October 2012

E

منظمة الأغذية والزراعة للأمم المتحدة | 联合国粮食及农业组织 | Food and Agriculture Organization of the United Nations | Organisation des Nations Unies pour l'alimentation et l'agriculture | Продовольственная и сельскохозяйственная организация Объединенных Наций | Organización de las Naciones Unidas para la Alimentación y la Agricultura

Asia and Pacific Commission on Agricultural Statistics

Twenty-fourth Session

Da Lat, Viet Nam, 8-12 October 2012

Agenda Item 9

Use of Smart-phones for food and agricultural price collection and dissemination

Background:

The FAO Statistical Division is currently evaluating and experimenting new approaches and technologies to support national institutions in collecting data using smart phone-based surveys and/or automated email services. The main objective of this initiative is to verify the effectiveness and efficiency of alternative methods to gather field data, starting from food prices of marketplaces located in major towns and in remote areas. The collection of these data is usually carried out using paper-based methods, which are generally quite inefficient and sometimes unreliable. The inefficiency is due to the slow and cumbersome data transmission process from the field to HQs, which involves entering and checking the same data a number of times. More in detail, the data collection and transmission process is responsible for a) the time gap between the data collection and the analysis, and b) non-sampling errors in the data that propagate in the various steps to transmit data to HQ. The use of new technologies is expected to solve these inefficiencies and to streamline the data collection process. This new approach is currently tested on the collection of food market prices, but if successful, it would be extended to other types of socio-economic and environmental data.

An efficient collection of data on food market prices is essential to provide the most reliable evidence of ongoing or forecasted trends and volatility and to allow government to promptly intervene and prevent uncontrolled spread of price fluctuation. Timely and reliable data are key elements for developing short- and long-term strategies to face localized or widely spread actual or potential emergencies.

The new technologies are currently being tested in Haiti and are about to be implemented in India, Bangladesh, Nigeria, Kenya and Uganda. In the context of this pilot study, two technologies are being evaluated:

- Use of smartphones to transfer field survey data to the FENIX database
- Use of an automated email system to transfer field survey data to the FENIX database

As mentioned above, the objective is to reduce paper-based surveys to the minimum and move to a more secure and streamlined data collection system. The two methods are complementary and may apply in different conditions. However, both of them contribute to:

1. Reducing the time and cost of inputting data collected from paper-based surveys into a computer database;
2. Reducing the probability of errors from the data input process and therefore reducing the need for data cleaning and management;
3. Potentially increasing data security and confidentiality;
4. Creating a template/platform that could be easily adapted to other projects;
5. Increasing efficiency to make collected data available by eliminating transporting or converting intermediate physical (e.g. on paper) or digital (e.g. Excel) forms;
6. Enabling real-time monitoring of key field information, also in not-easily accessible areas

With the growth in availability and decline in cost of smartphones, these become ideal technologies to improve the accuracy and reliability of food price data collected directly from the markets and to reduce the time needed for providing data to those who need them. Similarly, automated email services may provide inexpensive and easy-to-use tools for collecting data in the field and loading them to the database.

System Architecture

The new food market price data collection system uses the FENIX platform as a basis. This is an internet-based and scalable system developed by FAO (initially for supporting food security and early warning) that serves data and functionalities to users and other web-applications through web services. Because of its modularity it can meet different user requirements and can be adapted to support global, regional and national level projects and initiatives.

The application is entirely based on cutting-edge open-source technology to avoid license constraints and to allow free distribution to FAO partners and beneficiaries.

A key objective of FENIX is to support harmonization an integration of national and international data by enabling application-to-application connections, providing data conversion and transformation tools and supporting data collection/visualization using a variety of software platforms, including mobile devices. FENX is capable of handling different types of data such as remote sensing, GIS layers, databases and texts. It is meant to work in a network environment for sharing information and for linking non-FENIX-based applications through appropriate data and metadata standards. FENIX provides different data upload and query tools, including an automated email system, directly managed by FENIX.

FENIX includes advanced data visualization tools and can provide client applications with advanced analytical functionalities through the "R statistical package", a powerful open source statistical library. FENIX-generated tables and charts may be viewed on smartphones, which might also be used by appropriate FENIX applications to collect data from the field through digital questionnaires.

FENIX promotes and implements international standards to manage and integrate data across national/international databases. Examples of these standards are: a) the FAOSTAT coding system, b) the Harmonized Commodity Classification (HS), c) the Global Administrative Unit Layers (GAUL), d) the OECD/DAC Creditor Reporting System (CRS) classifications for donors, sectors and geographic areas and various others. National standard and classification might also be included in, and handled by the application.

FENIX is used by other applications such as the new FAOSTAT data dissemination system, the statistical module of the Agricultural Market Information System (AMIS Statistics) sponsored by the G20, and the CountrySTAT application, which is already in use in 21 countries.

Data and metadata characteristics

Independently from the technology used, FENIX receives/sends both data and metadata information at every data exchange transaction. Metadata information sent via smartphone or email includes the minimum set of elements needed to handle the data correctly. Additional metadata information is linked to the data once data is fully uploaded.

The metadata information sent through the smartphone or the Excel file consists of three elements:

Metadata structure	Data Type	Source	Periodicity
Example	Price	Ministry of Agriculture	Daily data

Both data and metadata information use codes to store information in the database. Coding systems might be conformant to both national and international standards or can be user-specific. This flexibility is needed to ensure usability of the data collection/management system in different contexts.

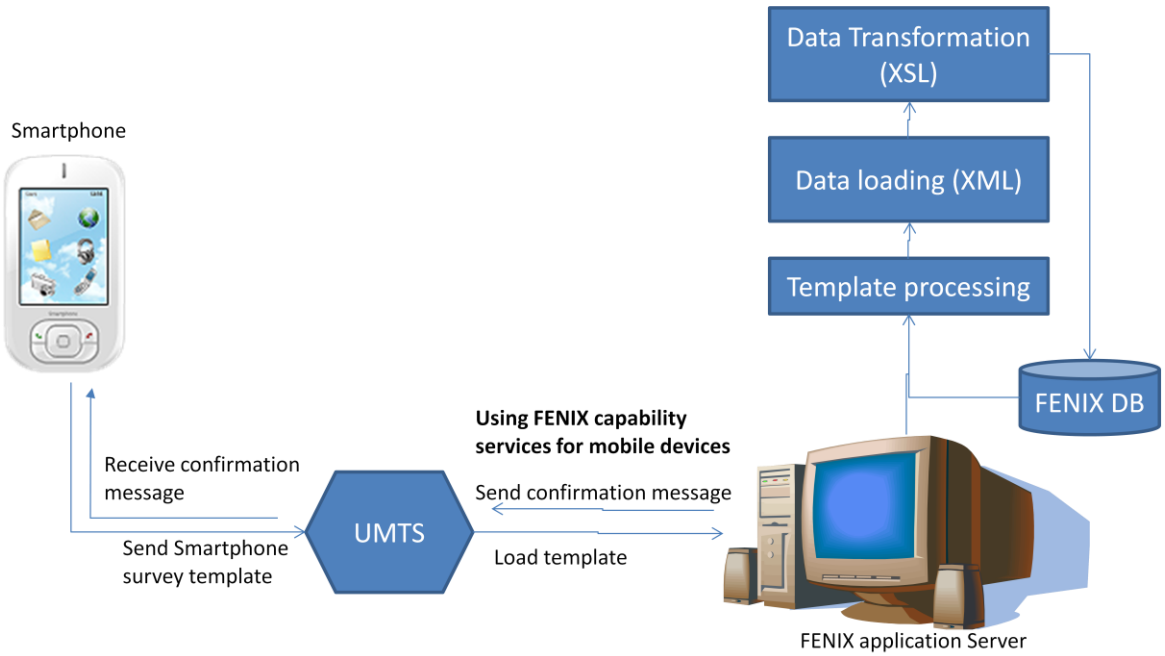
Technologies

Smartphone-based surveys

Templateswere created to provide the users with simple and intuitive user interfaces including textual data, numeric data, dates, single choice options, dropdown down lists to efficiently enter data and to minimize errors. Once data is sent by the user, the values entered in the template are accepted and managed at runtime. The templates are processed in 2 steps:

1. Replace parameter placeholders with the actual values of the parameters
2. Contact the FENIX database through the appropriate service and transfer data

The picture below provides an overview of the main steps for uploading data using smartphones.



Data collected in the field using the system describes above are stored in a temporary database, where data is checked and validated before it is sent to the production database and published.

Some of the characteristics of this system are:

- Multiple devices can be used to conduct a survey. There are no restrictions on the number of devices used to gather data.
- No Internet connection is required while conducting surveys and results are temporarily stored on the device and uploaded to the database when an internet connection is available.
- Raw survey data can be downloaded in the widely used CSV (comma separated value) file format. CSV files are most commonly viewed using Microsoft Excel.
- Smartphones can be configured to record the GPS location at the time the questionnaire is conducted. Longitude and latitude coordinates are included in survey results when downloaded in CSV file format or locations can be mapped using any mapping applications. This feature further enhances the detail included in survey results.

Although at the moment the system has been tested with the UMTS network only, it is possible to make it working with both the UMTS and the GSM networks. The latter requires a GSM gateway connected to the FENIX server. The use of the GSM network would considerably extend the geographic coverage, reaching even remote areas. The surveyors using smartphones would be transparent to the network they use to operate the system and to send data. The switch to the appropriate network would be entirely automatic.

The use of smartphones for data collection presents the opportunity to develop more responsive and extensive market information collection networks. In addition, the system can be used to broadcast up-to-date market information to farmers and other users in remote areas. The system by-passes several

steps needed to transfer data from the field to HQs and to provide opportunities to move the analysis process closer to the data generation process.

Automated email data collection

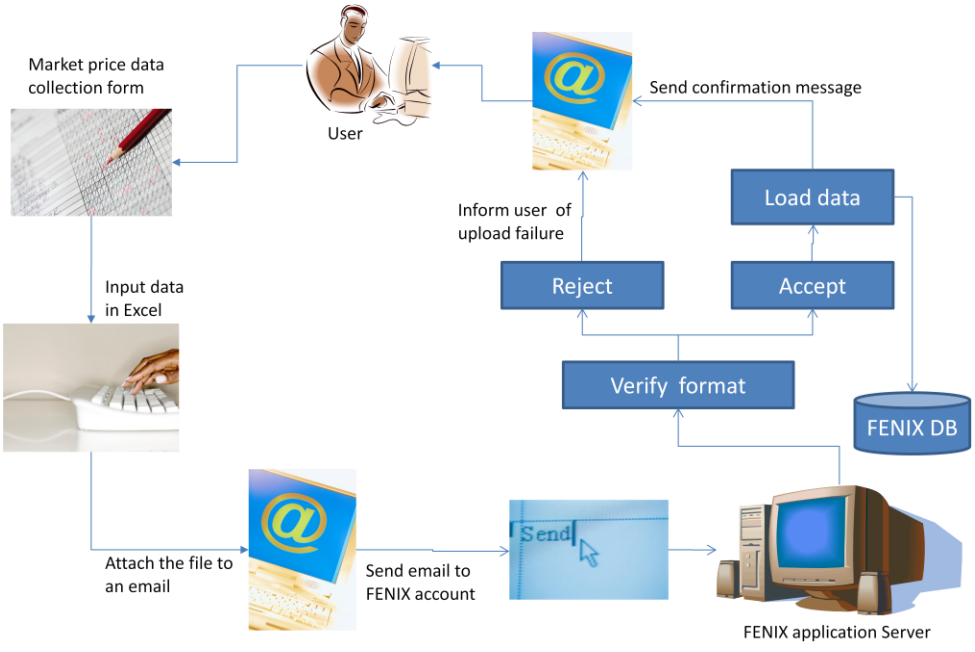
FENIX has implemented an automated email system that allows users to upload data to the database using email attachments. Accordingly, food market price data collected using the traditional paper forms are inputted in Excel files first and then sent to the application using the FENIX email account.

This method has the advantages to be inexpensive and to require low level of training. Data is collected on paper and subsequently (e.g. at the end of the day) entered in computer using Excel. These tasks can be carried out off-line without internet connection. Once the Excel file is ready, it can be transmitted to a FENIX node simply by sending an email to an account managed entirely by the FENIX application.

The FENIX email system performs the following operations:

1. Checks for new emails every 4 hours
2. If email is received:
    a. The system verifies the existence of attachments
    b. If attachments exist, it opens the attachment
    c. Verifies the compatibility of the format of the data and metadata with the FENIX data structure
    d. Once validated, it checks if the file to be uploaded is to be appended to an existing dataset or loaded as a new dataset
    e. Uploads the data and verifies if the operation has correctly been executed
    f. Sends a return email to the sender to communicate the result of the operation (successful or unsuccessful)

The picture below shows the workflow of the data collection process using automated email system.



The drawback of this system is that typing/formatting mistakes may occur while data collected in the field is entered in an Excel file before uploading to the FENIX database. Consequently, data verification against the original filed data collection form may be needed.

The automated email system may also be used to download data from the FENIX database.Users may send emails with keywords that are interpreted by the FENIX application and used as filters to query the database to extract data. Extracted data is then sent back to the user as email attachments.

Pilot testing in Haiti

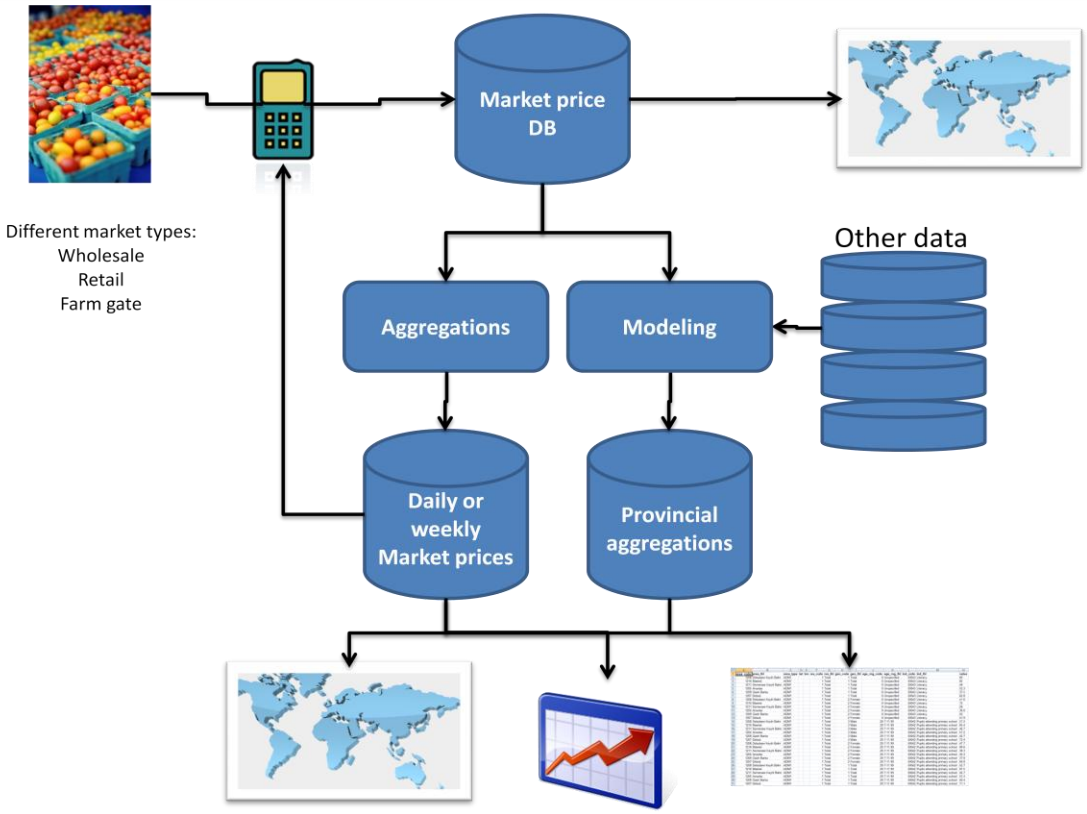
The FAO Statistical Division is collaborating with the Ministry of Agriculture in Haiti (MARNDR) to develop and test a new software application to support food market data collection in the field using smartphones. The objective of this activity is to verify efficiency and feasibility of this technology by implementing it for one (or two) market(s). The application allows to handle specific Haiti commodities and varieties and to enter prices in local currency or USD.

Two smartphones have been loaded with a preliminary version of the application and shipped to Haiti.

Data collected through smartphones is sent to a central database that validates the data, collects the geo-location of the smartphone, associates it with the closest market and displays data on real-time basis through a map.

Further development involves the statistical processing of the collected price data in order to derive daily, weekly and monthly statistics. This data will be made available online in a dedicated website.

The diagram below provides an overview of the data flow and the high-level processing involved.



Data is collected from markets using ad-hoc forms built in the smartphone application. The smartphone send data to the database that displays them on a map on real-time basis. The processing aspect involves a) aggregations to derive daily weekly and monthly prices and b) modeling to estimate market price data at provincial level by combining market price data with other relevant information from other data sources.

A successful test of this technology may encourage planning for a country-wise expansion of this data collection method.

Conclusions

The FAO Statistical Division is testing new ways to facilitate data collection in developing countries using alternative methods that can replace or strengthen current practices and technologies. Although only used to collect food market prices at the moment, the described data collection systems are flexible enough to be easily adapted forhandling any other type of data (e.g. land use, population, etc.).

If the requirement is to provide real-time data, the smartphone-based survey are more effective, although more expensive due to the cost of the equipment. The email system requires less training (data collectors continue to use the paper forms) and are considerably cheaper, but data has to be entered twice (in the field and in the office) with the risk of additional errors and is inevitably transmitted to the database with a certain delay.

In addition to the described methods, FENIX includes facilities to upload Excel/CSV files using an import routine or to directly edit the tables online.

In the evolution of this activity it is important to include the use of the GSM network which provides a much larger coverage, especially in the developing countries. This would require a GSM gateway connected to the FENIX server. The system then will be programmed to verify the availability of the UMTS network first and the GSM network as second option. The client application on the smartphone will be used indifferently with the two networks with no impact at user level.

By adding flexibility to the data collection system, the FAO Statistical Division will be in a better position to gather information from a much larger number of countries and areas. It will also be able to increase efficiently in information management through a series of database tools for handling, validating and publishing data. In order to increase data accessibility, web services are also being developed to access and upload data from countries that already implement database-driven applications and are willing to automate data exchange with FAO and among their partners. The FAO Statistical Division will therefore be able to advise on best options to establish/strengthen data exchange networks between national institutions and international/regional organizations aiming at harmonizing and optimizing inter-institutional data sharing.

Points for discussion

1) Have APCAS members any experience in this field?

2) Do APCAS members think that this system can streamline data collection?

3) Would APCAS members be interested in implementing it?