MEDINFO 2015: eHealth-enabled Health
I.N. Sarkar et al. (Eds.)

doi:10.3233/978-1-61499-564-7-396

Applying a Geospatial Visualization Based on USSD Messages to Real Time Identification of Epidemiological Risk Areas in Developing Countries: A Case of Study of Paraguay

Silvia Ochoa[a], Julia Talavera[a], Julio Paciello[a]

[a] *Polytechnic School, National University of Asuncion*

Abstract

The identification of epidemiological risk areas is one of the major problems in public health. Information management strategies are needed to facilitate prevention and control of disease in the affected areas. This paper presents a model to optimize geographical data collection of suspected or confirmed disease occurrences using the Unstructured Supplementary Service Data (USSD) mobile technology, considering its wide adoption even in developing countries such as Paraguay. A Geographic Information System (GIS) is proposed for visualizing potential epidemiological risk areas in real time, that aims to support decision making and to implement prevention or contingency programs for public health.

Keywords:

eHealth, mHealth, GSM, USSD, GIS.

Introduction

Health is one of the most critical areas regarding the use of information due to many existing regulations and primarily, public management.

In Latin America and the Caribbean, considerable inequity exists in the access to public health services . This is a result of a number of factors , including: lack of human resources, infrastructure, equipment, and medication. Further, a physical and cultural gap exists between the public health system and the population, related to low incomes and different ethnic backgrounds in developing countries. Thus, vulnerability is determined by income levels, geographic location, and ethnicity, causing an exclusion of millions of households in the region [1].

Information and Communication Technologies (ICT) increase the chance of improving the life quality in the community, as shown in previous works [2,3]. This paper focuses on applying ICT as a tool for public health surveillance. Thus, this initiative can be considered within the topic of *eHealth*. According to the *World Health Organization* (WHO) [4] eHealth is *the cost-effective and secure use of information and communications technologies in support of health and health-related fields, including healthcare services, health surveillance, health literature, and health education, knowledge, and research.*

The purpose of the present work is to optimize the geolocated data collection process for suspected or confirmed disease occurrences through the use of mobile devices to enable the identification of potential epidemiological risk areas in real time. Visualization of collected data is proposed through a *Geographic Information System* (GIS).

The proposed model can be implemented via mHealth as an application that allows users to report disease occurrence from their mobile devices through the exchange of *Unstructured Supplementary Service Data* (USSD) messages. USSD is a feature included in *Global System for Mobile Communications* (GSM) mobile technology (known as “2G”), that allows bidirectional transmission of information between the mobile device and an operator-defined application. With the collected information, this work proposes a geospatial visualization to identify epidemiological risk areas. Also, it can be used as a supporting tool for improving the actions of the organizations responsible for health.

We propose involvement of the community members in data collection, using the community as an actor in the process of epidemiological surveillance. The model of *Community Health Workers* [5] arises as part of the community auxiliary health team, where workers are selected and trained in order to work for their own communities.

In the next sections of this paper, a description of the problem is presented, including statistical information about the penetration of technologies in Paraguay. Also, the proposed model and considered scenarios are presented in detail. Finally, a comparative analysis between the proposed model and other alternatives is presented, along with conclusions of the work.

Background

In Paraguay, the *General Office of Health Surveillance* (DGVS, for its acronym in spanish) of the *Ministry of Health and Welfare* (MSPBS, for its acronym in spanish) is responsible for providing continuous information about monitored diseases in order to support decision making regarding prevention, control, and elimination of these diseases.

All the information related to mandatory surveillance of diseases in Paraguay is centralized by DGVS. Every week, the *Data Management Unit* (UGD, for its acronym in spanish) receives a notification form – which is printed (fax or mail) or in digital format (e-mail) – from the health regions. The forms are received mainly in printed format. Each health region has an Epidemiological Unit, which is responsible for collecting information from all reporting units.

The UGD unit performs a manual process of digitizing and unifying the collected information in order to obtain an electronic spreadsheet, which indicates for each record the reported disease and the specific health region during a epidemiological week, as defined by the MSPBS.

Finally, the information is sent to the *National Center of Communications* (CNE, for its acronym in spanish), where statistical analyses of the collected data are performed. An epidemiological newsletter is published weekly, showing the current status of the country. Typically, risk maps are used for visualizing the information.

According to statistical information from Paraguay in 2013 – with an estimated population of 6.700.000 – 95.5% owned a mobile device (97.1% urban areas, 93.4% rural areas). Additionally, 19.4% had landline communication subscription [6].

The General Office of Statistics and Census (DGEEC, for its acronym in spanish) announced that of all Paraguayan households, 26.6% have Internet-connected computers. However, it should be noted that Paraguayans have recently increased their access to the web using mobile devices (such as smartphones and tablets). For this reason it is estimated that the actual Internet penetration is approximately 30% [7].

From the mentioned statistics, the mobile technology – particularly GSM networks – represents a communication channel with a greater penetration in households than landlines or the Internet. This makes it an ideal tool to optimize the process of gathering information.

Based on these statistical values, we propose that an efficient way to penetrate more extensively in the territory of Paraguay is through GSM technology. Currently, internet and smartphone penetration remains insufficient in a developing country such as Paraguay, where the territory is mostly rural. Also, it should be considered that individuals typically excluded from public health services are those with low incomes or of different ethnicities.

Materials and Methods

Natural disasters, such as wildfires, floods, epidemics, and oil spills, can be modeled and presented through a GIS [8]. In public health, GISs have been used mainly for processing medical and epidemiological research that explores the magnitude and distribution of several health problems. In addition,, for analyzing, monitoring, and decision making related to this field.

The use of USSD communication protocol is considered, due to its ability to send information and instructions bidirectionally (between the mobile device and the applications that process the data). Thus, real time information can be provided by users, which will feed into a GIS, to allow the surveillance of epidemic risk areas. This will allow saving both economic and human effort, considering that, having this tool, the citizens and health workers would spend less time in the early identification of epidemiologic risk areas in order to support establishing policies, strengthen planning processes, and guiding actions to improve citizens' quality of life.

As shown in Figure 1, the information was collected in a centralized system, which may be accessible in real time to the actors involved in the information collection process. This will support the decision making at all levels and ensure that plans and policies are fully justified based on the evidence gathered. Thus, an improvement over the conventional non-automatic model was obtained, where the information originally reaches the levels required for decision making with a considerable delay.

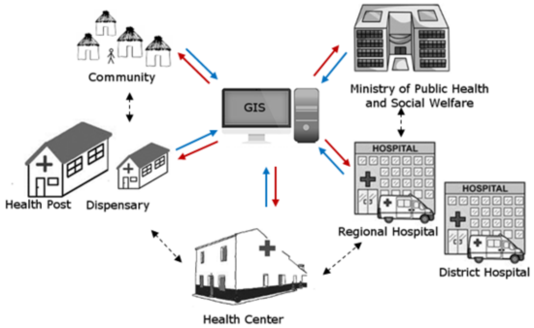


Figure 1 - Real time access to information

The posed problem in this work has been detached into two parts: data collection and visualization. A USSD module was designed for the first one, which provides greater coverage in terms of the number of citizens who will be able to access the service. For the second one, a GIS module was implemented, responsible for rendering (in real time) the information gathered for geographic visualization. The proposed model including the modules mentioned is shown in Figure 2.

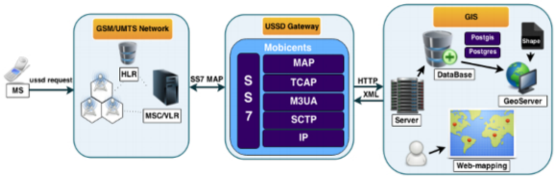


Figure 2 - Proposed model

The proposed model begins with the user or subscriber, who initiates a USSD dialogue by dialing a *Service Code* (SC) from the *Mobile Station* (MS). This message is encapsulated in a message *Mobile Application Part* (MAP) and is sent through the GSM network elements to the USSD Gateway (USSD module) which is responsible for receiving the request and identifying (based on the SC) which application has to be accessed in order to forward the request to it establishing a session. The application returns the initial menu to the USSD Gateway and this back to the user.

In this way, a USSD dialogue starts between the mobile device and the USSD application, where each entered option is logged in a georeferenced data base.

Finally, the data collected through the USSD module was processed and analyzed in the GIS module, which allows for visualizing real time information obtained from such data. In the following sections, both modules are explained in detail.

USSD Module

This module is responsible for data collection through a USSD application that uses RESTful web services to get an interactive menu of disease reports and registration.

The collected information was categorized according to the user who has made the notification. Every type of user will have a different notification accuracy level. This distinction has been deemed important because it is necessary to ensure the quality of the information. Inconsistent data or false positive notifications would not reflect reality, so the actions taken or applied on these data cannot be used to improve or eliminate the identified health problems. Thus, the notifications can be visualized and compared in real time and for each user type. Among the possible types of users, the proposal includes: a citizen, a community health worker, and/or an official of the Ministry of Public Health. The last two types of users are able to implement more rigorous and reliable checks and verification to reduce false positives.

The Mobicents [9] platform was used to simulate the USSD dialogue, which is an implementation of *Open Source Software* (OSS) for development of USSD applications based on the *Signalling System No. 7* (SS7) stack protocol. In Figure 3, the message flow of a typical MAP protocol is presented for transferring data between a network node, the Mobicents platform, and the USSD application.

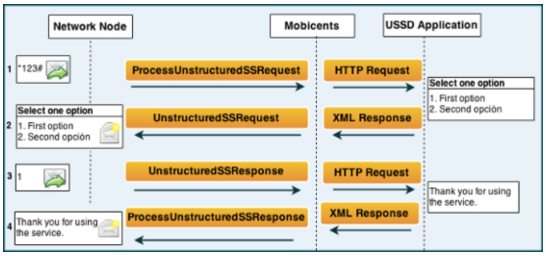


Figure 3 – USSD Dialog

The user dials the SC USSD code (e.g., *123#). A network node can represent the MSC, HLR, or VLR. The MAP protocol is used by the Mobicents platform to communicate with the network node and HTTP and XML protocols are used to communicate with the USSD application. The USSD application handles the business logic and sends the response back to the user in the same session.

GIS Module

The GIS Module is responsible for processing, analyzing, and categorizing data collected in the previous module, and allows the definition of thresholds on when to consider an area as being at risk for an epidemic. Finally, analysis of information is performed by a map viewer in real time.

This module aims to provide to the government, non-government organizations, and agencies that work on health services a tool to obtain georeferenced information of areas or regions affected by some kind of disease.

As shown in Figure 2, the GIS module consists of several components: a server application that is used by the USSD module to obtain interactive menus and persist the notifications of the users in the database, and a map server, which is responsible for analyzing the collected data and publishing the diseases layers with styles specific to each one of them. Finally, a web map visualization is included, where the available layers are passed to the map server and rendered to the final users.

Considered Scenarios

The scenarios considered in this work are differentiated by the locating mechanism used (that is, the mechanism that allows the mobile device's position to be determined). This information may be provided by the GSM network operator or by the user.

Automatic localization based on Cell-ID

There are several methods to determine the location of the MS. They can be different according to the technologies used to determine the location and the expected precision. In our scenario, the location method using the MS's Cell ID was used, considering its low cost and its inclusion in the GSM network, in order to determine the mobile device's location. The antenna identifier (the Cell ID) to which the device is connected is used to geographically locate the antenna, which is an approximation of the user's location.

For this scenario, the operator must provide geographic information of their *Base Transceiver Station* (BTS) antennas, so they can be identified by the visualization component of the GIS module. Since the information regarding the connected mobile device's antenna is handled by the telecommunication company, once a user's request is received, the USSD gateway simulates the consult to the HLR in order to determine the antenna to which the device is connected [10], and returns a MAP message with the Cell ID for that particular antenna.

The accuracy of this method depends on the cell size. We assumed 500 meters of accuracy within a city and about 10 kilometers in rural areas, representing an acceptable range considering the necessary resolution for identification of risk areas.

Manual localization

For this investigation two additional menu fields were included in the USSD application: the department (state) and city must be selected by the users in order to map to their corresponding latitude and longitude in the GIS, based on its own geographic data.

This scenario was considered as an alternative way to determine location when the system fails to obtain the proper information from the the telecommunication company. It is noteworthy that in these cases the accuracy of the model decreased significantly, and also required the user to know and select their department and city.

Results

The software evaluation process allows us to justify investing in the technology, based on the goals and needs of the consumer.

The development and selection of high quality software represents a critical point in the process, considering that the development and implementation of our method drives the success or failure of processes that are supported by these tools. This can be achieved by defining appropriate quality characteristics, taking into account the purpose and use of the software [11].

To verify the benefits and drawbacks of the conventional model currently used by the DGVS, and of the proposed model, a comparative study was conducted , considering each of the characteristics and subcharacteristics defined in Table 1.

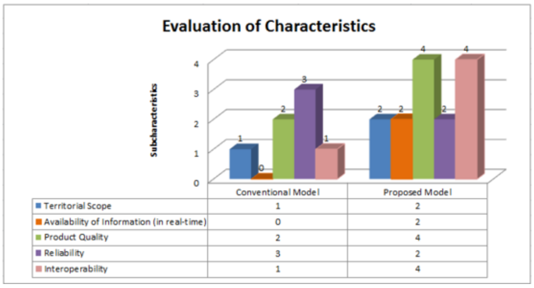
The comparison is based on a set of general characteristics and subcharacteristics which were defined in order to fix the problems that emerged during the present study and considering the characteristics and subcharacteristics of ISO 25010 that are comparable to the evaluation models.

Table 1– Quality characteristics

Characteristic	Subcharacteristic	Evaluation	Weight	Conventional Model	Proposed Model
Territorial Scope	Data Collection	Complies with this subcharacteristic if data collection covers at least 70% of the population as potential notifiers.	1	0	1
	Disemination Information	Complies with this subcharacteristic if the model allows access to information to support decision making.	1	1	1
Availability of Information	Data Collection	Complies with this subcharacteristic if data collection is done in real time.	1	0	1
	Disemination Information	Complies with this subcharacteristic if the information is disseminated in real time.	1	0	1
Product quality	Functional Pertinence	This subcharacteristic counts one point for each tool supported by the model.	3	2	3
	Documentation	Complies with this subcharacteristic if the model has adequate documentation to facilitate the use of the product to the end user.	1	0	1
	Accessibility	Complies with this subcharacteristic if the product can be used by illiterate or blind users.	1	0	0
Reliability	Objectivity	Complies with this subcharacteristic when the information provided is not biased towards a particular opinion.	1	1	1
	Topicality	Complies with this subcharacteristic if the displayed information is updated on a weekly basis, considering the epidemiological calendar.	1	1	1
	Credibility	Complies with this subcharacteristic if is possible to determine that the information provided is reliable and relevant with respect to the subject matter.	1	1	0
Interoperability	Exchange Formats	This subcharacteristic counts one point for each format (GIS, Web Services, or images) supported by the model.	3	1	3
	Open Data	Complies with this subcharacteristic if the data are made openly available.	1	0	1

Discussion

The score obtained by the models evaluated in each characteristic can be seen in Figure 4, where the scores for our proposed model far exceed those of the conventional model. With regard to *Territorial Scope* – and considering the process of data collection through the use of USSD technology – the proposed model outperforms the conventional model. The *Availability of the Information* (in real time) is a characteristic of vital importance, because our study is within the health sector, and whose subcharacteristics are dependendientes. In other words, since the information is available in real time, data should be collected in real time. Considering *Product Quality* (specifically when users interact with the collection or visualization tools) the conventional model provides the option to view historical information from the epidemiological newsletter from previous weeks, but has no tools to interact with the information provided. Of the three subcharacteristics evaluated with respect to functional pertinence, the conventional model complies with two, while the proposed model complies with three. Additionally, the conventional model provides little documentation, which hinders the understanding and interpretation of risk maps displayed in the epidemiological newsletter. This is opposite for the proposed model, which displays official documents in the same application. Finally, considering accessibility, both the proposed model and the conventional model cannot be used by people with certain physical impairments, such as people who cannot read and/or write, or who are blind. A drawback of using the proposed model is its low score in *Reliability*. This is because the reporting user base is much broader than that of the conventional model. The *Interoperability* characteristic was not highly rated for the conventional model since it only allows the display of static information, which does not allow share and reuse of the information. Instead, the proposed model complies with all considered formats of information exchange, and also openly shares data.



	Conventional Model	Proposed Model
Territorial Scope	1	2
Availability of Information (in real-time)	0	2
Product Quality	2	4
Reliability	3	2
Interoperability	1	4

Figure 4 – Final Evaluation of Characteristics

Figure 5 shows the final scores of the models evaluated. The proposed model has a final score of 89.33%, the highest score, wheras the conventional model has a final score of 43%.

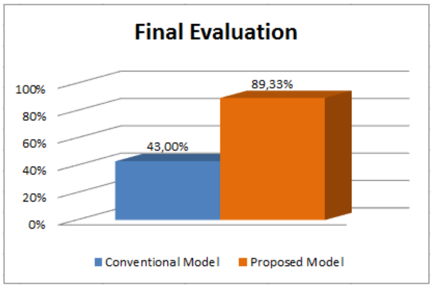


Figure 5 – Final Evaluation

Conclusion

This work highlighted the potential and benefits of a proposed model used for ICT-based epidemiological surveillance, which can easily be extended to other areas, towards the implementation of *Electronic Government* (eGov) and *Mobile Health* (mHealth) paradigms.

The emergence of new strategies for public health play an important role in Paraguay, for parties including primary health care (and family health units), community workers, and community empowerment towards personal health.

We have designed a USSD application that utilizes web services to provide an interactive disease menu on mobile devices connected to a GSM network, so that they can send notifications to a georeferenced database. A Web GIS module, which obtains spatial data from the database mentioned previously, allows the visualization of potential risk areas, and makes them available in real time to all stakeholders through web-mapping.

The proposed model is provided to governmental organizations, NGOs, and agencies working on health areas, and contains a tool for obtaining georeferenced information on areas or regions affected by some kind of disease while providing decision support for public health professionals. Note that the effectiveness of a GIS system generally depends on the nature of the problem and the selected software, as well as the reliability of the data and the ability of the users to understand their results.

Future work includes an implementation plan that will be drafted for applying the model to a Healthcare institution like the MSPBS in Paraguay. Additionally, a refined cost-comparison model is needed in order to better justify the benefits of the proposed tool.

References

[1] Comisión Económica para América Latina y el Caribe (CEPAL), “TIC y salud: promesas y desafíos para la inclusión social” presented at III Conferencia Ministerial sobre la Sociedad de la Información de América Latina y el Caribe, Lima (Perú), Santiago (Chile), 2010.

[2] Wouters B, Barjis J, Maponya G, Maritz J and Mashiri M. “Supporting home based health care in South African rural communities using USSD technology” presented at Americas Conference on Information Systems (AMCIS), 2009.

[3] Pan American Health Organization (PAHO), World Health Organization (WHO). “New Mobile Phone-based System facilitates Cholera Surveillance and Control in Haiti”. Washington, D.C., 2013. [Online]. Available: http://www.paho.org/ict4health/index.php

[4] Han JH, Sunderland N, Kendall E, Gudes O and Henniker G. “Professional Practice and Innovation: Chronic disease, geographic location and socioeconomic disadvantage as obstacles to equitable access to ehealth”. Health Information Managment *Journal*, vol. 39, no. 2, p.30, 2010.

[5] Lehmann U and Sanders D. “Community health workers: What do we know about them? The state of the evidence on programmes, activities, costs and impact on health outcomes of using community health workers”. School of Public Health, University of the Western Cape. Geneva. 2007.

[6] Direccion General de Estadística, Encuesta y Censo (DGEEC). “Entre los anos 2010 y 2013 se sumaron aproximadamente 780.000 usuarios de la Telefonía Móvil en Paraguay”. 2014. [Online]. Available: http://www.dgeec.gov.py/newspaper/news77.html

[7] Ultima Hora (UH), “El Acceso a Internet de los Hogares”, 2014. [Online]. Available: http://www.ultimahora.com/el-acceso-internet-loshogares-n795345.html

[8] Johnson R. “GIS Technology for Disasters and Emergency Management”. Environmental Systems Research Institute (ESRI). 2000.

[9] Mobicents. “USSD Gateway”. 2008. [Online]. Available: http://www.mobicents.org/

[10] Bertuzzi Leonelli RC and Salazar Fernandez JP. “Diseño e implementación de un servicio de localización y visualización de mapas utilizando J2ME para dispositivos moviles y herramientas de libre distribucion”. Síntesis Tecnológica, 2007.

[11] Siabato Willington,“Métricas aplicadas a los modelos de calidad: caso de uso en los SIG”, Congreso Internacional de Ingeniería Geomática y Topográfica. IX Congreso Nacional Top-Cart, 2008.

Address for correspondence

Silvia Ochoa; nochoa@pol.una.py
Julia Talavera; julia.talavera@pol.una.py
Julio Paciello; julio.paciello@pol.una.py