

An Approach to Develop a Geographic Information Database Using Dot Spatial Open Source Platform and Google Search Engine

Maher Ibrahim Sameen*, Ali Adil Ali, Nibras Othman Abdul Wahid and Mohammed Ahmed Al Kubaisy

Department of Civil Engineering, Putra University, Selangor, Malaysia

Abstract

Currently, Internet has become a great database of knowledge and information for different disciplines. For instance, today, we can see the weather conditions for any place in any time online. All those data and information are important for geographic studies. However, transferring such information from Internet websites to GIS databases is a tedious work. This paper presents an automatic approach to transfer Internet information to a GIS database using open source platform. The case study of Malaysia database is used to be updated with temperature data from Google engine by the developed application. The designed application is developed by DotSpatial platform and Visual Basic.Net programming environment. It consists of four modules: (a) connection to Google and search for temperature data. (b) The process module to process the inputs and calculate the HTML parameters. (c) Temperature data extraction module to find out the temperature data. (d) Shape file update module to append the temperature data into GIS database. The results of testing the developed application showed good results of updating the loaded shape file of Malaysia states with the temperature information.

Keywords: DotSpatial; MapWin GIS; Open source GIS; HTML; Google; Vb.Net; GIS database

Introduction

Data is the core of Geographical Information System (GIS). Collecting and updating geographically referenced and descriptive data is necessary for GISs. Ignoring the updating process will obviously decrease the usefulness of GIS; therefore, GIS databases must be updated to make better use. There are different methods to collect and update GIS data, for example, by field survey, remote sensing, digitizing, or by analyzing old data to produce new data.

In terms of updating geographical GIS data, conventional methods can't satisfy the need of GIS development with some disadvantages such as low digital degree, high cost and long time consuming. The investigation results of a research by Min Han [1] proved that the data collection and database update of GIS based on GPS technology can make the local database update more convenient and flexible. In addition, many researchers are studying the GIS geographical data updating. Jan Ping Pan [2] summed up three kinds of GIS data update ways: raster to raster data, raster to vector data and vector to vector data. Then the author proved that these methods are effective and feasible. While Wang Mi studied on the design aspect of the efficient software system of GIS database update based on geophysical images.

Attributes are alphanumeric data that are associated with graphical features and stored in separate database files called attribute tables Hakan Maras [3]. On the other hand, for updating a GIS database attribute data, almost a survey notebook and high technology handheld GPS are used. But these methods are costly and time consuming. Recently, Internet provides large amount informative data which it can be used in GISs such as temperature, humidity, precipitations, populations, and economical data and so on. However, transferring those data into a GIS database is a tedious work.

In this paper, we will present a new approach for transferring such data based on HTML Object Library and Open Source GIS. The new approach automatically transfers HTML data into a GIS database by using a friendly user interface designed by Visual Basic.Net and DotSpatial open source platform.

This research takes Malaysia and temperature data as a case study. The designed application uses Google engine to get the temperature data by searching in the. Then, the data are automatically transferred to a loaded shape file in the application and saved. In this project, Malaysia states are used as a shape file which contains different states and information. The application will help as to add a field to this shape file and then fill the attributes with temperature values. The general construction of the proposed application consists of four modules: (a) connection to Google and search on temperature data. (b) The process module to process the inputs and calculate the HTML parameters. (c) Temperature data extraction modules to find out the temperature data. (d) Shape file update modules to append the temperature data into a GIS database.

Hyper Text Markup Language (HTML)

Hyper Text Markup Language (HTML) is a platform-independent method of identifying a document structure and references. Developed by scientist Tim Berners-Lee in 1990, HTML is the "hidden" code that helps us to communicate with others on the World Wide Web (WWW). It acts as a source code for web. It is a combination of variety of elements called tags, which are used to define type styles, headings, inserting specialized contents such as images, sounds, virtual reading words, Java applets and many more [4]. These tags tell the browser how to display the text or graphics in the document.

***Corresponding author:** Maher Ibrahim Sameen, Department of Civil Engineering, Putra University, Selangor, Malaysia, Tel: 0060183102 041; E-mail: maherrsgis@gmail.com

Received August 16, 2014; **Accepted** September 19, 2014; **Published** September 21, 2014

Citation: Sameen MI, Ali AA, Wahid NOA, Kubaisy MAA (2014) An Approach to Develop a Geographic Information Database Using Dot Spatial Open Source Platform and Google Search Engine. J Comput Sci Syst Biol 7: 217-220. doi:10.4172/jcsb.1000159

Copyright: © 2014 Sameen MI, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

The availability of information in large quantities on the Web makes it difficult for user selects resources about their information needs. The good link between the internet users and this information is Search engine. Search engine is kind of Information Retrieval (IR).

What are search engines?

Search Engine is a software program, help the Internet users to search and access to the data existing in the Web, like PDF, Video, Images, MS word and etc. Search Engine works according to the user query, which means search engine takes the keyword from the user and search in the Web, then search engine return the results to users in the form of hits. It uses the special software program called Web Crawler. It works to visit the Uniform Resource Locator (URLs) in the Web and index them. So when the user searches for something in the Web via search engine, just the user needs to put his/her keyword then search engine send the Web Crawler to collect the information from the database.

What is Google search engine?

Google search is a Web search engine owned by Google Inc., and is the most used search engine on the Web. Google receives several hundred million queries each day through its various services [5,6].

Google history

Google began in January 1996 as a research project by Larry Page, a Ph.D. student at Stanford he was soon joined by Sergey Brin, a fellow Stanford Ph.D. student they were in search for a dissertation theme, Page considered among other things exploring the mathematical properties of the World Wide Web, understanding its link and structure [5].

Dot Spatial open source platform

Dot Spatial is an open source GIS library written for .NET 4.0 frameworks. It helps the developer to build mapping applications utilizing visualize, query, analyze, browse and edit geo-spatial (raster and vector) and attribute data in Visual Studio environment. It also provides the developer with libraries to add various geo-processing tools in the application. Since its first release, Dot Spatial has been downloaded over 40,000 times and it currently receives approximately 200 downloads per day by users and developers for exploring the free and open source alternatives for GIS enabled custom software [7,8].

Dot Spatial is being used by a lot of open source GIS development including LiDAR applications and water resource management and hydrological modeling [9-11]. It is being developed by members of the MapWindow GIS open source team, and the broader OS Geo DotNet developer community. Currently, the 30 tools available in the Dot Spatial library are classified into 9 categories: Analysis, Terrain Analysis, Conversion, Spatial Reference, Vector Overlay, Raster Overlay, Interpolation, Statistics and Generalization.

Application architecture

The developed application architecture consists of five major modules used to update a GIS database using information from Google. First of all, the user enters the city name in a textbox offered by the software interface. Then, the software takes this string and put it in Google website (www.google.com) in order to search for the temperature data of the city. Next, the software extracts the needed information from the searched website and by using tags element of HTML object. When the application gets the needed data, it adds a

new field to a loaded shape file in the application layers and updates the attribute data by matching the city name strings. The Figure 1 below shows the major process of the application architecture.

Methodology

The design process of this software started from looking for a website provides temperature information of Malaysia states. After searching in Google, we found that if the city name with the word “weather” written in Google, we will get the temperature information of that city. Then, we identified the HTML elements by inspecting the page source in order to be used later for information extraction. In order to extract the needed information from a website, we need to tags name used in that website. Therefore, we extracted the tags name that contains temperature information and it was (“span”) with an index of (30).

Next stage of design aspect was the interface design using Visual Studio.Net 2012. As shown in Figure 2, the user interface of the software consists of basic tools of GIS such as add data, select, zoom, and information buttons. Also, it contains functions allow the user to interact with the software easily by major four functions. “Go to Google” function is developed to navigate to Google page. Next, the “Progress” function is used to store the city name with the word “weather” together and navigate the web browser to the combined string. In the “Calculate” function, the software extracts the temperature information from the website using tags name. Finally, the application uses “Update shape file” function to add a new field to the loaded shape file. After adding the new filed, it updates the attribute data of that field by the information that we got in the previous stage by comparing city names that the user entered and the city name in the shape file attribute table. Figure 3 below shows the attribute table of the used shape file and the States field.

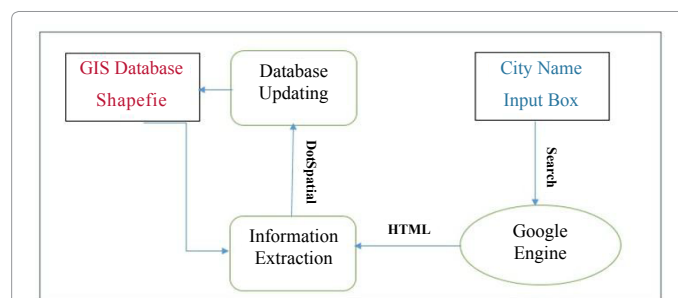


Figure 1: Application Architecture.

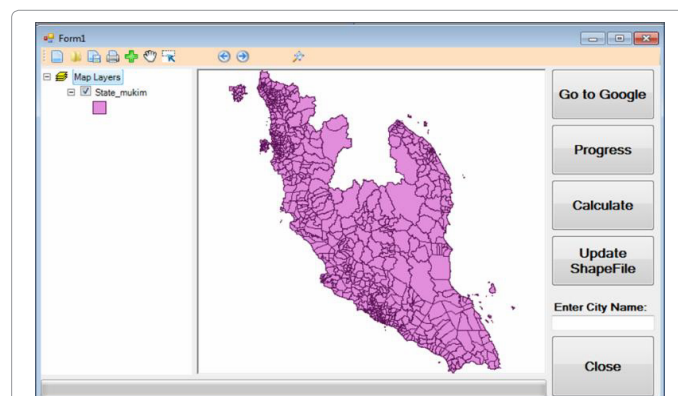


Figure 2: Application Interface with a Loaded Shapefile.

ID	MUKOM	STATE	DISTRICT	STATE_CODE	DISTR_CODE	SHAPE_Leng	SHAPE_Area	no	Temperature	FID
	Pengerang	Johor	Heading	01	005	163.2165	1400.961			930
	Pantea Tencr	Johor	Kota Tinggi	01	004	43.85795	116.9094			937
	Pantea Tencr	Johor	Kota Tinggi	01	004	376.188	9350.869			938
	Pantea Tencr	Johor	Kota Tinggi	01	004	806.9359	31794.92			939
	Pantea Tencr	Johor	Kota Tinggi	01	004	531.8851	17662.63			940
	Pantea Tencr	Johor	Kota Tinggi	01	004	270.5957	4985.588			941
	Pantea Tencr	Johor	Kota Tinggi	01	004	220.1694	5356.505			942
	Pengerang	Johor	Kota Tinggi	01	004	98.76251	726.8259			943
	Pengerang	Johor	Kota Tinggi	01	004	147.9325	1646.915			944
	Pantea Tencr	Johor	Kota Tinggi	01	004	875.7905	36794.46			945
	Puteh Kukup (P)	Johor	Pondean	01	007	11108.42	6300273			946
	Pantea Tencr	Johor	Kota Tinggi	01	004	122.3025	940.0765			947

Figure 3: Attribute Table.

Below code shows the complete code lines of the developed application by VB.Net programming language. First of all, it shows the necessary libraries that we used in the programming form DotSpatial Open Source libraries. Then, it shows some code lines deals with the HTML page and to set the attribute information into Google page. Finally, the four basic functions: “Go to Google”, “Progress”, Calculate” and “Update shape file” codes are shown respectively.

```
Imports SHDocVw, DotSpatial. Data, DotSpatial. Symbology
Public Class Form1
    Dim automate As Boolean
    Dim txt, btn As HtmlElement
    Dim temp (), t, tt () As String
    Private Sub Button1_Click (sender As Object, e
    As EventArgs) Handles gotoGoogle.Click automate = True
    WebBrowser1.Navigate (“http://www.google.com/”) End Sub
    Private Sub WebBrowser1_DocumentCompleted (sender As
    Object, e As WebBrowserDocumentCompletedEventArgs) Handles
    WebBrowser1.DocumentCompleted
        If automate = True Then automate = False
        Else Exit Sub
        txt = WebBrowser1.Document.GetElementById (“q”) txt.
        SetAttribute (“value”, TextBox1.Text & “ “ & “weather”)
        btn = WebBrowser1.Document.GetElementById (“btnG”) Do
        Loop Until WebBrowser1.ReadyState =
        tagREADYSTATE.READystatechange_COMPLETE ProgressBar1.
        Value = 100
    End Sub
    Private Sub Button3_Click (sender As Object, e As
    EventArgs) Handles progress. Click
    If ProgressBar1.Value = 100 Then btn.RaiseEvent (“onclick”)
    End If
    End Sub
    Private Sub Button4_Click (sender As Object, e As
    EventArgs) Handles calculate. Click
```

```
t = WebBrowser1.Document.GetElementsByTagName
(“span”)(30).InnerText
tt = Split(t, “”) End Sub
Private Sub Button2_Click (sender As Object, e As
    EventArgs) Handles update. Click, update. Click
    Dim fs As New FeatureSet
    Dim query As String
    Dim ff As FeatureLayer
    fs = Map1.Layers.Item(0).DataSet
    query = “[STATE] =” + “” & TextBox1.Text & “” Map1.
    GetFeatureLayers (0).SelectByAttribute (query)
    Dim dts As DataTable
    ff = Map1.GetFeatureLayers(0)
    dts = ff.Selection.ToFeatureSet.DataTable
    DataGridView1.DataSource = dts
    Dim i As Integer
    Dim str As New List (Of String)
    For Each row As DataGridViewRow In
    DataGridView1.Rows
    If Not row.IsNewRow Then
    str.Add (row. Cells (0).Value)
    End If
    Next
    Dim i1 As Integer
    MsgBox (DataGridView1.Rows (0).Cells (8).Value) For i = 0 To
    str.Count - 1
    i1 = Me.DataGridView1.Rows (i).Cells
    (8).Value
    fs.DataTable.Rows (i1).Item (“Temperature”) = tt (0) Next
    fs.Save () End Sub
    End Class
```

Results and Discussion

The application user interface is shown in Figure 4 below with the Malaysia shape file that loaded into the software. The designed interface indicates that the software easily can be used for GIS database updating due to the large icons and self-explanatory buttons which can be used by non-GIS specialists effectively. It also offers basic GIS toolbars to implement basic functions such as zoom, select, information and pan function.

Data used were Malaysia shape file contains districts, states and their identification codes as shown below in figures. The shape file is used in order to be updated with temperature values. The example below is applied for Johor state as shown in Figure 4.

The software is tested in 1-7-2014 with good internet service on a

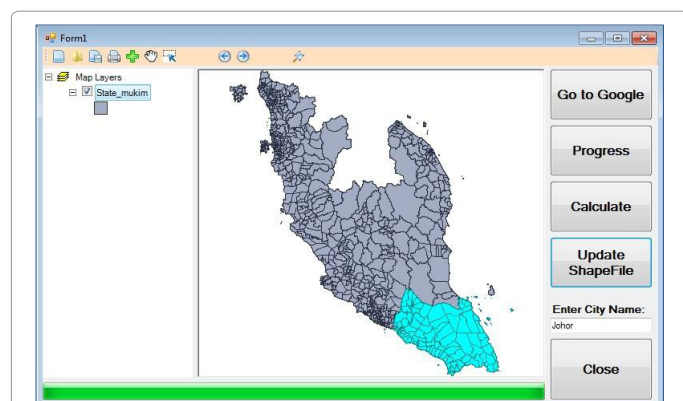


Figure 4: Selection by City Name.

OBJECTID	MUNIM	STATE	DISTRICT	STATE_CODE	DISTRICT_CODE	MUNIM_CODE	PCODE	SHAPE_Leng	SHAPE_Area	Temperature
296	Paritbangsa	Johor	Medan	01	005	05	CA0050	296	162891	33
299	Paritbangsa	Johor	Medan	01	005	05	CA0050	297	2478608	33
300	Paritbangsa	Johor	Medan	01	005	05	CA0050	298	2983293	33
301	Paritbangsa	Johor	Medan	01	005	05	CA0050	299	7914283	33
302	Paritbangsa	Johor	Medan	01	005	05	CA0050	300	1419163	33
303	Paritbangsa	Johor	Medan	01	005	05	CA0050	301	2544221	33
304	Paritbangsa	Johor	Medan	01	005	05	CA0050	302	3365811	33
305	Paritbangsa	Johor	Medan	01	005	05	CA0050	303	8779184	33
306	Paritbangsa	Johor	Medan	01	005	05	CA0050	304	1658432	33
307	Paritbangsa	Johor	Medan	01	005	05	CA0050	305	1462363	33
308	Paritbangsa	Johor	Medan	01	005	05	CA0050	306	1221285	33
309	Paritbangsa	Johor	Medan	01	005	05	CA0050	307	3871833	33
310	Paritbangsa	Johor	Medan	01	005	05	CA0050	308	2733790	33
311	Paritbangsa	Johor	Medan	01	005	05	CA0050	309	5285247	33
312	Paritbangsa	Johor	Medan	01	005	05	CA0050	310	6596856	33
313	Paritbangsa	Johor	Medan	01	005	05	CA0050	311	1254538	33

Figure 5: Updated Attribute Table.

personal computer with the properties: Core i5 CPU 2.5GHz, 64-bit operation system and 6GB RAM. The software worked properly and the shape file is updated with the temperature information as shown below in Figure 5. The new field is added named “Temperature” and loaded with the temperature information from Google engine on that time.

Conclusion

Collecting geographical and descriptive data are necessary for GIS. In this research we implement Google engine to update GIS attribute data. Malaysia is used as a case study with updating a database with temperature values. The developed application showed the ability on updating GIS databases with temperature information very easily and effectively with normal internet network connectivity and PC. Visual basic.Net and DotSpatial platforms gave the power to the software in terms of design and dealing with GIS files aspects. Developing GIS applications by Visual basic has made the user interface very friendly and efficient that can be used by non-specialists as well. However, the developed software has some limitations that we recommend to other researchers to investigate this approach further as following:

1. Develop a broad application that able to update GIS databases with different types of data such as humidity, population, and economic data.
2. Design user interfaces for non-specialists to be used in all organizations that they dream to become GIS users and help their organizations.
3. List important websites which offer significant information that is useful for GISs.

Acknowledgment

The authors are grateful to the Universiti Putra Malaysia, Malaysia, for the digital resources support. Also they thank Mr. Qayssar Mahmood for providing the lab facility in conducting this research work. And we would like to thank the Minister and the employees of Ministry of Higher Education and Scientific Research who have helped enrich my knowledge and for all their support and encouragement.

References

1. Han M, Tian X, Xu S (2005) Research on data collection and database update of GIS based on GPS technology. Proceedings of IEEE International Geoscience and Remote Sensing Symposium.
2. Pan J, Xu Q, Yang C (2014) Research and Application of GIS Data Update Technology. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 4: 195-200.
3. Maras HH, Altan MO (2000) Updating of a Geographic Database: an Application and Design of a Geographic Information System. International Archives of Photogrammetry and Remote Sensing 33: 616-623.
4. Soomro TR, Zheng K, Pan Y (1999) HTML and multimedia Web GIS. Third International Conference on Computational Intelligence and Multimedia Applications. 371-382.
5. Manchanda R (2009) Google Search Engine.
6. Brin S, Page L (1998) The Anatomy of a Large-Scale Hypertextual Web Search Engine.
7. Ratha B (2014) Search Engine.
8. Ames DP, Horsburgh JS, Cao Y, Kadlec J, Whiteaker T, et al. (2012) Hydro Desktop: Web services-based software for hydrologic data discovery, download, visualization, and analysis. Environmental Modeling & Software 37: 146-156.
9. Ames DP, Michaelis C, Dunsford T (2007) Introducing the Map Window GIS project. OS Geo Journal.
10. Dianzhi S, Xinyue YE, Tian G (2014) Open GIS for big data: opportunities and impediments. Progress in Geography 33: 723-737.
11. Teodoro AC, Duarte L (2013) Forest fire risk maps: a GIS open source application—a case study in Norwest of Portugal. International Journal of Geographical Information Science 27: 699-720.