

Building a sustainable GIS through community input using free and open source technologies

G.N. Jayasena, D.D. Karunaratna
University of Colombo School of Computing,
No:35, Reid Avenue, Colombo 7,
Sri Lanka.

nalakajayasena@gmail.com, ddk@ucsc.cmb.ac.lk

Abstract

Building a comprehensive GIS for a country is an expensive and resource intensive process. Even if built, it is vulnerable to the changes of natural and socio-economic forces, rendering its information content at least partially obsolete and outdated not long after its deployment. One solution to this problem is to make more frequent updates- by deploying a larger force of data collection personal and by utilizing automated feature extraction methods from aerial and satellite images- which unfortunately, is not a financially viable option for a developing country such as Sri Lanka.

In this paper, we present an alternative with negligible financial cost where we engage, empower, and utilize the citizens themselves to play a major part in creating the GIS within a web-based, collaborative, wiki-like interaction model. The concerns of security and accountability are addressed by a role-based model of participation while accuracy of data is ensured against the reliability of the contributor as evaluated by their peers. All the components used are free and open source while standards are adhered to where available. MapGuide Open Source is used as the web mapping framework with MySQL as the spatial database.

1. Introduction

A Geographical Information System works with a representation of reality, oftentimes very crude and inaccurate. The reality represented through the GIS constantly changes. New roads are constructed while existing ones expand or fall to disuse. New buildings spring to life where previously bare lands existed; land and political boundaries change; floods and other disasters change the effective land-surface for the duration of that event; political, economical and cultural shifts change the way the land is utilized. In short, a Geographical Information System, no matter how comprehensive, is vulnerable to the changes of natural and socio-economic forces, rendering its information content at least partially obsolete and outdated with the passage of time.

One may try to overcome this problem by trying to make more frequent updates by utilizing a larger force of data collection personal. And by utilizing hi-tech automated feature extraction methods from aerial and satellite images. But invariably, all these methods fall short. The reason is straight-forward, the reality to be modeled by a comprehensive Geographical Information System is simply far more complex and volatile than can be handled by one or more institutions, no matter how large and powerful. Add to this the fact, that in developing countries like Sri Lanka, the available resources are merger at best, we have a situation where living with outdated GIS data is the norm. This is why we have to think outside the traditional GIS mindset; government or a powerful private agency collecting and distributing data and everybody else depending on them.

There is one agency that is more resourceful and powerful than any government, non-government or international agency. The agency is you. You, me, us; the common people. Voted “New York Times Person of the Year for 2006” [3], the power of the individual, boosted by the ever pervasive Internet, is increasingly being recognized for its true power. So it is towards the community that we have looked towards, as the ultimate resource to harness in building a sustainable GIS.

2. Background

In this section, we discuss the background in which our work originated: the key concepts on which this work is based and related work by other researchers.

2.1 Geographic Information Systems

A Geographic Information System (GIS) may be defined as a system for capturing, storing, analyzing and managing spatially referenced data and associated attributes.

There are primarily two types of spatially referenced data: raster data, and vector data. Raster data include scanned topographic maps as well as orthorectified (measurement errors removed) imagery such as satellite and aerial photographs. Although raster data are basically images, taking up a lot of space in storage, vector data tend to be more economical on storage since they are mathematical representations of the geographical features. Vector data can be captured either from ground surveys or by digitizing raster data. Although automated algorithms for doing this conversion exists, their accuracy unfortunately is not yet as reliable as manual digitizing.

Attribute data on any theme can be associated with the core spatial data of a GIS, allowing the analysis of such thematic data to use knowledge of the spatial dimension. Since most information directly or indirectly have a spatial component, the tools and techniques of GIS are becoming increasingly important in the government and industry.

2.2 Sustaining a GIS

A Geographic Information System is not static. It is impossible for it to be such, since a GIS is a partial representation of reality and that reality as we know only too well, is subjected to constant change. For this reason it is necessary for the data in a GIS to be periodically updated.

Updating GIS data can be done through ground resurveys by field teams; which while having the advantage of being highly accurate is prohibitively expensive. A more feasible approach is by using aerial photographs and satellite imagery to identify what has changed and to update the data by an appropriate combination of field visits and digitizing. There are several attempts at automating this change detection process such as described in [1] and [2].

2.3 Web GIS

Web GIS, the use of the World Wide Web as a platform for GIS, has become very popular in the last few years, even finally succeeding in the task of taking GIS to the masses with such huge success as Google Maps. Without doubt, the Web is the ideal platform for the delivery of maps. Additionally, with the advent of Web 2.0 and the rediscovery of Web as a general application platform [4] Web GIS is in a position to support most, if not all functionality offered by a conventional desktop GIS.

2.4 Collaborative Content Generation

Collaborative generation of content is a well tested concept recognized for its immense potential. One of the most popular examples is Wikipedia [5], an encyclopedia with community contributed content.

The major criticism of collaborative content generation with open access is the doubt that the resulting content could be accurate. However, an encouraging precedent is that Wikipedia has consistently defied its critics on such issues and emerged as arguably one of the best and most accurate encyclopedias, at least in some subject areas [6].

Using the concepts of collaborative content generation to web mapping is a recent development with relatively few precedents. One of the best known is WikiMapia [7]. It integrated satellite images from Google Maps and allows people to mark points and areas with attached descriptions. Another such effort is OpenStreetMap [8]

There are several challenges inherent in a collaborative web GIS, the primary one being that of quality assurance. With open distributed access and all the benefits it brings, the accompanying issue of concurrent editing of the same feature by multiple people has to be handled as well.

3. Praja GIS

The GIS Lab of the University of Colombo School of Computing has been working on the problem of effectively utilizing the potential of common people in building a sustainable GIS. This research was motivated by the need for the availability of low-cost, up-to-date spatial data for national development. We have named the resulting software Praja GIS, a web based collaborative system for creating and validating spatial feature data by the common people.

In both Sinhala and Sanskrit languages, the term ‘Praja’ means ‘the common people’, aptly describing our intention: a GIS for and by the common people.

3.1 Application Logic

When Praja GIS is initially deployed, it starts-off with the available base data for the region. This dataset is gradually expanded by contributors. All contributed data are maintained in separate layers.

The primary usage scenario of Praja GIS can be illustrated as follows:

A contributor performs a query; for example a query for ‘Vishaka Road’. A list of matching features is displayed. The contributor selects the one she is interested in. The map view displays the chosen ‘Vishaka Road’ and its immediate surroundings. The contributor realizes that some important features on either side of ‘Vishaka Road’ are missing from the displayed map. The contributor selects the ‘Edit’ option and the map view changes to the edit mode. Thereafter, the contributor uses the editing tools in the interface and creates new features and modifies existing features.

When the same location is queried for again by the same contributor or any other contributor, the map is displayed with both the original base data and the additional user-contributed data. If another contributor identifies that a particular user-contributed feature is in error, that second user is allowed the option to edit the feature and fix it. However, the previous contributor’s version is not lost but is saved in history.

3.2 Maturity Model for Contributors

Users of the GIS can be identified with two distinct roles; Contributors and Viewers. Contributors input and validate spatial data into Praja GIS whereas Viewers simply make use of the web mapping functionality the GIS has to offer.

The role of a contributor is further specialized by the access level to the system and by responsibility towards various region and theme specific datasets. Similarly the role of a viewer is further specialized by what datasets they have access to and the level of web mapping functionality available to them.

Contributors are evaluated by the user community as a whole based on the quality of their contribution. This will give rise to an interaction pattern where contributors will 'mature' within the system and 'earn' more power on decision making.

3.3 Interface Design

The Praja GIS in 'Map Edit Mode' is illustrated in fig. 1. The feature selected for editing is highlighted on the map. The information pane on the right hand side displays its metadata, edit history, and details. The edit tool bar contains tools for creating various types of features and for other relevant actions.



Figure 1: Interface for editing a spatial feature and associated attributes

The details tabbed page contains information such as: the name of the feature, the number of total edits for the feature together with a confidence level expressed as a percentage calculated based on the contributor(s) maturity and the data entry procedure used, and the associated attributes. The metadata tabbed page contains among others, information such as the original source, and method of entry (GPS, heads-up-digitizing, approximate drawing).

When creating new features the best case scenario is that the contributor is using a highly accurate GPS unit with which she has marked the boundaries of the feature and it is those coordinates that she is using for defining the feature. But the reality, especially in Sri Lanka,

is that not one among thousand contributors will have access to a GPS device. Therefore Praja GIS allows users to contribute in other more simple ways as well. This is an example for such a scenario:

The contributor selects a road segment, possibly as the segment between two road intersections. The user specifies the number of total land parcels abutting the left hand side of the road segment and the number abutting the right hand side. The land parcels are drawn along the road segment using suitable default dimensions. The user can move the vertices of the automatically created shapes to reflect the relative sizes of the actual land parcels. The user then selects each feature and specifies its name and optionally a description. Furthermore the user may, if she so chooses, specify additional metadata and attributes. The attributes relevant to the feature is displayed on the information pane based on the theme to which the selected feature belongs to.

Although not very accurate, this approach has the benefit of allowing the capture of spatial data by just about anybody without any special GIS knowledge. If or when more accurate data becomes available, these roughly accurate entries will be replaced by them.

3.4 Application Design

The application design as illustrated by the roles and responsibilities of its important classes and their public class members is as illustrated in fig. 2.

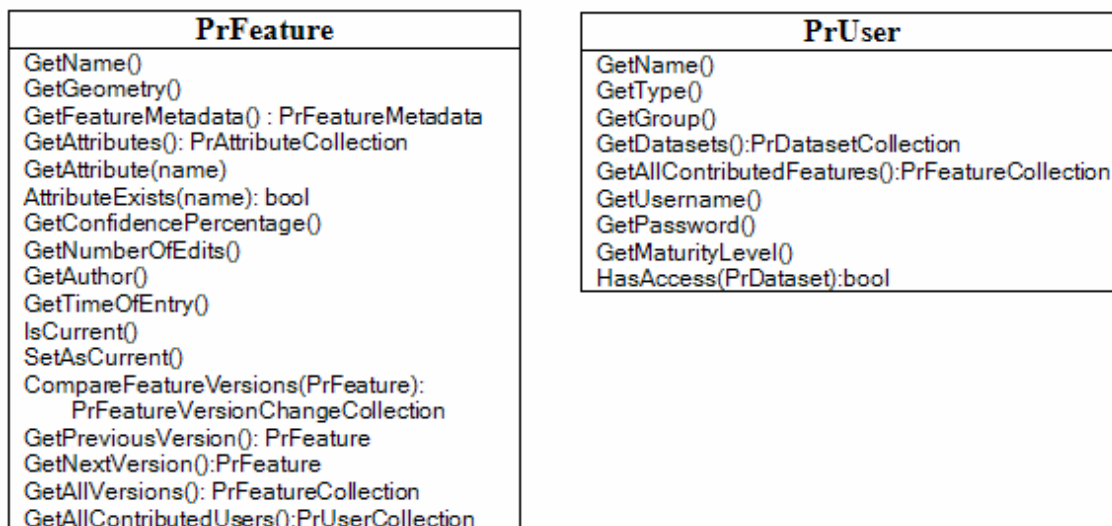


Figure 2: Abbreviated illustration of PrFeature and PrUser class members

The PrFeature class encapsulates a feature, both its spatial and non-spatial aspects. GetGeometry() returns the actual geometry as an OGC well-known binary (WKB) string. The PrFeatureMetadata class encapsulates the concept of metadata associated with a feature. The GetFeatureMetadata() operation returns a PrFeatureMetadata object representing all the metadata for a feature. Operations for accessing attributes and for supporting the wiki-like functionality is also provided.

4. Tools & Technologies

All tools and technologies used in Praja GIS are free and open source software. As the spatial database, which stores both attribute and feature data, MySQL with spatial extensions is used. Geo-processing functionality is obtained through the MapGuide Open Source geo-spatial framework and the application is developed using PHP while deployed in an Apache web server on a Linux environment.

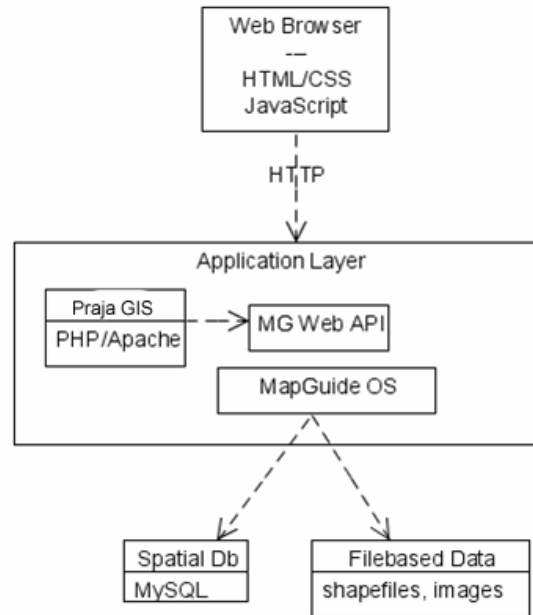


Figure 3: The technology stack

5. Conclusion

Building a comprehensive and sustainable GIS is a difficult and labor intensive process which by definition eludes the grasp of developing countries like Sri Lanka. Harnessing the community to power the GIS is an alternative. This paper has outlined the techniques and tools used by the GIS Research Group at University of Colombo School of Computing in their ongoing research effort in exploring this promising alternative and making it a viable option.

A major design goal of Praja GIS is security and reliability, of primary importance because eventually we intend this technology to be used at the local and national government level. With this in mind, a maturity model for contributors has been introduced to keep the accuracy of the generated data as high as possible.

One possible future development of Praja GIS is the facility to visually illustrate how a particular area changes with time using all the previous versions of features in that geographical area. This could be especially useful in visualizing an area undergoing rapid development or an area being developed after a natural disaster.

Although our focus is currently on providing open access to digital spatial data on Sri Lanka, it is quite feasible for the system to be ported for use in other countries because only open source technologies are used and our work itself is freely released to the public in source form.

References

- [1] A Gruen, O Kuebler, P Agouris, “Automatic Extraction of Man-Made Objects from Aerial and Space Images,” Springer 1995.
- [2] C. Eidenbenz, C. Käser, E. Baltsavias, “Atomi Automated Reconstruction Of Topographic Objects From Aerial Images Using Vectorized Map Information”, International Archives of Photogrammetry and Remote Sensing, 2000
- [3] Lev Grossman, “Time's Person of the Year: You” [online], Time Inc., Dec 2006, available at: <http://www.time.com/time/magazine/article/0,9171,1569514,00.html>
- [4] Wikipedia article on Ajax Programming [online], available at: [http://en.wikipedia.org/wiki/Ajax_\(programming\)](http://en.wikipedia.org/wiki/Ajax_(programming))
- [5] Wikipedia [online], available at: <http://en.wikipedia.org>
- [6] J. Giles, “Internet encyclopaedias go head to head” [online], Nature Magazine, Dec 2005, available at: <http://www.nature.com/news/2005/051212/full/438900a.html>
- [7] Wikimapia [online], available at: <http://wikimapia.org/>
- [8] Open StreetMap [online], available at: <http://www.openstreetmap.org/>