Progress Monitoring for the Field Operations of the 2007 Philippine Census of Population[1]

By

Valentino C. Abuan[2]

A. INTRODUCTION

The Philippines National Statistics Office (NSO) conducted its 2007 Census of Population (POPCEN) in August 2007. This is the twelfth census of population undertaken since the turn of the 20th century. Previous censuses of population were carried out in 1903, 1918, 1939, 1948, 1960, 1970, 1975, 1980, 1990 and 2000. With a projected population of 88 million to enumerate, the 2007 census round is a massive undertaking requiring 37,000 enumerators, 7,000 team supervisors and 2,000 census area supervisors spread throughout the country.

Such a wide-scale data collection effort involving so many interviewers and field supervisors within a 30-day enumeration period requires a good tool to effectively monitor and track the status of enumeration work. Management action relevant to the census operations relies mainly on the progress information received from the field to be able to respond to the situation on the ground. Reallocation and mobilization of resources should be possible when and where needed.

This paper describes the Progress Monitoring System (PMS) that was implemented for the 2007 POPCEN with focus on the use of Short Message Service (SMS) technology, or more popularly known as text messaging, in the transmission of data. The paper also presents some of the benefits derived in the adoption of text messaging technology for this progress monitoring aspect of the census. At the same time, the technical and operational problems encountered in the implementation of the system are further described.

B. REQUIREMENTS FOR A PROGRESS MONITORING SYSTEM

A Progress Monitoring System was installed for the 2007 POPCEN mainly to track the status of enumeration throughout the country. This system is meant to facilitate

---

1 Country paper prepared for the Expert Group Meeting on Effective Use of IT in Population Censuses held in Bangkok, Thailand on 10-12 December 2007.

2 Director of Information Resources Department, National Statistics Office, Philippines.

the transmission of household and population counts for each enumeration area (EA) and the generation of summaries. These counts together with a status of enumeration for the EA serve as the indicators for tracking the progress of the field operations at the barangay, municipal/city, provincial, regional and national levels.

The basic input to the PMS comes from the CP Form 10, which is the EN's Accomplishment/Progress Monitoring Report. This form is used to record the total number of households, institutional living quarters (ILQ), male and female household members and ILQ residents. Entries in the CP Form 10 are derived from the CP Form 1, which are the listing sheets.

The outputs from the PMS are in the form of reports, which are accessible to authorized NSO officials as printouts, webpages and as files suitable for further processing. When queried by an authorized cellphone, regional and national-level summary counts are also sent via text messages. Among the reports generated by the PMS are:

PM Report 1 - Daily Transaction Report. This report shows the PM text messages accepted at the Central Office.

PM Report 2 – EA Coverage Report. This report details the population and household covered in an EA per day by an enumerator (EN). This will only be generated after completing enumeration of an EA.

PM Report 3 – Progress of Enumeration Report. This report shows the progress of enumeration in each geographic area selected by the user.

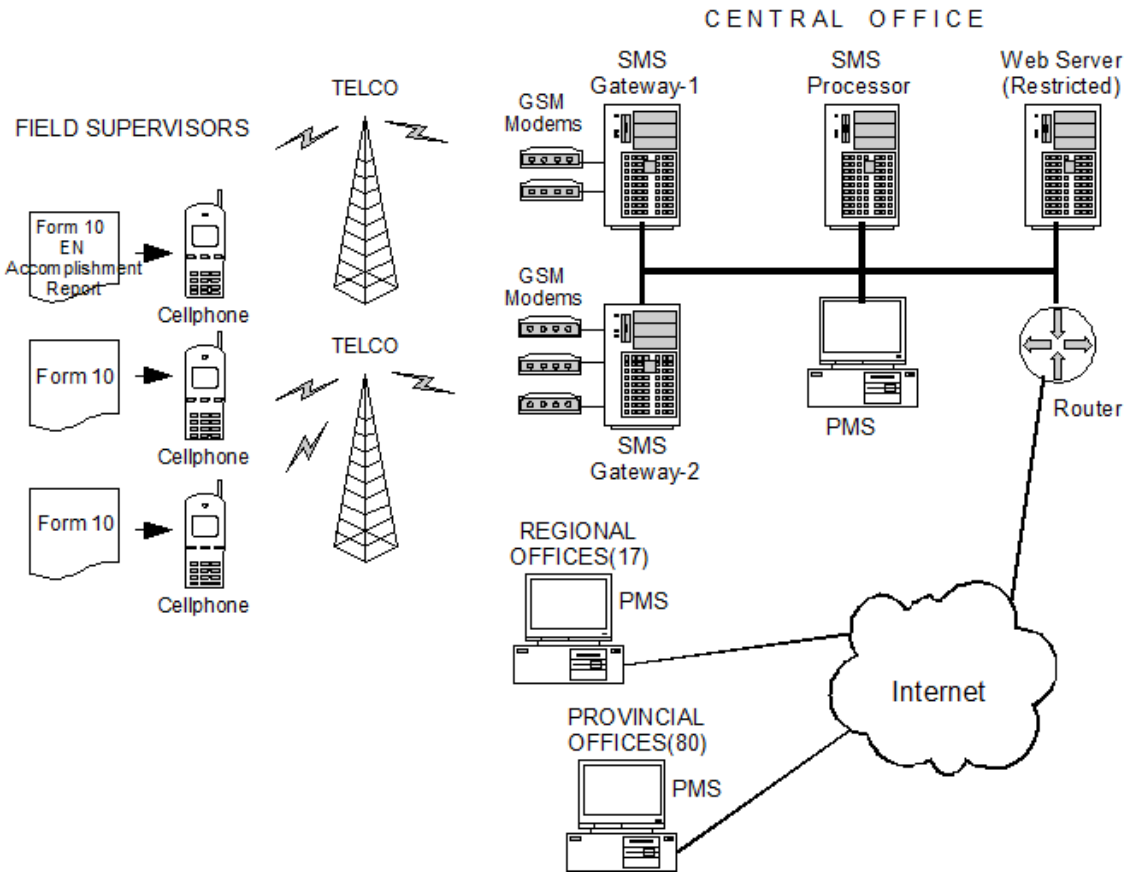
PM Report 4 – Data Evaluation Report. Serves to provide an initial evaluation of the reported counts based on selected derived indicators such as growth rates, household size and sex ratio. This is made possible with the inclusion of reference data based on past censuses.

It may be noted that while the PMS is primarily developed to provide progress information to as close to real time as possible, some reports are also made available to support quick assessments on possible padding and undercounts at an early stage. The assessment can be done by comparing the current counts against past census data as well the projected values. This early evaluation and assessment can trigger additional visits to the enumeration area.

The secondary objective of the PMS of making possible an assessment is to enable the NSO to publish the PMS-summarized counts as preliminary results. Having undergone initial assessments, the counts can be useful as reference during the actual data processing of the questionnaires and corresponding evaluation of the final results. It is also worthy to note that this census round did not provide for the conduct of a post-enumeration survey due to its time proximity to the regular census being planned for 2010.

C. HARDWARE AND SOFTWARE SUPPORT FOR PMS

The overall infrastructure needed to support the PMS implementation is shown in the figure below. The Census Area Supervisors (CAS) assigned to the field need to be equipped with cellphones to enable them to transmit EA-level data through text messages. Due to the large requirement for census area supervisors, the regular field officers of the NSO had to be augmented with hired supervisors for the entire duration of the census operation. The cellphones were personal property of the supervisors but the NSO provided an amount for the prepaid load credits for the phone units.



The cellphones utilized the nationwide network of the various telecommunication carriers in the Philippines. At the Central Office, five GSM modems are available to receive the text messages coming from the field supervisors. The GSM modems are connected to Linux servers functioning as SMS Gateways. These gateways control the operation of the modems, monitor incoming text messages and make decisions as to which machine the text message will be forwarded to based on the keyword contained in the message. Note that each of the GSM modem has a SIM card with corresponding assigned cellphone number. Furthermore, the SMS Gateways and modems are not solely for the census but also used in other text-based services of the NSO.

Text messages received by the SMS Gateway machines and found to have PMS-related keywords are forwarded to the SMS Processor box. This box makes a detailed scrutiny of the text message and validates the sender cellphone number. It also checks validity of the geographic codes and other required data against its reference database. When the text message is deemed valid, the SMS Processor machine stores the message into a database, generates a unique SMS Receipt Number and sends out an acknowledgment message to the sender. Text messages found to have problems either in its contents or coming from non-registered cellphone units are also responded to with error messages and without a SMS Receipt Number. The text message containing the acknowledgment or error message is sent out again through the SMS Gateway. Thus, all text messages received from or sent to the field always pass through any of the two SMS Gateway machines.

The PMS workstation in the Central Office accesses the database records created out of the valid text messages to generate any of the reports. This workstation also contains additional reference data from past censuses that are necessary to make evaluation of the current data. It also generates files to be used by the provincial version of the PMS program. The generated files are reviewed at the provincial offices by checking against the census forms submitted by the enumerators.

The web server also accesses the database of stored text messages to create summary counts for each geographic level and makes these available as web pages. Access to these web pages is restricted and limited only to NSO offices and officials through the use of user-id and passwords. This server creates web pages based on the updated databases every two hours.

PMS workstations at the provincial and regional offices serve the same functions as that of the PMS station at the Central Office. However, its data scope is confined to its specific area of concern. The provincial offices, after comparing the messages sent

against the forms, may also update the data using their PMS programs. This allows the handling of non-transmitted data as well as the correction of miscoded EA records.

The key softwares used in the implementation of the PMS are summarized in the table below:

Machine	Functionality Requirement	Software Used	Remarks
SMS Gateway	Operating System (OS)	Linux (Fedora 6)	Open source
	SMS Handling/Routing	Kannel	- do -
SMS Processor	OS	Windows XP	
	Database	MySQL	Open source
	Web service	Apache	- do -
	Customized functions (SMS scrutiny and handling)	Visual Basic 6	
PMS Workstation	OS	Windows XP	
	Customized Functions	Visual Basic 6	
Web Server	OS	Linux (Fedora 6)	Open source
	Web server	Apache	- do -
	Database	MySQL	- do -
	Scripting	PHP	- do -

The previous plan for handling of the SMS messages was to develop a program that will manage and control the GSM modems. However, a search at the Internet pointed the NSO to an open source software called Kannel. Initial trials on Kannel revealed that the software satisfies the needed functionality and reliability. Adoption of Kannel reduced the development period for the PMS.

D. SMS-BASED DATA TRANSMISSION

One of the desired characteristics of the PMS is a reduced time lag between the data collection and the availability of progress information to census management. Past implementations in the Philippines of progress tracking systems (also referred as Quick Count systems) made use of Internet facilities to bring the data to the Central Office. While the actual transmission process in an Internet-based setup is fast, the old method required bringing the forms to the provincial office for encoding and transmission via email. The bringing of forms to the provincial office can not be made more frequent due to cost considerations. Thus, the time lag for the progress monitoring data to reach the

central office can be long. It is not surprising to get the initial progress information on the last week of the enumeration period, which usually lasts only for a month.

A key consideration in the current design of the PMS is the recognition of widespread use and availability of cellphones in the country. The deployment of many texting services in the Philippines provided a good demonstration of the feasibility of using text messaging. In a census application, the transmission of data for the progress monitoring has been considered a good candidate for text messaging. SMS technology is mature and the only challenge left is to configure said technology to suit the specific application of the NSO.

Keywords for PMS Text Messages

For proper handling and routing of a text message, the message sent by the field supervisors should follow a specific format. Such format generally consists of a pre-defined keyword followed by one or more values. A one or two-character identifier prefixes each value in the message.

The keywords used in the PMS are as follows:

PMREG – to register a cellphone
PM – to send EA-level data after the cellphone has been registered
PMCOR – to send corrections to a previously sent and acknowledged data
PMREM – to send a free-form message, query or feedback.

Other keywords were also formed but were restricted and used only for system administration. For instance, there is keyword to enable the system maintainer to query performance statistics of the text message processor. Another keyword is available to check on the status of the server.

Registration of Cellphone

To prevent unauthorized or unwanted data from getting into the PMS database, each cellphone used in the transmission of text data must be registered first. The one-time registration process requires sending a text message containing the PMREG keyword, the geographic codes of the specific area of assignment, and the Access Code. The Access Code is a unique 5-digit number pre-assigned to a municipality/city and is given to the Census Area Supervisor at the start of the enumeration. The CAS receives a confirmation of successful registration also through SMS. From there, the SMS Processor machine checks each message received to ensure that only the data sent by registered cellphone numbers are recognized.

Frequency of Sending EA Data

CAS sends the EA summaries direct to the Central Office through text messages. Each EA-level data include the geographic codes, EA code, household counts, private and institutional population counts, and status of enumeration of the EA (whether ongoing or completed). The CAS is expected to send such information twice a week for each EA until such time that the EA is completely enumerated.

Validity Checks of SMS Data

Data received through SMS undergo some validity and consistency checks. Validity check includes matching the geographic codes against a reference database. Consistency check can be as simple as the total must equal the sum of the details. Error messages are sent back while text messages with valid contents are acknowledged with a SMS Receipt Number. This Receipt Number is to be recorded on the Form 10 by the CAS to indicate that the EA data has been successfully transmitted. The receipt number must also be included in the text message each time the EA record is updated through SMS.

Statistics on Text Messages Received

Text messages received were archived into a database. The statistics presented below were derived from such database.

Description	Number
Unique cellphone numbers used	2,389
Messages Received (valid and invalid)	243,079
Valid messages for registration (Successful registration)	3,929
Invalid messages for registration	876
Unsuccessful due to invalid geocodes	613
Unsuccessful due to wrong access codes	263
Valid messages for EA data	197,317

E. BENEFITS

A number of beneficial effects were realized in the implementation of the SMS-based Progress Monitoring System. The promptness by which the first progress information becomes available considerably improved. Right on the very first day of enumeration, status information started to pour in at the Central Office giving management a good picture of the overall operations. Past censuses had similar information available, at the earliest, on the second week of the enumeration period. The early availability resulted in much eagerness in the overall progress of the POPCEN not only in the Central Office but also in the provincial and regional offices.

Another interesting effect is the removal of time constraints in the transmission of data. Transmission of progress information used to be done during the daytime in past censuses. The arrival time of text messages was monitored during this census round and it was noted that a substantial number of data submissions continue to arrive even at night each day. It can therefore be said that the PMS is a 24x7 operation.

Early detection of some of problems in the listing operation as well as their rectification is achieved. The text message response from the system provided immediate feedback on the occurrence of invalid codes and inconsistencies. This allowed the CAS to call the attention of the concerned enumerator while they were still in the field. Thus, the bulk of the validity checks had been shifted from an office-based review, as what was done in the past, to a review while still in the field.

Reporting of events by field supervisors and seeking of clarification are facilitated. Accidents and other situations in the area such as dog bites, flooding, backing out of a hired enumerator, etc. are quickly reported to the Central Office. The facility also gave opportunities for the field supervisors to make direct queries with Central Office for clarifications on exception cases as they relate to the concepts given in the manuals or during the training.

The sending of data through text messages generates much enthusiasm among the field supervisors. Given the popularity of text messaging in the country, participants were eager to send immediately the data and to see the outcome. The enthusiasm and acceptance of the technology was seen as early as during the training stage. The large number of text messages received during the actual operation may give credence to the statement about their enthusiasm.

F. PROBLEMS AND ISSUES

The PMS is not perfect. There were difficulties encountered some of which were anticipated. For instance, it was expected that not all the CAS might have a cellphone unit. Likewise, some CAS may be slow to type data in their cellphones, which can take as long as five minutes to send a message. Among the measures taken to address these concerns was to allow the Team Supervisor (TS) with cellphone unit to also send data. The TS usually supervised five enumerators.

On the technical front, there were instances when the arrival of SMS messages was delayed. This problem was initially blamed as caused either by congestion in the SMS network of the telecommunication carrier or that the available GSM modems cannot handle the volume of incoming messages. Since there were only three GSM modems available during the start of PMS implementation, an additional two units were later hooked up to ease the traffic load.

While the solution of additional modems helped reduce the time delays in the incoming SMS messages, it was later determined that the root cause was the inability of the modem to recover from some errors inherent in the simultaneous receiving and sending of messages. Such modem errors required a power reset of the modem unit to recover. To ensure continued operation, the server log of the SMS gateway machine where the modems are hooked up had to be regularly reviewed. Nevertheless, a better solution was found midway into the enumeration period by simply restarting the Kannel SMS gateway software. The restart effectively reset the modems. The automatic restarting of the Kannel software at specific times everyday allowed a fully unattended server operation during the second half of the enumeration period.

In the early period of PMS operation, some CAS were unable to register their cellphone units. After repeated attempts to register with unsuccessful result, some degree of frustration may have gotten to the CAS. As it turned out, some of the CAS did not receive on time their final access codes and were using the earlier access code, which is applicable for training purposes only. The final access codes were emailed to the provincial offices and were supposed to be issued to the CAS before getting out for fieldwork. While this problem can be considered a lapse in the communication chain, the operation of the PMS in some areas was still affected.

G. CONCLUDING REMARKS

The Philippines NSO has successfully demonstrated the use of text messaging technology in monitoring the progress of enumeration in the 2007 POPCEN. As described earlier, there were some significant benefits realized from the use of the technology. Encouraged by these benefits and taking note of the lessons learned in the application of this specific technology, it is likely that the use of SMS would be repeated for the upcoming 2010 census round.

Some open source softwares were used in the current PMS. To a certain degree, the open source software category is often, or simply, neglected by many organizations in favor of well-known proprietary or commercial softwares. The Philippine experience shows that open source alternatives present a viable option. It is however important to emphasize that adequate research and trials be done to ensure that the identified open source solution is sound and robust enough.

Costs of software licenses continue to be a concern especially in census processing where many workstations would be required. While alternative platforms such as Linux prove stable and robust, it is unfortunate that not too many software packages for census or statistical processing are designed for such alternative platforms. It is hoped that the international funding organizations support the development of platform-independent statistical processing software that can also be used for census applications.

#