PINdex + 4: A SPATIAL IDENTIFICATION SYSTEM FOR POINTS OF INTEREST IN INDIAN CITIES

Santosh, NANDI[1], Patrick, NICHOLS[2], Madhavi Latha, NANDI[3]

[1]L1 Technologies Inc., 591 Camino De La Raina, 92108, San Diego, USA
santosh@l1inc.com
[2]L1 Technologies Inc., 591 Camino De La Raina, 92108, San Diego, USA
patrick.nichols@l1inc.com
[3]Department of Information Management, Xavier Institute of Management, Chandrasekharpur, 751013, Bhubaneswar, India
u506001@stu.ximb.ac.in

Abstract

Location addressing system in India is associated with drawbacks that limit the scope of geocoding the points of interest and subsequently carrying out spatial related analyses. In this paper, we present the details of a map mashup called *PINdex+4* spatial identification system that is proposed to overcome the shortcomings of the existing location addressing system in geocoding the points of interest and contribute to conducting business analyses more effectively. The proposed system leverages on an existing web mapping service to accurately locate a point of interest. Further, a spatial code called *PINdex+4* is generated for each point of interest by appending four digits to the corresponding postal index number (PIN) of the point of interest. The authors suggest that the *PINdex+4* code can serve as a spatial tag for the points of interest, which can be integrated with other business information and enable the organizations to conduct business analyses and generate business reports at different spatial levels including within the PIN code regions. The paper concludes with a brief discussion on the practical implications and limitations of the proposed system.

Keywords: geocoding, spatial Database, GIS Data, PIN code, map mashup

INTRODUCTION

Information has long been one of the important resources used by the organizations to deliver their services (Frenzel et al., 2004). Developments in information technology (IT) and the computer based information systems (CBIS) in the past three decades have changed the ways, in which organizations operate in the markets. Firms are increasingly relying on IT and CBIS to enhance their operational efficiencies and service effectiveness. Geographic information systems (GIS) are one such CBIS that has the potential to improve the efficiency of transaction systems and support organizations in making their business related decisions (Pick, 2004). Availability of the location data with offerings such as Google Earth and Microsoft Virtual Earth has opened up many possibilities of integrating location data with the business data in the business decision processes (Pick, 2007). The applications of GIS in business related analyses encompass asset management, route planning, service planning, customer profiling, and sales planning. Despite the advancements in GIS and related technologies, organizations in developing nations such as India, Pakistan, Phillipines and other nations are still struggling to reap the potential advantages of the GIS, when compared to their counterparts in developed nations. The problems hindering the potential use of GIS in the organizations include lack of consistent and accurate source data to generate the digital GIS data and lack of consistent location and demographic information (West et al., 1998). This paper focuses on the difficulties faced by the organizations in India in geocoding their points of interest to carry out various analyses such as customer profiling, asset management, and route planning. The paper further proposes a spatial identification system called *PINdex+4* spatial identification system that can enable the organizations to overcome the difficulties associated with geocoding the points of interest (POIs) and facilitate different business analyses.

GEOCODING THE POINTS OF INTEREST IN INDIA

Geocoding of the points of interest involves processing their textual addresses to add the positional co-ordinates that usually are the latitude and longitude of the location, to each address. These coordinates are then indexed to enable the addresses to be searched geographically (Bhaduri, 2003). There are several

methods to add the positional coordinates of the POIs. Positional coordinates of POIs could be derived physically using global positioning system (GPS) devices and integrating them with the address information. This process however, can be cumbersome and expensive, when there are huge numbers of POIs. Another method of geocoding the POIs is by plotting them manually, using the textual address information and the GIS data layers. Accuracy of the geocoded location however, is a problem with this method, especially when the user does not have a first-hand knowledge of the POI location.

Lack of urban planning and standardized location addressing system in the Indian cities make it difficult to incorporate standard logics for geocoding POIs. An address for a location in an Indian city typically consists of the building number or plot number followed by the nearest landmark, the area name, the city name, and the postal index number. Building numbers and plot numbers are allotted by the local Government entities that is, municipal corporations, following different methods. These numbers are not necessarily sequential and are not easily comprehensible for general usage. Landmarks and area names cannot be technically used as an input for geocoding the POIs. Landmark reference to a POI usually is associated with phrases such as “besides”, “near”, “behind” and so on, which are difficult to interpret. A similar difficulty in locating the addresses was pointed out by Razzak et al. (2011) in the context of locating the addresses of the road crash sites from medico-legal records in Pakistan. Concerning the area names, no definitive information is available on the area names and their corresponding boundaries, and hence serve as a loose reference to the location. Postal index number otherwise referred to as PIN code is the only available standardized and reliable resource for finding the geographical location of a POI till date. PIN code however, encompasses a larger geographical area. In the background of these difficulties associated with geocoding the POIs, many organizations are settling in for the business analyses at the PIN code level. The present paper proposes *PINdex+4* spatial identification system that allows users to graphically identify the address on a map while providing the addresses. The system further generates spatial codes called *PINdex+4* codes for the addresses of the POI, that can serve as spatial tags for the POIs in carrying out the business related analyses. Details of the proposed spatial identification system are discussed in the following sections of the paper.

PINDEX+4 SPATIAL IDENTIFICATION SYSTEM

System Architecture

The *PINdex+4* spatial identification system consists of a web application that is integrated with a spatial database, *PINdex+4* grid maps, and a web mapping service application such as Google maps, Bing maps, and Open street maps. In web development terminology, the proposed system can be called as a map mashup. A map mashup is a web application based on web 2.0 that uses and combines at least one map data source with added information to create a new map or a spatial database (Bitzer et al., 2009; Floyd et al., 2007; Gong, 2007). The general architecture of the mashup as shown in Figure 1 consists of three levels. The first level in the bottom contains data sources and services to be used in the mashup. The access to data sources and services is achieved through using the corresponding set of application program interfaces (API). In the case of *PINdex+4* spatial identification system, the first level consists of a set of map APIs from Google, Bing, or Open Street Maps, the *PINdex+4* grid database, and a spatial database. The spatial database is a database that supports storing and querying of spatial elements such as points, lines and polygons. The second level of the mashup holds the mashup applications, which are typically software routines. The mashup applications enable communication between the mashup components and bring the data sources and services together for a set of logics and functions the mashup provides. Last level on the top of the mashup architecture has the web user interfaces required to access the mashup sites (Gong, 2007).

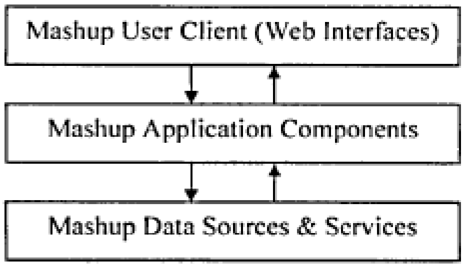

Fig. 1. General mashup architecture
(*Source: Gong, 2007)*

Functionality

While several functionalities can be integrated with the *PINdex+4* spatial identification system, the main functionality that potentially can help in overcoming the geocoding of the POIs in India is discussed in detail in this paper. The user interface essentially consists of a frame to enter the address information of the point of interest and other POI related information and map interface that displays the location of the POI. When the user enters the address information, a system software program retrieves and processes the state, city, and area information and sends a location search request to the mapping service application using the corresponding API. The map with an approximate location of the POI from the mapping service application is presented to the user. The user checks for the accuracy of the POI location, makes a correction if needed and approves the location. After the user approves the accuracy of the POI location, the *PINdex+4* code is generated by the system and stored in the spatial database along with the other information of the POI. The *PINdex+4* code serves as a spatial tag and the dataset can be used for different analyses at different spatial levels including regions within the PIN code.

Generation of PINdex+4 code

A PIN code in India consists of a six-digit number indicating the region, sub-region, revenue district, and the delivery post-office (India Post, 2012). For instance, in the PIN code 110006, the first digit that is, 1 refers to the northern zone, the second digit that is, 1 refers to the region of Delhi within the Northern zone, the third digit refers to the Delhi revenue district and the last three digits that is, 006 refer to the post office numbers in that PIN code boundary. The *PINdex+4* spatial identification system, appends four digits to the corresponding PIN codes of the POIs by recursively decomposing the rectangular grid generated from the extents of the PIN code boundary. The PIN code boundary grid is divided into four equal rectangular grids. The first digit following the PIN code of the POI corresponds to its location within these four grids. The digit is assigned in a clockwise direction starting from the upper-left grid, that is, the upper-left grid is numbered as 1, the upper-right grid is numbered as 2 followed by lower-right and lower-left as 3 and 4 respectively. This step is repeated for each of the constituent grids in which the POI is located, until four digits are achieved. An example for the procedure is shown in Figure 4, where the *PINDEX+4* code of 110006-2113 generated for a POI in Delhi. The logic behind the appending additional four digits to the PIN code can be compared to the point region quadtree data structures, which are generally used to store the vertices of polygon map data (Samet et al., 1985)

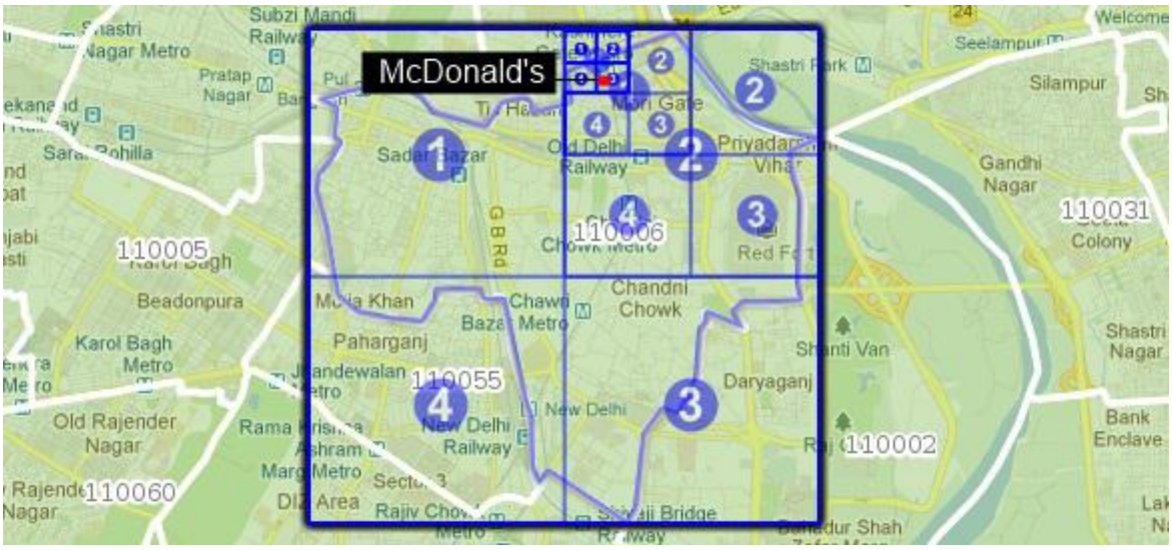

Fig. 2. Generation of PINdex+4 code

DESIGN AND DEPLOYMENT

The architecture of the proposed spatial identification system is shown in Figure 3. Presently, the system is in a conceptual stage. The authors intend to build a prototype of the proposed system and conduct a pilot case study with a direct broadcast satellite service provider in India. Initially the *PINdex+4* is planned to be deployed as a mashup for the use of enterprises catering to their specific business planning needs. In the later stages, when the system stabilizes, the *PINdex+4* code generation tool along with some analytic tools are planned to be hosted for general public use.

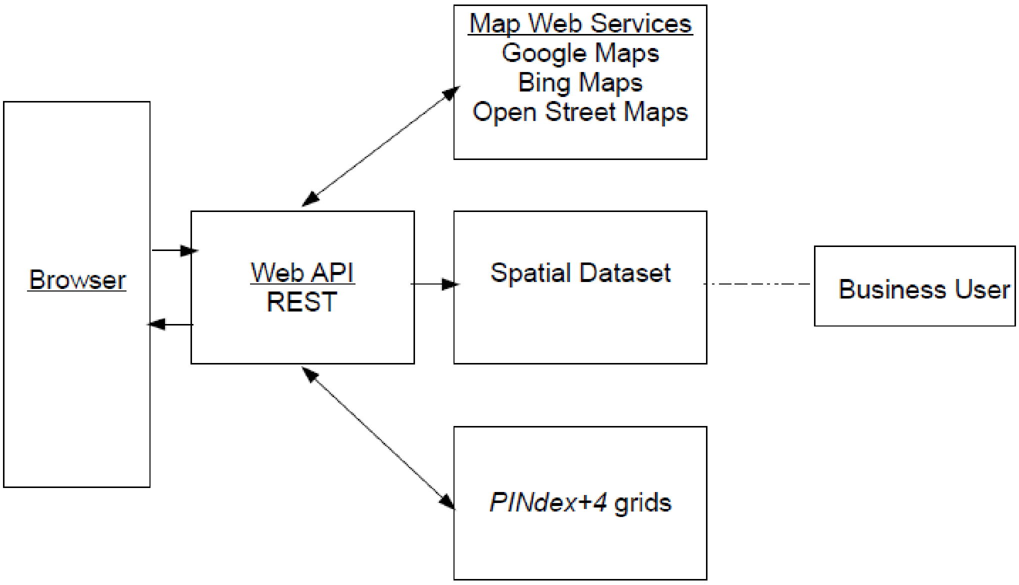

Fig. 3. *PINdex+4* spatial identification system architecture

PRACTICAL IMPLICATIONS

Firstly, the *PINdex+4* spatial identification system enables geocoding of the POIs and generation of a spatial database for the POIs with a simple additional step in the address information entry at the user level. These *PINdex+4* codes can simply be integrated with other enterprise applications and enable conducting business analyses and generating reports at different spatial levels independent of any GIS software platform. With

PINdex+4 codes, the level of business analysis can be enhanced to spatial levels within the PIN code regions. This feature is especially important to telecom and direct broadcasting satellite service providers to manage their assets and service outlets based on the customer profiles within the PIN code regions. Similarly, enterprises in the banking sector can plan their all time money (ATM) machine installations and design targeted promotions based on the customer profiles within the PIN code region. At a later stage when the system stabilizes, the general public users can communicate their address with the *PINdex+4* code, rather than the existing address system, which in most cases is lengthy.

LIMITATIONS

An important limitation of the *PINdex+4* spatial identification system is that the *PINdex+4* codes are generated from the PIN code boundaries. With changing political situations, such as addition of a new states and districts, PIN code boundaries change. For instance, new states have been added consistently until as recently as 2000 in India. These situations can have negative implications on the standardization of the *PINdex+4* codes. Also, the level of accuracy of locating a POI is dependent on the size of the PIN code boundary. Consequently, there is a possibility of having the same PINdex+code as a spatial identifier for multiple POIs. Therefore, it needs to be noted that the *PINdex+4* code can serve as an approximate location reference to a POI and not as its unique reference. Moreover, the accuracy of the POI location in the proposed system is subjective to the users' input and cannot be compared to the geocoding accuracy using the GPS survey method. Lastly, it is possible that different software providers can come up with different logics for geocoding that can potentially result in a conflict and hinder in creating a standardized spatial identification system. While this can be considered as a negative implication to the proposed system, it still could be a positive indication to addressing the geocoding problem that is persistently hindering the potential benefits that could be reaped from the GIS applications by the organizations.

REFERENCES

Bhaduri, A. (2003) User controlled privacy protection in location-based services, Electronic Theses and Dissertations, University of Maine, Paper 580.

Bitzer, S. and Schumann, M. (2009) Mashups: An approach to overcoming the business/IT gap in service-oriented architectures, AMCIS, San Francisco, CA, USA, AMCIS 2009 Proceedings, Paper 359, http://aisel.aisnet.org/amcis2009/359

Floyd, I. R., Jones, M. C., Rathi, D. and Twidale, M. B. (2007) Web mash-ups and patchwork prototyping: User-driven technological innovation with Web 2.0 and open source software, HICSS, Big Island, HI, USA Proceedings of HICSS-40.

Frenzel, C. W., and Frenzel, J. C. (2004). Management of Information Technology, Cengage India, New Delhi, India.

Gong, J. (2007) Mashing up web-based participatory GIS: Potential for future e-government services, Theses and dissertations, Ryerson University, Paper 260.

India Post. http://www.indiapost.gov.in/Pincode.aspx, October 8, 2012.

Pick, J. (2004) Geographic information systems, AMCIS, New York City, USA, AMCIS 2004 Proceedings, Paper 568, http://aisel.aisnet.org/amcis2004/568

Pick, J. (2007) Spatial technologies: An evolutionary framework and case study analysis of company strategies, AMCIS, Keystone, CO, USA, AMCIS 2007 Proceedings, Paper 362, http://aisel.aisnet.org/amcis2007/362

Razzak, J., Khan, U., and Jalal, S. (2011). Application of geographical information system (GIS) for mapping road traffic injuries using existing source of data in Karachi, Pakistan - A pilot study. Journal of the Pakistan Medical Association, 61, 640-643.

Samet, H. and Webber, R. E. (1985) Storing a collection of polygons using quadtrees, ACM Transactions on Graphics, 4, 182-222.

West, L. and Mennecke, B. (1998) Geographic information systems in developing countries: Issues in data collection, management, and use, AMCIS, Baltimore, AMCIS 1998 Proceedings, Paper 138, http://aisel.aisnet.org/amcis1998/138