THEMATIC MAPPING SERVICE FOR TIME SERIES GEOSPATIAL DATA VISUALIZATION

Teerayut Horanont[*1] and Ryosuke Shibasaki[2]
[1]Researcher, Institute of Industrial Science, the University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan; Tel +81-3-5452-6417
E-mail: teerayut@iis.u-tokyo.ac.jp

[2]Professor, Center for Spatial Information Science, the University of Tokyo
5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8568, Japan; Tel +81-4-7136-4290
E-mail: shiba@csis.u-tokyo.ac.jp

KEY WORDS: Thematic engine, Time-series visualization, Web services, Spatial analysis

ABSTRACT: Interactive spatial analysis package are available in a number of traditional desktop GIS. However, visualizing multi-temporal thematic maps through the Internet is still limit to fix contents and restrict change of the input data. The users with limited GIS knowledge are normally have difficulties to create and output thematic maps from generic data. In this study, we develop thematic mapping services that can be applied to non-spatial data format served through powerful map services solutions. Naive users who have no GIS software experience can simply input a plain text file with location identifier field such as place name or gazetteer to generate time-series thematic maps online. The system designed by this paper has great priorities in following three aspects: 1) Serving as a thematic engine for non-GIS data format; 2) Accelerating the common workflow and data redistribution by using online service; 3) Automate geographic coordinates and features discovery by utilizing location identifier. The system dedicates a great deal of effort to the initiated study of time-series geospatial analysis and visualization for naive users including those with no past experience using Geographic Information Systems.

1. INTRODUCTION

A thematic is a type of map or chart especially designed to show a particular theme connected with a specific geographic area. It generally displays the spatial pattern of a real world phenomenon and widely uses as qualitative analytic method to depict and analyze daily live information in geographical space. The most common use of thematic maps is the choropleth map in which areas are shaded or patterned to shows the measurement of the statistical variable on the map, population density or per-capita income is a common example. Thematic mapping has a long history in cartography and widely operate in PC software platform. However, as the Internet developed further and as online applications began to appear and play more important role in contemporary geospatial processing, especially in the context of geospatial visualization including thematic displays. What it has shown recently is the increased use of Web Map Services (WMS) for publishing maps in the web environment. Hong, J. et al (2005) has introduced OpenGIS technology for the creation of thematic maps based on a number of OGC standards like GDAS, GLS, and WMS. The proposed service architecture works on the basis of standardized interface among different servers and allow easier data integration and interoperability. Analysis of spatio-temporal data is especially challenging for advanced GIS interactive web mapping. It requires special treat of difference aspect including information (attributes), time (temporal) and space (spatial). Voss, H. et al (2000) implements a Descartes system which specifically design to support thematic mapping by automatically proposing alternative visualization of geo-referenced statistical variables, particularly addressing new ways of comparing multiple attributes in time-series data. Nevertheless, the system still relies on the predefined base map and input data. Since the thematic map can covers large geographic area, the amount of base map data needed is enormous, finding good base-map data and thematic data for spatial questions has always been difficult (Cammack, 2005). Commack proposes a prototypical example of data gathering by using WMS solution to collect base-map data from providers in real time and allows the cartographer to focus on developing the best thematic content of the map.

There are few of existing services available on the Internet at the current review of this study. The Plug and Play Maps created by GISmatters (http://www.plugandplaymaps.com/) is a free web service that lets you embed interactive thematic maps in your web pages with a single line of code that indicates the data you have to map and the way you want to display. Another example is Geobrowser (http://thematicmapping.org/) this site provide Thematic Mapping API that allows you to create KML based thematic maps from your own data source. None of them offer a wide range of flexible delivery of the thematic map based on simple data input in term of location discovery and spatio-temporal aspects. The Plug and Play Maps provide a built-in coordinate database for common features. However it is limited only to the use of state or country codes.

2. THE THEMATIC CLASSIFICATION SCHEMES

The main aim of thematic map classification scheme is to organize the information in a group, so that the users can interpret the classed maps easily. The concepts and methods associated with designing and producing thematic maps including data standardization and classification. The classification schemes are described in this section. The choropleth mapping capabilities, which are implemented in most GIS software, is suggested for the testing of this service. The classification scheme is the process of dividing raw data into classes or groups, each class represented by a unique symbol or color. As results, classed map makes it easier for the reader to discriminate among a large number of different color shades. In this study, the user interface for the proposed system provides a major classification schemes for creating choropleth maps including equal-interval, quantiles, natural breaks, and standard deviations.

2.1 Equal-Interval

The equal-interval classification scheme recommended only if the statistical distribution of a mapped variable is uniform. In this case, the predefined number of data classes divides the range of the entire dataset, such that each class occupies an equal interval along the range of data values. The major disadvantage of the equal-interval classification is that the data distribution is not considered when determining class breaks for the intervals.

2.2 Quantiles

The quantiles classification or equal count creates classes with equal numbers of data points in each class. This method produces maps that have an apparent balance so that each class is represented equally. The major disadvantage of the quantile classification is that the process does not consider how the data are distributed. This may give rise to misleading the map if the data have a highly skewed distribution.

2.3 Natural Breaks

The Natural Breaks classification uses an iterative algorithm to optimally assign data to classes such that the variances within all classes are minimized, while the variances among classes are maximized. In this sense, the data distribution is explicitly considered for determining class breaks. The major disadvantage is that it may be difficult to understood by the naive users and the legends values may not be intuitive.

2.4 Standard Deviations

The standard deviations classification assigns the data to classes based on the mean and standard deviations of the data distribution. This method works well if data is normally distributed since class breaking method reflects the variance of data values from the mean and the data range. The major disadvantage is that the classification method requires a basic understanding of statistical concepts and may not be easy for naive users to interpret.

Since choropleth maps use map color to represent statistical data. A simple gradual color schemes and division numbers are provided as standard visualization and present technique. Naive users who have no GIS software experience can easily create and understand a choropleth map by using classification scheme, setting the number of classes, and defining classification methods. Each of these factors can be manipulated to yield large collections of maps.

3. SYSTEM ACHITECHTURE AND DESIGN

The system is designed to cope with the heterogeneity of the Internet components and to ensure the interoperability among other services. This means that the standard data structure and interfaces are basic requirement of the system. Implementations of web service standards recommended by the Open Geospatial Consortium (OGC) such as, the Web Map Service (WMS), Web Feature Service (WFS) and Styled Layer Descriptor (SLD) provide a principle for communication and allow users to visualize spatial information as thematic maps. The core system working behind as a thematic engine composes of two important modules. The first module is used for spatial data discovery process and another is thematic cartography, where the map classification and rendering are performed at this stage. Figure 1 show the flow diagram how thematic maps creation process work in procedural steps. The two core functions were developed mainly using open data and openGIS library and software. Another core function of this system is a temporal control where the operation take place at the client side. The series of dataset can be integrated to prepare a spatial temporal visualization as static frame and animation. The detailed procedures of the core functions are described in this section.

3.1 Spatial Data Discovery Process Flow

The Spatial Data Discovery (SDD) Process is an automated system, which allows the input data prepared in a simple form. For instant, a simple Copy and Paste from a website or import a comma-separated-values (csv) file. The data should contain at least one column that represents geographic location or place name and the structure provide a cell separator with equal column in each row. The SDD process search for the matching place name and automate assign the geographic features of the dataset.

3.2 Thematic Cartography Process Flow

The Thematic Cartography (TC) Process performs thematic classification process based on a predefined configuration. The prototype system support 4 classification schemes as discussed in the section 2. Each of data classes are styled based on the Styled Layer Descriptor (SLD) and Symbology in support with Mapserver engine. When MapServer gets a valid SLD through a request, it parses this SLD to extract all the styles attached to the NamedLayers, and then it applies these styles to the map before it is returned to the client. Map legend is also created by a support of Mapserver after the class break or data range has been defined from the last step. The successive layer will be parsed to render on top of the client mapping system.

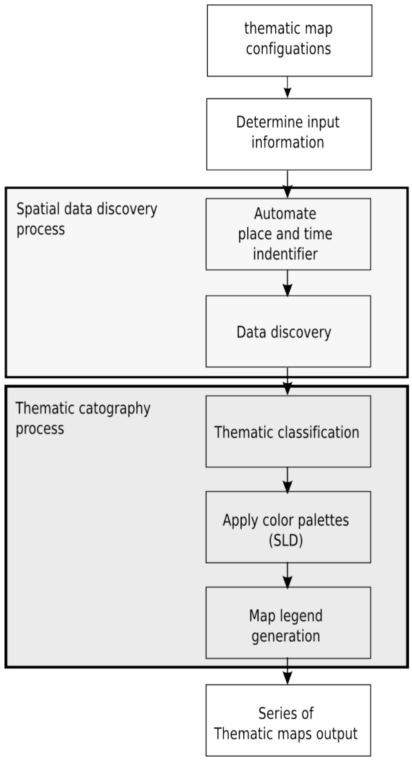

Figure 1. Flow diagram of thematic engine system architecture

3.3 Temporal Control

The temporal discovery (TD) process takes place at the same time when SDD is executed. The series of thematic maps are cached and parse on the client side. The time series maps can be deployed and displayed in static map frame or animate with time control. The feature is controllable via a user-friendly interface.

4. DYNAMIC POPULATION MAP

We demonstrate the use case of thematic maps engine with our previous research output. The dataset is the estimated dynamic population from the usage of the mobile phones.

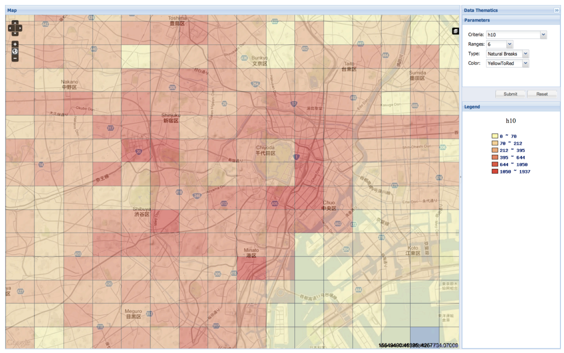

Figure 2. Web based prototype system

gid	code	h5	h6	h7	h8	h9	h10	h11	h12	h13	h14	h15	h16	h17	h18	h19	h20	h21	h22	h23
27	53392589	77	77	151	97	213	149	149	149	149	149	149	97	97	166	132	156	92	92	92
28	53392593	0	0	170	134	0	43	0	0	0	0	0	76	0	50	0	0	0	0	0
29	53392594	508	508	658	248	198	116	197	125	94	278	142	320	411	493	485	508	516	538	566
30	53392595	263	263	297	258	74	74	130	45	82	92	227	228	226	260	263	317	263	263	263
31	53392596	176	176	225	216	194	80	101	164	164	298	201	156	325	80	165	80	108	138	145
32	53392597	526	631	679	894	589	420	388	517	504	407	342	423	630	425	525	682	470	451	480
33	53392598	716	844	908	1117	733	577	538	381	429	361	508	519	788	557	509	803	578	688	655
34	53392599	0	41	99	201	106	106	106	148	194	212	152	106	106	42	0	0	0	0	0
44	53392650	0	35	0	84	107	26	26	26	26	71	61	26	26	0	91	0	0	0	0
49	53392660	11	116	59	25	105	25	25	25	14	14	49	14	14	14	128	14	28	28	0
51	53392662	15	49	15	162	242	242	242	242	228	228	272	272	228	228	137	0	0	0	0

Table 1. Raw data derived from the mobile sensing platforms

The population has been dynamically calculated from the location of active mobile phone devices in the area. The result provided the estimated population distribution in modular grid tile system with an hour interval. Table 1 give an image of simple raw data. Code attribute column is recognized as place identifier and will be automatically assigned with unique geographic feature. Please note that code name in this study is basically a standard mesh id of Japanese mesh system and that of available in the SDD database. Figure 3 illustrate the sample use case of this system where changing of map configuration can affect directly to the interpretation of output thematic map.

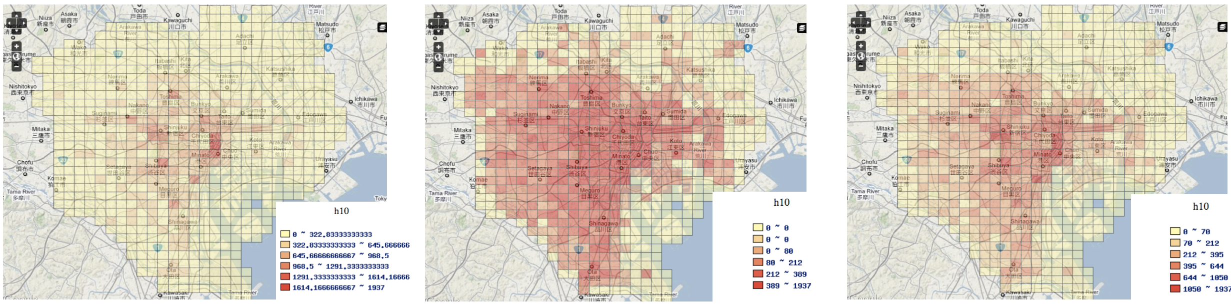

Figure 3. Thematic classification schemes, Equal-Interval (left), Quantiles (middle), Natural break (right)

5. DISCUSSIONS AND FUTURE WORKS

Given the technical possibilities, the use cases and the program prototype, the dynamic mapping and online spatial analysis services are at the beginning of a potentially explosive development cycle. Our goal with the research described in this study focus goes beyond the visualization issues of thematic maps. We aim to automate the thematic map generation so that novice users are able to visualize spatial data as thematic maps over the Web. The service allows naive users with no GIS software experience to generate online time-series thematic maps from a plain text input file containing a location identifier field such as place name. For future works, we aim to expand this study to discoverable data sources in a Spatial Data Infrastructure (SDI). In other words, we want to automatically generate spatial information (thematic maps) from spatial data (SDI data sources). The SDI data sources are not necessarily georeferenced; they might just contain location identifiers, similar to the input files in this study.

References

Cammack, R.G., 2005. Web Mapping Services: A Tool for Thematic Internet Maps, Proceeding in XXII International Cartographic Conference (ICC2005), A Coruna, Spain.

Dietze, L., Zipf, A., 2007. Extending OGC Styled Layer Descriptor (SLD) for Thematic Cartography - Towards the ubiquitous use of advanced mapping functions through standardized visualization rules, in 4th International Symposium on LBS and Telecartography, Hong Kong.

Hong, J., Lin, S., 2005. Web-based Thematic Map Service in OpenGIS Environment, Proceeding in the 26th Asian Conference on Remote Sensing (ACRS 2005), Hanoi, Vietnam.

Sae-Tang, A. & Ertz, O., 2007. Towards web service dedicated to thematic mapping. OSGeo Journal, Vol. 3, pp. 31-34.

Sandvik, B., 2008. Thematic Mapping Engine. MSc Thesis in Geographical Information Science, Institute of Geography, School of GeoSciences, University of Edinburgh.

Voss, H., Andrienko, N., Andrienko, G., Gatalsky, P., 2000. Web-based Spatio-Temporal Presentation and Analysis of Thematic Maps. The Journal of Cities and Regions, Journal of SCORUS, the Standing Committee on Regional and Urban Statistics and Research, November 2000, pp.51-61