost of 7 bytes per packet for every message after the first, which is a significant amount of overhead given the payload size. Also, in our testing, 1337 bytes proved to be too few to take advantage of any generalized compression algorithms, with compression of most payloads resulting in an *increased* size of between 5 and 60 percent, depending on the algorithm and the structure of the data.

In areas where GPRS is available, most providers in the developing world provide options for Multimedia Message Service (MMS) messaging, an extension of SMS which allows arbitrary payload sizes. MTN provides MMS messages of up to 100 kilobytes for approximately USD 0.125 in urban areas of Uganda, which represents close to a 300-fold increase in messaging budget per unit cost. Naturally, as this technology expands, marginal costs will decrease commensurately. Furthermore, even where GSM coverage is not available, SMS service is available via satellite phone coverage (e.g. Thuraya[29]), which allows any system built using an SMS metaphor to be generalized to practically any location on the globe where a clear line of sight to the sky is available, albeit at far greater cost.

3.2.1 SMS Delivery and Submission Rates

On-field tests conducted once every 30 minutes from Adjumani (labeled in the northeast quadrant of Figure 2) for the 12 hour window from 8AM to 8PM to a handset in Kampala using a handset sending 50 pre-written 140-byte (160 characters with septet encoding) dummy messages by hand yielded a minimum send rate of approximately 8 messages per minute during peak times. Tests using a script submitting the same 50 messages over the same period on a subsequent day to MTN's internet SMS gateway[24] bound for the same handset yielded a minimum receipt rate of approximately 12 messages per minute. We consider these to be sane limits when the network is operational as we assume that the handset equipment used has a lower-powered radio and significantly less processing power than an ATM might, but we do not maintain that this rate is an absolute lower bound.

Loss rates during these tests were 0%, but we do not treat this result as conclusive due to the small sample size and the fact that these tests were conducted only across one day in each direction. Additionally, all messages were delivered in-order, which we also do not treat as conclusive, for similar reasons. Tseng et al. found a loss rate of 0.66%[15] after sending 915 messages over a 38 day period in 2005 in Taiwan. As Taiwan has high GSM penetration and a more modern cellular infrastructure (all but two of Taiwan's GSM providers had high-bandwidth 3G networks by October 2005) we assume that we will see greater loss rates on the whole in our testbed.

The median latency for our group of messages was asymmetric, at approximately 36 seconds from Adjumani to Kampala, and approximately 21 seconds in the other direction, including any latency introduced by the internet SMS gateway.

These latency and throughput numbers are used only to garner an overall idea of the timescale these messages operate on and are not assumed to be statistically significant.

TABLE I
TP-VP GRANULARITY AND RANGE

TP-VP	Implied Validity Period
0 - 143	(TP – VP +1) * 5 minutes
144 - 167	12 hours + (TP-VP – 143) * 30 minutes
168 - 196	(TP – VP – 166) days
197 - 255	(TP – VP – 192) weeks

3.2.2 Payload Optimizations

At a minimum, the message payloads must contain some information about depositor accounts and their balances. Based on our design principle of reducing network traffic, we wish for each payload to contain information about more than one transaction. Due to the small payload size allowed in SMS messages, we made several design decisions to decrease the size of data passed via SMS.

We observed that the smallest denomination available in Uganda is 50 Ugandan Shillings (Ushs), or approximately USD 0.03, so we chose to encode balance information as multiples of this amount using 20 bits, which gives us a maximum amount of Ushs26214350 or approximately USD 15,000, fully an order of magnitude more than any microdepositor is ever likely to have on deposit, while maintaining signing. Ugandan checking account numbers are mandated to have exactly ten digits[26], which would require 34 bits to encode, but we can establish a mapping from account numbers to a smaller unique integer; in this case we use 20 bits which gives us a maximum of 1048575 discrete accounts. This gives us a total of 40 bits for a (account number, amount) pair which sets an upper bound of 28 of these pairs per 140-byte payload.

3.3 Protocol

Our overall strategy is to encode a withdrawal transaction by indicating a negative delta to the balance indicated on a given machine, and encode balance updates from the financial institution's central authority (CA) reflecting verified deposits and redistributions of balances among machines with positive or negative deltas to the balance.

The CA maintains tables containing the GSM phone numbers of every ATM (by which incoming SMS messages are identified), precomputed pairwise distance metrics of each machine to every other machine, cash balance in each machine, home ATM(s) for each account, and account and balance records.

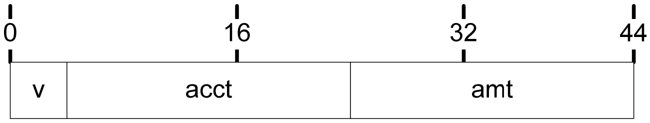

Fig. 3. Record structure

3.3.1 Message Structure

As SMS does not guarantee delivery or in-order delivery, the addition of a sequence number at some level is critical to error-free communication. While SMS has no built-in sequence number, a sent SMS message does have a user-specified validity period indicator (TP-VP[12]) not contained within the user data area (and therefore essentially free) which guarantees expiration of the message if it cannot be delivered to the destination system within a given period of time. In addition, a received SMS contains a timestamp indicating when the message was received by the SMS Center (SMSC), which is in-order for a given source as a device cannot send another SMS message until one sent prior has been either acknowledged by the SMSC or timed out. We encode each (account number, amount) pair with four additional bits which indicate a version number from 0-15 which wraps around. Combined with the timestamp on delivery, this allows us to keep updates on a given account from a given source in-order. We also set the daily transaction count limit for a given account at a given ATM to 4, meaning that an account will not experience version number wraparound for 4 days, and set the TP-VP to be 2 hours, which guarantees against version number collision.

Of the 1120 bits available to the payload, we use 1056 for 24 (version, account, amount) tuples of 44 bits each, and add a 20-bit checksum to protect against bit errors. We order these tuples by account number. Given our one-per-minute estimate of maximum transaction execution rate due to human interface considerations, this gives us an upper bound of 30 SMS messages per day for our 12 hour, 8AM to 8PM day, with the rate of messages sent by the ATM never exceeding one per 24 minutes.

3.3.2 Message Flows

3.3.2.1 ATM to CA

Each ATM sends a message either when a set of 24 transactions has been completed or some time *t* has elapsed and at least one transaction is ready to send, where *t* is 120 minutes with a random jitter up to 5 minutes in order to prevent synchronization among ATMs. If a message is not acknowledged by the SMSC within 30 seconds, sending is reattempted at most three times, then the versions of the accounts contained in the message are flagged as non-acknowledged (NAK), and operation continues. In order to reduce the number of updates sent, multiple transactions on the same account within the window of a single message are aggregated and treated as a single update and are given a single version number. However, each transaction still counts towards the transaction limit imposed by the ATM.

During a 30-minute window after the ATM ceases public operation each day, it will attempt to resend any messages that are NAK for the duration of the window to compensate for extended periods of loss.

3.3.2.2 CA to ATM

The server at the central authority records any incoming transaction records from an ATM and applies the balance adjustments to the appropriate account, updating the version number in the process. If a version number received is not the one expected based on the CA's record, it is applied, but the missing sequence number is noted.

3.3.2.3 Error Detection and Correction

In order to protect against silent failure of the CA server or an ATM, we add a heartbeat/echo message containing a random string which is triggered once every four hours by each ATM as necessary (i.e. after that amount of time has passed without a message having been received), jittered by up to 20 minutes. Any echo message received by the CA is replied to immediately with the same contents. If the CA server does not hear a heartbeat or transaction message from an ATM for 360 minutes, the ATM is flagged as 'down'.

An ATM being flagged as 'down' ensures that the operation of the local ATM continues as normal for a given account until that either account's available balance at that ATM or the cash in the ATM is depleted, during which time connectivity may be restored. An ATM is brought back 'up' whenever an update is received from that ATM. Thrashing is possible, but the timescales involved reduce the problem, and no additional resources are consumed through the CA considering a machine 'up'. Any machine which is flagged as 'down' during two consecutive batches is flagged as 'disconnected' and becomes ineligible for any balance distribution updates from the server until manual intervention can be scheduled.

At the end of the day after the ATMs have ceased public operation, the CA waits for 1 hour for any incoming updates that may have been lost, and applies any received changes, then explicitly queries ATMs for any messages still missing using a series of messages containing only the missing version number and the account number. Any ATM which still has questionable data at the end of one hour after the final update request is sent is flagged as being 'down' and is removed from the set of ATMs eligible for balance redistribution.

After the day-end synchronization batch described above, the CA sends each ATM that is 'up' a message containing a hash of the concatenation of all versions, accounts, and balances that ATM is known to be aware of, ordered by account number and version number. Each

ATM computes the same hash upon receipt, and an echo containing the ATM's hash value is sent back. Any ATM which returns an inconsistent value is flagged as 'disconnected' as above.

Messages which are received but fail checksum and cannot be repaired are discarded and considered lost.

3.4 Security

SMS over GSM is susceptible to a variety of attacks[7]. In particular, it was indicated in a personal communication to the author by a member of MTN Uganda staff that SMS is transmitted in plaintext throughout its network, and this is the default behavior[7]. As this is the case, and as both Originating Address (OA) spoofing and realtime sniffing of SMS messages are available in the wild, it is necessary for the system to have mechanisms to protect the secrecy of balance/account information and to provide some measure of authentication of messages sent between the ATMs and the CA.

In order to maintain effective payload size while providing this type of protection, a stream cipher is used, which produces ciphertext of the same length as the plaintext. This system uses RC4 due to its ease of implementation, but any such encryption system will serve. Each ATM maintains a separate 256-bit key, and the CA maintains a keyring of all of these keys, using them as appropriate in communication to a given ATM.

A 32-bit timestamp is embedded into all messages in order to mitigate the problem of replay attacks; a timestamp will only be accepted once and any subsequent message from a given source bearing the same timestamp will be discarded. In order to mitigate substitution attacks, this timestamp is placed between arbitrary (version, account, amount) tuples and will be padded with 12 bits of 1's in order both to maintain tuple offsets and to indicate which block contains the timestamp. Version number 15 then becomes reserved as not to restrict the range of account numbers which may be used.

3.5 Balance Distribution

The practice of available balance distribution among ATMs is a hedge against users who may wish to game the system by trying to take advantage of network outages or periods of higher-than-normal latency in order to withdraw more than their account balance. We assume for the purposes of this practice that ATM use multifactor authentication where at least one factor is a biometric, such as a fingerprint. The hash of this biometric is stored on the card in order to defend against card duplication.

We begin by associating each given user with a 'primary' ATM, which is the machine closest to that user's residence, and then consider the other ATMs which are present within one 3 hour period's travel of that location. In our testbed, the geography is advantageous as it offers only a few inter-town river routes, and no tarmac-paved roads. We also use the intuition that (1) accounts with higher available balances are more attractive to overwithdraw, (2) people with higher balances are similarly more likely to have access to faster transport such as *boda-bodas*, (local motorcycle taxis), *matatus* (intra- and inter-town minibuses), and interdistrict buses or to own or be able to borrow a private vehicle, (3) that fraud associated with higher-balance accounts costs the host institution more, and (4) holders of high-balance accounts are less likely to need to withdraw all or most of their available balances at once.

We consider the maximum feasible distance covered by walking over a 3 hour period (~10km) and the maximum feasible distance covered by driving along unpaved roads during that same period (~125km). Given that the average daily unskilled rural labor wage is in the neighborhood of $1.2 per day[13] (approximately Ushs2100), we arbitrarily set the 'high balance' mark at Ushs900000 (USD ~5000). Accounts with balances of 450000 and under have a maximum of 100% of their account balance distributed across all the ATMs in the system, decreasing linearly to 75% at 900000 and beyond.

Only ATMs within a specific distance of the primary ATM are considered for this distribution, specified by the formula

$$(d_{max}/2 - d_0) * \min(1, b/c) + d_0$$

where d_{max} represents the greatest distance between any two ATMs in the network, d_0 represents the minimum radius to be considered (40km in our case), b represents the balance of the account in question, and c is the cutoff 'high balance' mark at which maximum distribution spread is reached. This encapsulates our intuition that we want to distribute the balance for those with higher balances over a larger area, and the linear increase in radius produces an increase in area covered proportional to the square of this value.

The proportion of the available balance held in the primary ATM varies according to the formula

$$b_{primary} = b * p_{max} - ((p_{max} - p_{min}) * \min(1, b/c))$$

where p_{max} and p_{min} are the maximum and minimum proportions, respectively (.8 to .4 in our simulation). These numbers are based on assumptions we make about the behavior of clients as listed below, and the formula again reflects our intuition that the further the client is likely to travel, the further afield the balances should be distributed. The balances to the ATMs within the effective radius defined above other than the clients' primary ATMs are distributed proportionately to the inverse of the square of the distance of the machines from the primary ATM. Any available balances less than Ushs250 and any residues of allocated balances modulo Ushs50 are removed and added to the primary ATM's available balance.

These two formulas together have the effect of concentrating more of a user's balance nearer to his home ATM as his balance decreases. In practice the constants and proportions are adjustable to ensure that any balances under a given amount are held entirely at the user's primary ATM.

3.5.1 Migration

The initial germ idea for this project called for support for highly migratory populations to have continuous access to funds. We solve this problem by introducing a notion of carrying 'virtual' available balance on a credential token used to access the system, such as a smart card. This 'virtual' balance would involve 'locking' a specific amount of available balance on one ATM and writing a device-signed token for the amount onto the card including a unique transaction ID allowing the user to take the card to some specific other ATM and withdraw the amount 'locked' plus any available balance on the local ATM. This could be useful for trips where larger amounts of capital are required but carrying cash is dangerous, e.g. seasonal trips to buy seed for planting or taking animals or produce to market. Using the extra balance would cause the local machine to create a new signed token with any remaining 'locked' balance and inserting it into the original machine would unlock any such remaining portion and treat the spent portion as a withdrawal. This has the advantage that if the card is lost or stolen, funds which are locked on the card but have not been unlocked anywhere across the ATM network can be recovered from the institution.

3.5.2 Multihoming

We wish to model one commonplace situation, in which a head of household or other member may work in a remote region for most of the year and return for several months or a season. In such cases, the person who is working away from home typically wants to send money back to his household – this is the problem that approaches like M-PESA address. We model this situation by explicitly adding the single ATM closest to the worker to the set of ATMs considered for balance distribution, if it is not already in the set. While the typical usage of such an ATM will be solely to deposit funds left over after cost of living is paid for in cash, we arbitrarily allocate 20% of balance available at the primary ATM to this ATM. The rationale is the idea that the breadwinner represents, on average, 1/5 of the total size of the household based on the average family size of 4.7 and assuming a single primary source of income. This is trivially extended to accommodate arbitrary family configurations. The proportion that is allocated is drawn from the bottom up: any available balance outside the primary ATM is consumed first, and then primary ATM balance is reallocated until the correct proportion is reached.

4 SIMULATION

4.1 Simplifying Assumptions

There are several factors present in the real-world region upon which our simulated testbed is modeled which we ignore completely. One is familial seasonal migration from place to place – for the one-year duration of our simulation we assume that people's residences remain fixed, i.e. the primary ATM of a given account never changes; individual members may go on excursions but never the entire family at once. Another is that of immigration to the region or other sources of population growth such as refugee influx or organic growth of the number of households through marriage. We assume that all members of our simulated environment start life with accounts, that is, we have a 100% adoption rate, with some proportion of accounts beginning with a zero balance. As a final detail, we have found that we do not have enough information on the overall prevalence of bitwise errors within SMS delivery, as the literature has various accounts, and have no information at all on its prevalence in the rural context, where distance may increase and technology may be older or cheaper but there are far fewer sources of electromagnetic noise, so we ignore it, assuming that those messages not correctible by the checksum are incorporated into the loss rate.

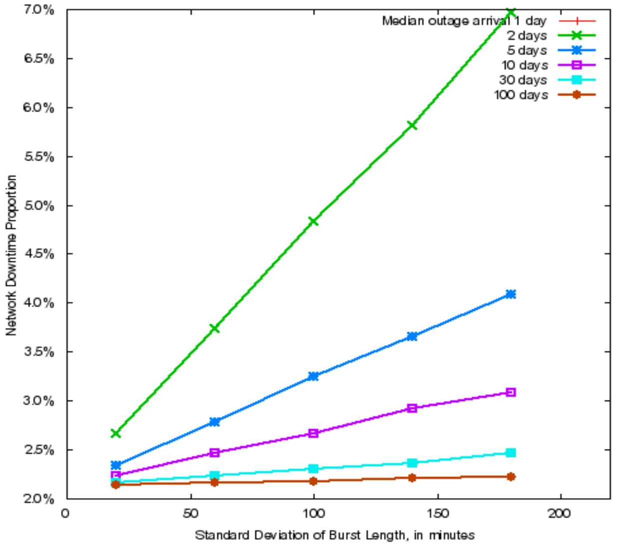

Fig. 4. Proportion of downtime versus standard deviation outage burst length for 4 values of λ.

4.2 Simulator Overview

The simulator uses minute-scale timeslices over the course of a calendar year and contains metaphors for customers, ATMs, the CA server, and the SMS network connecting the various components of the overall system. In our experiments we change various parameters of the system to see what effect they have on the cost model and customer request satisfaction rate. We consider a withdrawal request successful if, having specified some amount desired which is less than his account balance, the machine both has enough cash on hand to satisfy the request and has enough of the balance allocated to the given machine to allow the machine to dispense that cash.

4.3 Network Behavior

Network behavior parameters are the factors most directly related to the cost model. To model bursty outage behavior, we use a Poisson distribution to model burst arrival and a truncated normal distribution (with mean at 0 in which we only consider positive values) to model burst length, and change λ and σ^2, respectively for each to test the robustness of the system under various outage models, as well as how the level of customer satisfaction changes as the network and therefore the application built atop it becomes less reliable. We would also like to vary latency with a truncated normal distribution and μ set to 36 seconds as the experimental data in 3.2.1 indicated, and the σ^2 likewise set to 5 to show how the effects of variable propagation delays; however the timescale we are using renders this variability negligible. As such we set μ and σ^2 to 1 minute, which biases slower than experimentally detected, which cannot artificially improve results. Finally, we add an additional Poisson distribution-based loss model for 'black hole' loss where messages simply disappear after having been successfully sent to the SMSC.

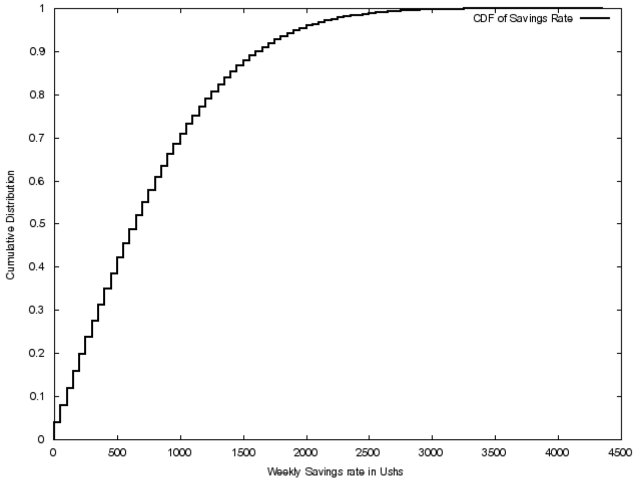

Fig. 5. CDF of customer weekly savings rate in Ushs.

4.4 Household Behavior

The 136541 households in our testbed are assigned to population centers according to real-world population levels. Each account is held by a household having certain specific characteristics with which we attempt to reasonably emulate a real population:

1) Primary ATM, allocated randomly using a uniform distribution

2) Weekly savings rate, generated by taking an arbitrary minimum savings rate (Ushs200) and adding it to Ushs1000 multiplied by the absolute value of a random element drawn from a Gaussian distribution with mean 0 and standard deviation

3) Visit frequency, a step function generated by another random Gaussian with mean 0 and standard deviation 1, by assigning 7 days if the random value is within 1 standard deviation, 14 if within 2, and 28 if more than 2

4) Withdrawal strategy (whether the client will withdraw a proportion of his balance or of his weekly withdrawal rate, half of the clients being in each category), selected using a uniform random variable

5) Probability of withdrawing from an ATM other than his or her primary ATM, between 0% and 20%, generated using a Gaussian distribution with mean 10% and standard deviation 5%. Any random value selected outside of 0-20% is reselected

6) 0-4 arbitrarily placed seasonal brief migrations where 10% of the population has one or more migrations. Presence of a migration is decided by a uniform random variable, and the number of migrations is generated in an analogous manner to 3)

and 7) multihome (second primary) ATM, where 10% of the population has a secondary home as described above. Both the presence of a second primary ATM and its location are selected using a uniform random variable.

For simplicity, we register deposits proportionate to the weekly savings rate each time a client visits an ATM for withdrawal. Deposits are processed every two weeks, and only deposits received at least one week prior to processing are applied to balances.

4.5 ATM Parameters

We use a parameter n to determine how many ATMs to use to saturate our network, i.e. one ATM per n users. When this figure causes the number of ATMs to exceed one per population center, we use the notion that a given account can only use one ATM within the center and any others are treated as nonexistent, to obviate complexity of balance distribution among the multiple ATMs while maintaining semantic equivalence. We also use a ratio p to specify the proportion of the sum of the total available balances from a given ATM that the ATM should have on hand in cash, which affects cash restocking. Finally we specify how many days the server waits between balance redistributions, t_b, which affects message traffic volume as well as funds availability following a deposit.

To limit the number of variables in the simulation, we chose to fix the frequency of cash refills at once every two weeks.

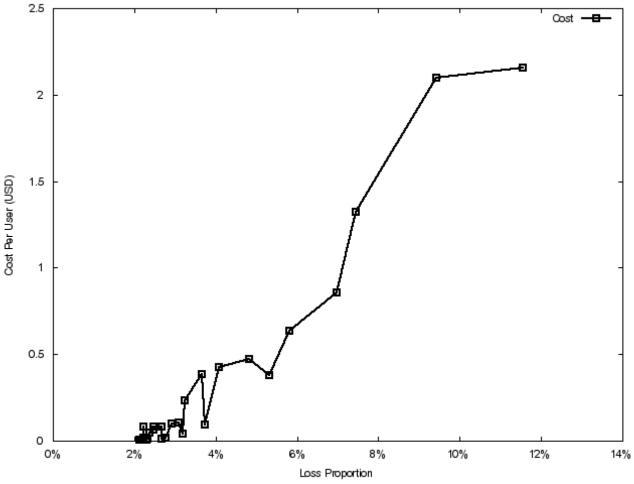

Fig. 6. Loss versus Cost per User in US Dollars, averaged across all runs of the simulator.

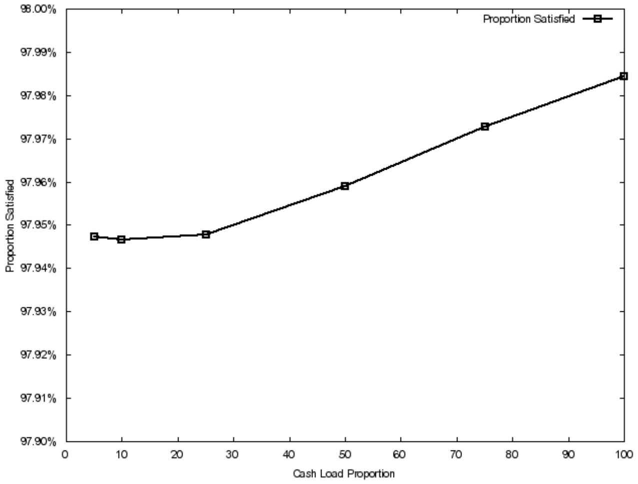

Fig. 8. Cash load proportion versus successful transaction proportion, averaged across all runs of the simulator.

5 RESULTS

In running our simulations, we are interested in seeing the effect of our various parameters on two primary results: cost per user per year, which indicates the marginal cost of adding more users to the network, and withdrawal request satisfaction rate.

5.1 *Loss Rates*

Simulation shows that the system is affected not specifically by outage frequency or outage burst length, but by overall downtime proportion. As shown in Figure 6, as loss increases from 1% to 12%, cost increases linearly as the number of messages which need to be resent increases. It is an obvious conclusion that messaging cost would increase as loss rate increased, but it is encouraging that the cost increases linearly rather than exponentially. The jaggedness of the curve is due to our using precomputed loss models across all trials to speed the simulator up, which exacerbates certain lossy periods as they occur in the same place in every run.

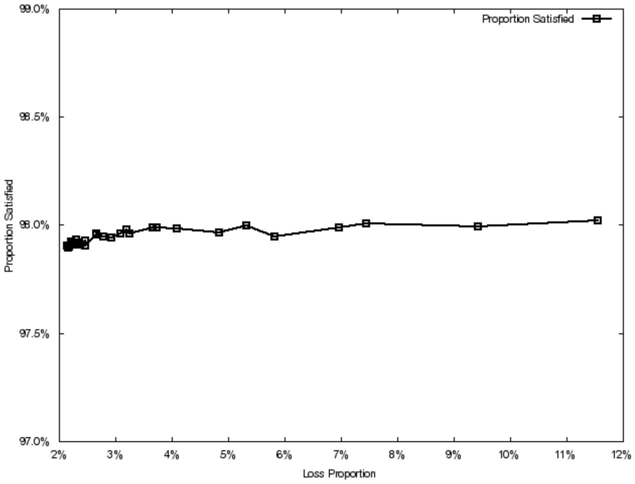

Fig. 7. Loss versus successful transaction proportion, averaged across all runs of the simulator.

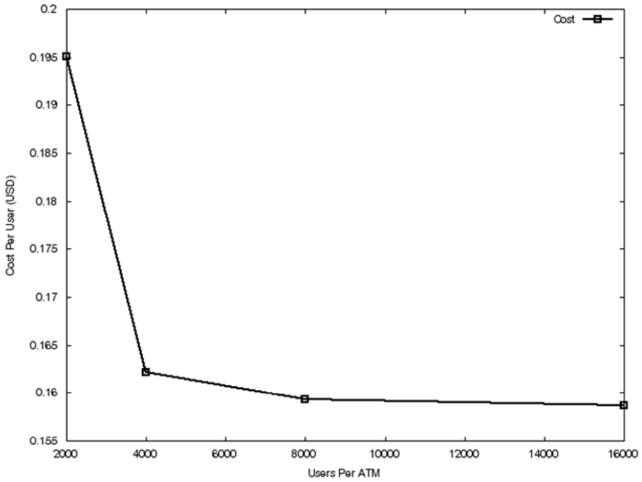

Fig. 9. Number of users per ATM (*n*) versus Cost per User in US Dollars averaged across all runs of the simulator.

Figure 7 shows that satisfaction is unaffected by the number of messages sent, which is an encouraging indicator that our protocol recovers well from loss.

5.2 *Cash Exposure*

Figure 8 illustrates the relationship between the cash load proportion and the proportion of requests satisfied. While the difference is absolutely minute, constituting less than 0.04% difference, the trend is clear that loading the ATMs with more cash provides slightly better request satisfaction levels. However, it is encouraging to see that even with at most 5% of the cash payable across all accounts in a given ATM loaded, satisfaction is approximately 98% of all requests under this model.

5.3 *ATM Density*

Varying the number of users served by a given ATM has a dramatic effect on the number of messages sent as each ATM sends messages after a given period of time

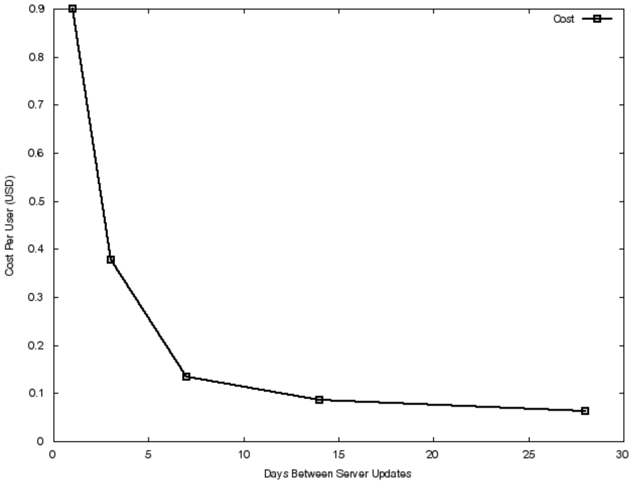

Fig. 10. Days between balance redistributions versus cost

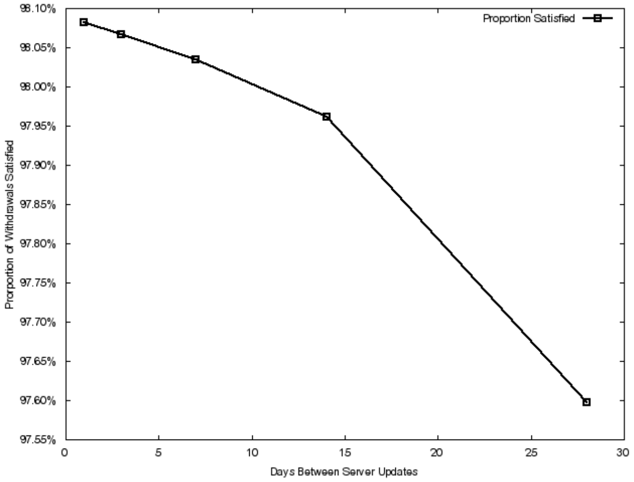

Fig. 11. Days between balance redistributions versus satisfaction rate

regardless of whether the message contains the maximum number of transactions or not, given that at least one transaction has occurred during that period of time and the server must send messages to each ATM during any balance redistributions.

As shown in Figure 9, cost per user drops off dramatically between 2000 users per ATM and 4000 users per ATM and levels off thereafter. This sharp elbow is explainable by the fact that many of the population centers used in the simulator have between 2000 and 4000 accounts, and therefore increasing the number of users served per ATM decreases the number of ATMs per location from 2 to 1. In our simulation, this change in density decreases the number of ATMs from 92 to 61. On the other end of the scale, few locations have more than 8,000 accounts, so there is little change by increasing users per ATM beyond that point.

5.4 Balance Distribution Frequency

As shown in Figure 10, cost per user falls off dramatically with less frequent updates at the cost of user satisfaction rate. The cost, fortunately, falls off much faster than the request satisfaction rate as seen in Figure 11. As each balance redistribution accounts for one complete set of account balances sent to each ATM in the system, this sharp increase in cost as updates become more frequent is compounded if the number of users per ATM is low enough to introduce an unnecessarily large number of ATMs into the system.

5.5 Example Optimization

To attempt to select a set of parameters to produce an optimal solution, we first need to observe that cost and satisfaction rate are opposed to each other, as are cash load and satisfaction rate. An implementing institution would have to determine what mixture of cost, risk, and customer satisfaction it would be comfortable with, but one example set of parameters is

$$p = 0.05$$
$$n = 16{,}000$$
$$t_b = 7$$

which yields over a 98% withdrawal success rate with a messaging cost of between USD 0.10 and 0.15 per user per year even accounting for nearly a 10% message loss rate.

6 DISCUSSION AND FUTURE WORK

The system as described has several practical weaknesses. An important limitation is the occasional need of accountholders to withdraw their entire savings. This can, however, be ameliorated by augmenting the system with an on-demand functionality whereby a user could pay for the additional SMS messages required to withdraw his or her full balance immediately.

Other weaknesses of the analysis include the explicit exclusion of physical security considerations and legal restrictions, both of which are often of critical importance in the developing world. The former in particular is important with deposits, as they need to be conveyed to some physical bank branch for validation.

The details of transaction auditing and reconciliation are also neglected, but it is assumed that standard practices will be followed.

A baseline study on the field performance of SMS would allow for further improvements to this system. We would like to better understand outage patterns for SMSCs and cellular networks as a whole. Some national governments, such as Nigeria’s, require cellular network operators to file reports on statistics affecting quality of service such as bit error rate, end-to-end error rate, etc., knowledge of which would allow us to better simulate this aspect of the environment. Also, detailed knowledge about the location of cellular towers and their respective power levels would greatly aid in the accurate representation of the region.

The next step in research for this system in particular will be investigation into whether decentralization can reduce network traffic and increase robustness against network failures, particularly at the cellular network cell hosting the institution. As balance distribution is dependent on geographical location, it may prove to be wise to have nodes collude with their nearest neighbors rather than receiving updates only from a central authority.

Other potential extensions of the project are increases in granularity or changes in the strategy of balance redistribution and tracking of loss and corruption rates of received messages in order to better allocate message budget. A redistribution strategy which is more adaptive to the past patterns of the user is also likely to yield positive results in terms of customer satisfaction.

As crafted, the system could be trivially extended to allow for payouts of microcredit installment loans or salaries for rural workers.

7 CONCLUSION

This paper presents ATMosphere, a protocol designed to provide ATM services to the world's rural poor. The system as simulated is based on currently deployed technologies and is implementable today using low-cost hardware. The result is a solution which has low start-up and marginal costs, represents low relative risk to financial institutions, and provides sorely needed access to savings and capital markets for the marginalized.

8 ACKNOWLEDGEMENT

We would like to thank various personal friends living in rural Uganda for firsthand assistance with baseline testing of MTN's cellular network. We would also like to thank Jinyang Li for her comments.

This material is supported by the National Science Foundation under Grant No. CNS-0831934.

REFERENCES

[1] E. Brewer, "Technology Insights for Rural Connectivity," in *Wireless Communication and Development: A Global Perspective.* Marina Del Rey, CA, 2005.

[2] L. Elser, A. Hannig, and S. Wisniwski, "Comparative Analysis of Savings Mobilization Strategies," Eschborn, GTZ, 1999.

[3] W. Enck, P. Traynor, P. McDaniel, and T. La Porta, "Exploting open functionality in SMS-capable cellular networks," on *12th ACM Conference on Computer and Communications Security*. Alexandria, VA, ACM, 2005, pp. 393-404.

[4] A. Hannig, and S. Wisniwski, "Mobilizing Microsavings: The Experience of Seven Deposit-taking Institutions," Eschborn: GTZ, 1999.

[5] G. Ivatury and M. Pickens, "Mobile phone banking and low-income customers: Evidence from South Africa," Washington, DC, 2006

[6] A. Jhunjhunwala, A. Ramachandran, and A. Bandyopadhyay, "n-Logue: The Story of a Rural Service Provider," *The Journal of Community Informatics,1* (1), pp. 30-38.

[7] S. Lord, "Trouble at the Telco: When GSM Goes Bad," *Network Security, 2003*(1), pp. 10-12.

[8] C. Olston, "Approximate Replication," Doctoral Dissertation, Stanford University, Stanford, 2003.

[9] C. Olston and J. Widom, "Offering a Precision-Performance Tradeoff for Aggregation Queries over Replicated Data," in *26th International Conference on Very Large Data Bases (VLDB).* Cairo, Egypt, 2000, pp. 144-145.

[10] R. Patra, S, Nedevschi, S. Surana, A. Sheth, L. Subramanian, and E. Brewer, "WiLDNet: Design and Implementation of High Performance WiFi Based Long Distance Networks," in *4th USENIX Symposium on Networked Systems Design & Implementation*. Cambridge, MA, 2007.

[11] L. Payne, "Rebuilding Communities in a Refugee Settlement: A Casebook from Uganda." Oxford: Oxfam, 1998.

[12] G. Peersman, P. Griffiths, H. Spear, S. Cvetkovic, and C. Smythe, "A tutorial overview of the short message service within GSM." *Computing & Control Engineering Journal, 11* (2), pp. 78-89.

[13] S. Staschen, "Microsavings, Microcredit and Microinsurance: Financial Products of Small Farmer Co-operatives Ltd. In Nepal." Cologne: AEF University of Cologne, 2001.
G. Taylor and M. Bekabye, "Employment Creation and Labour-Based Technology in Roadworks: Ugandan Case Study, 1999." [Online] Available: http://www4.worldbank.org/afr/ssatp/Resources/HTML/rural_transport/knowledge_base/English/Module%202%5C2_4b%20Case%20Study%20-%20Labour%20Based%20Techs.pdf

[14] C. Tseng, J. Jiang, R. Lee, F. Lu, C. Ouyang, Y. Chen, and C. Chang, "Feasibility study on application of GSM-SMS technology to field data acquisition." *Computers and Electronics in Agriculture, 53* (1), pp. 45-59.

[15] S. Wisniwski, "Microsavings Compared to Other Sources of Funds." Eschborn: GTZ, 1999.

[16] H. Yu, and A. Vahdat, "Design and Evaluation of a Continuous Consistency Model for Replicated Services," in *Fourth Symposium on Operating Systems Design and Implementation.* San Diego, CA, 2000.

[17] H. Yu, and A. Vahdat, "Efficient Numerical Error Bounding for Replicated Network Services," in *26th International Conference on Very Large Data Bases.* Cairo, Egypt, 2000, pp. 144-145.

[18] M. Zeller, "Models of Rural Finance Institutions," in *Paving the Way Forward for Rural Finance: An International Conference on Best Practices*. Washington, DC, 2003.

[19] ETSI GSM 3.40: 'Digital cellular telecommunication system (Phase 2); technical realisation of the short message service point-to-point', European Telecommunications Standards Institute TC SMG, Version 4.13.0, May 1996.

[20] ETSI GSM 3.41: 'Digital cellular telecommunication system (Phase 2); technical realisation of the short message service cell broadcast (SMSCB)', European Telecommunications Standards Institute TC SMG, Version 5.2.0, May 1996.

[21] GSM Doc 28/85: 'Services and Facilities to be provided in the GSM System' rev2, June 1985.

[22] MTN Uganda Ltd (MTN-UGANDA) – Interactive Coverage Map, 2007. [Online] Available: http://www.gsmworld.com/cgi-bin/ni_map.pl?cc=ug&net=mt.

[23] MTN Uganda: A Bright Y'ello World, 2007. [Online] Available: http://www.mtn.co.ug.

[24] MTN Uganda: Tariff Guide : Mobile Solutions : Business, 2008. [Online] Available: http://www.mtn.co.ug/business/talktime_tariffs.htm.

[25] The National Cheque Standard of Uganda, 1999. Available: http://www.bou.or.ug/NPSS/Cheque%20Standard.pdf.

[26] Safaricom: M-PESA, 2007. [Online] Available: http://www.safaricom.co.ke/index.php?id=745

[27] Quality of Service (QoS) Indicators for Mobile Services, Unknown Date. Available: http://www.ncc.gov.ng/TRS/QoS_Benchmark.pdf.

[28] Thuraya Satellite Communications, 2008. [Online] Available: http://www.thuraya.com

[29] Uganda Bureau of Statistics, 2007. [Online] Available: http://www.ubos.org

[30] WiMAX Forum – Technology, 2007. [Online] Available: http://www.wimaxforum.org/technology