Web-based Thematic Map Service in OpenGIS Environment

Jung-Hong Hong
Associate Professor, Dept. of Geomatics, National Cheng-Kung University
1, University Road, Tainan, China Taipei
junghong@mail.ncku.edu.tw

Shih-Yu Lin
Master, Dept. of Geomatics, National Cheng-Kung University
1, University Road, Tainan, China Taipei

Abstract: Thematic maps are nowadays widely used in a variety of research domains and human daily lives to depict and analyze spatial phenomena in reality. Though current Web-based GISs can already display thematic maps via internet, their content is nonetheless restricted to those maps already stored in databases. Naïve users, with limited cartographic knowledge and GIS software experience, can hardly submit their data and create thematic maps that meet their demands. In this paper, we proposed a web-based thematic map service that successfully integrates various services complying with numerous OpenGIS standards like GDAS, GLS, and WMS. The intelligent and automatic mapping capability of the proposed service is further improved by introducing built-in rule-based cartographic knowledge. Comparing to some of the current closed and inflexible centralized data archive; the proposed system architecture, though still preliminary, holds apparent advantages by having the standardized interface between servers, easier data integration and interoperability and the convenient map design environment to users without prerequisite knowledge.

Keywords: OpenGIS, Thematic Maps, Standards

1. Introduction

Thematic maps [1] are widely used map techniques to illustrate spatial distribution of real world phenomena. Even without professional mapping skill, a variety of users rely on thematic maps to depict and analyze spatial phenomena in their research domains. The rapid growth of GIS technology has made a huge difference regarding the creation of thematic maps. With easy-to-operate toolbox [2] and the availability of increasingly expanded digital data, naïve users found out the making of thematic maps was no longer a huge obstacle impossible to challenge, at least as far as the software technology is concerned. Though Vitek et al.[3] argued that such convenience may hide the risk that ill-quality maps are created much easier than before, yet the innovative GIS software environment and the potential of ubiquitous internet accessibility [4] arouse a world-wide technology revolution and this trend seems to be practically unstoppable. Unfortunately this fast development didn't mean Vitek's concerns will automatically disappear. For most of the GIS software, the successful creation of thematic maps would first require users to have correct data （in a particular format） beforehand, and further require users to have enough cartographic knowledge, and know how to issue appropriate commands. To many users, it is indeed true that GIS technology, as well as its applications in internet, largely removes the technology barrier to access thematic maps, but some difficult challenges remain to be resolved in the future. We argued in this paper that domain users not familiar with GIS technology have neither the ability to process software, nor enough skill to create data that holds correct corresponding relationship between spatial and aspatial attributes, a prerequisite demands for creating any thematic maps. They do have the needs for thematic maps, but they would also hope they don't need to deal with the complicated GIS operations and sincerely expect the thematic maps they need will be readily available as long as their data is ready.

Two major GIS technology threads in recent years are “open” and “distributed” [5]. The 19100 series of standards proposed by the ISO/TC211 and the abstract and implementation specifications proposed by the Open Geospatial Consortium (OGC) provided a unified and solid framework for the description, modeling, recording, transfer and internet distribution of geographic data [6]. With these standards, system operations, transferred data, as well as the communicating interface among geo-servers, can all comply with open and public standards and effectively improve the sharing and interoperability of geographic data [7]. By following selected standards, data and services can be created by authorized organizations in the distributed environment and linked thru standardized interface whenever necessary. With more and more commercial GIS software companies and communities supporting OpenGIS standards, the GIS technology is now under a tremendous revolution and it is surely interesting to investigate how this would impact the issue of thematic maps. We believe a web-based thematic map service should not be restricted to only display ready-made maps, it should be able to allow domain users to submit their own data, sometimes even without graphic description, and create thematic maps by built-in cartographic knowledge and services that comply with the OpenGIS standards. Furthermore, these services may serve different purposes and located in different places in the distributed OpenGIS environment. The immediate challenge is we longer have data with correct corresponding relationship between spatial and aspatial descriptions beforehand, and must create this mapping dynamically. Secondly, since all the required data may come from different servers, the thematic map service must have built-in knowledge to justify the received data content and make appropriate mapping decisions. This paper discusses our preliminary attempt to develop a web-based thematic map service that complies with various OpenGIS standards and the research result will also serve as the foundation of the service-oriented infrastructure for the years to come.

2. System Architecture

A thematic map depicts the spatial distribution of the selected theme, its successful creation therefore relates to at least three steps: (1) prepare data with correct spatial-aspatial corresponding relationship, (2) apply cartographic knowledge to determine mapping process, and (3) execute the mapping process. Current GIS software usually requires users to prepare data by themselves and provides wizard interface to guide users' mapping actions. To design a web-based thematic map service based on the OpenGIS technology, four major requirements are first identified:

1) Thematic data is provided by domain users
2) Thematic map service possesses built-in knowledge to judge data contents and perform necessary mapping actions
3) Data and services can be stored and processed separately in different servers
4) The communicating interface among different services must comply with OpenGIS standards

Fig. 1 shows the three-tier system architecture of the thematic map service we proposed. The three tiers respectively relate to (1) client that allow users to issue request and receive response, (2) thematic map servers that make cartographic decision and (3) basic services that provide necessary cartographic operations. These three tiers are connected in internet web-based environment, with the thematic map servers acting as the bridge between clients and basic services.

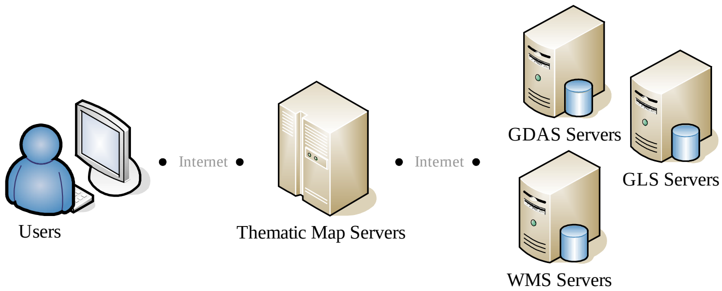

Fig 1. 3-Tier system architecture of the thematic map service

1) Client tier

To reduce software/hardware technical requirement, client should be able to interact with the thematic map servers with only standard browsers. This limits the request and response should follow HTTP protocol and the returned maps will be in image formats, such that it can be displayed in clients' browsers without further processing. Clients will be prompted a homepage to upload their thematic data and provide necessary descriptions about the uploaded data. Ideally the uploaded data should also comply with certain standards, but it is surely an overwhelming loading to clients. We instead choose Microsoft Excel as the designated format for the uploaded thematic data, because it is widely used in PC platform and has the capability to import various data formats like HTML table, CSV, TXT. MDB and DBF. This can of course be expanded in the future if necessary.

2) Thematic Map Server tier

As the center of the system architecture, the thematic map servers are in responsible fortasks that includes receiving client' requests, analyzing if maps can be created, accessing required data located in the distributed environment and applying cartographic knowledge to create correct thematic maps. Thru the homepage the thematic map servers provide, clients issue their mapping requests by identifying at least three types of information:

- geographic identification framework,
- data responsible organization, data theme, data time, and attributes for making thematic maps
- the uploaded data file.

Note the attribute data stored in Excel file does not have any graphic description, the thematic map server therefore needs to analyze if a correct corresponding spatial description file (e.g., digital map) is available. The successful linking of these two types of data allows the creation of correct thematic maps (Fig. 2). The geographic identification framework serves for this purpose and will be discussed in more detail later.

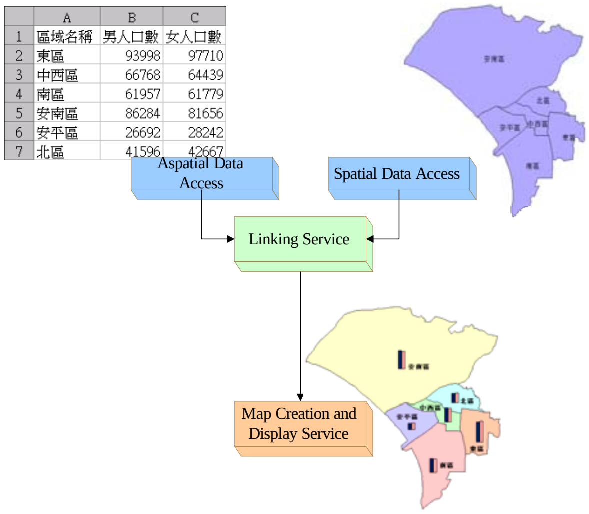

Fig. 2 The linking between spatial and aspatial data to create thematic maps

The various information about the uploaded data (i.e., metadata) is essential to the mapping process. Some are used to have a better understanding about uploaded data file, some are critical factors during the mapping process. For example, it is preferred the uploaded thematic data and the corresponding spatial description file referring to the same temporal extent, otherwise it is very likely some features will be located in their locations 10 years ago (if they have moved), and we will have no way to notice it. A series of thematic mapping rules were designed and included in thematic map servers as built-in knowledge to ensure the correctness of thematic maps. In the meantime, the thematic map servers actually do not hold any attribute or map data, it therefore must maintain a stable connection to the various servers in basic service tier.

3) Basic Service tier

One of the major merits of the OpenGIS standards is the standardized web service interface that takes away the barrier of commercial software format and thus allows easier data interoperability. A particular geo-server supporting selected OpenGIS standards is capable of fulfilling certain GIS needs and allowing standardized request/response from other clients and servers in the internet. After analyzing mapping procedures, three types of basic services are identified and can be operated with certain OpenGIS standards. These three services are Geolinked Data Access Service (GDAS), Geolinking Service (GLS) and Web Mapping Service (WMS) [8][9][10]. Servers complying with these standards are created accordingly in this paper.

1. GDAS Servers

A GDAS server complying with OGC GDAS standard can store attribute data with geographic identifiers (GID), and allow other clients or servers to request data that meet their demands. Geographic identifiers usually refer to a unique code that can be used to identify a feature with spatial extent, e.g., name of the political regions, street address of buildings, zip code, etc. Whenever data is collected in accordance with the spatial extent of GID (which is usually the case of thematic map data), it is advisory to include GID as one of the attribute and thus introduce an implicit spatial description to the attribute data. A GID framework therefore refers to a set of unique and systematic coding of GID, for example, there are 6 political districts in the Tainan city now, so the framework will consist of 6 different GID (East, South, Central-West, North, An-Nan and An-Ping) and can be named as "Tainan city political district framework." Since GID will be used as the reference for joining attribute and spatial data, GID framework is better created and maintained by authorized organizations. Whenever there is any change to a GID framework, a new version of GID framework should be created and the old version must still be saved for possible future use. For example, there were 7 political districts the Tainan city in the past, but the Central district and West district were merged to a new Central-West district beginning with the year of 2004. If the data is about the years before 2003, the 7-district version of GID framework should be chosen.

Following the GDAS standard, attribute data, together with its metadata, will be transformed and stored in XML format so that later internet communications between servers (e.g, data transfer) can be based on the XML technology. Fig. 3 shows an example of the recording data in the GDAS servers

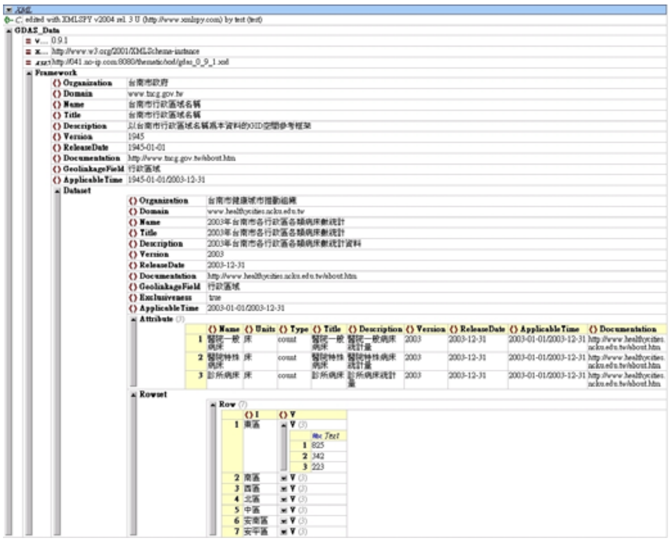

Fig. 3 Thematic data stored in GDAS Servers

2. GLS Server

A GLS server complying with the OGC GLS standard can build a link between the attribute data acquired from GDAS servers and its corresponding digital map data with common GID (similar to a "join" operation in relational database). A GLS server should have a digital map database to provide different GID framework and a mechanism to perform linkage operations. During the creation of thematic maps, it is a mandatory requirement that the data from GDAS server and GLS server must refer to the same GID framework, otherwise the mapping should not be executed any further. However, since these two types of services are created and maintained separately, it is certainly possible that no appropriate digital maps are available for creating correct thematic maps.

To improve the thematic map content, the proposed service includes a "reference map" design, such that clients can overlay roads or political districts to help map reading. From a cartographic perspective, the reference map data and GLS digital map should better refer to the same period of time and the same level of positional accuracy. Since this is not included in the specification of GLS standards, this is done primarily in thematic map servers with built-in cartographic knowledge.

3. WMS servers

Nowadays WMS capability is supported by almost every major commercial WebGIS. It works on the basis of map image transfer thru standardized request and response, and thus hide the detailed operations of commercial data format from the users. Since the response is in image format, it can be easily displayed and operated in internet browser or various types of image software. GeoServer（http://geoserver.sourcesforge.net/html/index.php）was chosen in this paper to serve as the WMS server. Since its default data format is ESRI shape format, the spatial data in GLS server is also in ESRI shape format to simplify mapping procedures. To appropriately adjust the layout and symbols in the thematic maps, OGC Styled Layer Descriptor (SLD) is used. Based on the types of thematic maps analyzed, different templates of SLD are applied.

3. Mapping knowledge and process

Based on the system architecture described in section 2, four different types of services can simultaneously locate in a distributed environment and fulfill certain task requirements. Though all services can run independently, it nevertheless requires a careful design to integrate all services and ensures the correctness of the thematic maps. This section provides a brief discussion about the mapping knowledge and the process flow of the mapping service.

1) Mapping knowledge

The successful creation of thematic maps requires professional mapping knowledge to bridge the gap between naïve users and mapping functions. Robinson et al.[11] had a detailed discussion about the factors influencing map design, e.g., measurement level, feature dimensionality, symbolization, etc. Built-in mapping rules in this paper are categorized into two groups: common rules applicable to all types of thematic maps and specific rules only applicable to a certain type of thematic map. Altogether 14 common rules regarding general properties, theme, reference data and annotations are developed. Each rule specifies the condition either data or service must fulfill. For example, rule 6 indicates that "thematic data must have corresponding spatial descriptions in graphic format", this would require the above-mentioned GID linking process. If the condition cannot be satisfied, the mapping process should be terminated. Some rules are optional during the mapping process, for example, rule 10 indicates "the scale of all reference map data should be within a specified range." Although this is certainly a preferred case, yet sometimes such data is simply not available. Another 23 specific rules regarding location maps, graduate symbol maps, choropleth maps, bar chart maps and pie chart maps are developed. Whenever the specific type of map is chosen, its corresponding specific rules must be enforced. For example, the pie chart maps require conditions like more than 2 submitted attributes must be selected, each GID must correspond to the same set of attributes, these attributes are mutual exclusive and their measurement levels are "ratio." These conditions must be verified one by one before any further action is issued.

Fig. 4 shows the four major steps the designed thematic map servers perform on the uploaded data. Only after the uploaded data can pass all the required tests, can the data be transformed into the format GDAS standard demands. Each step may involve a more detailed test procedure. Take GID framework judgment for example, a qualified GID framework in the GLS server means it maintain a correct corresponding relationship to the uploaded data from the perspective of (1)framework title, (2) temporal extent and (3) GID:

1. GID framework title

A digital map file with the same GID framework title is the first requirement that must fulfill. The thematic map servers therefore must be able to retrieve the GID framework title of all the digital map files in the GLS servers. The test returns a Boolean result of whether the mapping is possible. The title comparison is, however, not a rigorous constraint, and it is likely more than one digital map files may qualify. For examples, a number of digital map files may refer the same GID framework, but with different scales, positional accuracy, period of times, etc.

2. Temporal extent

It is mandatory for clients to specify the temporal extents of the uploaded data so that it can be compared against the temporal extents of digital map files in GLS servers. Following Allen's temporal logic [12], it is definitely preferred that a "contain" relationship exists between these two time records, meaning the digital map files can be safely applied to uploaded data. The Tainan city district adjustment case is a typical example of this issue. If a client submits a statistic data ranging from the year of 2000 to 2005, there is practically no single digital map file that can fulfill this mapping request because a new version is created in the year of 2004. No matter the time is modeled as a point or an interval, it is always possible that no overlap existing between the two recorded time. Though it does not necessarily mean the created thematic maps are wrong, certain risk is inevitable.

3. GID

Even if the title and temporal constraints are qualified, the linking relationship between the two types of data still requires one-to-one comparison. For example, a client may make a mistake about the reference time of the uploaded data（recorded as 2005 while it should be 2001）, the six-district version is found and returned. The GID comparison will immediately indicate that some of the data records do not have correct linkage and the completeness assumption is violated (in this case, statistics of the central and west district do not have correct corresponding relationships). We must, however, note that this violation does not always indicate mistakes. When a client has only the statistic data about three political districts, the linkage relationship is surely partial. A warning message will still be issued, but the thematic maps should still be created accordingly.

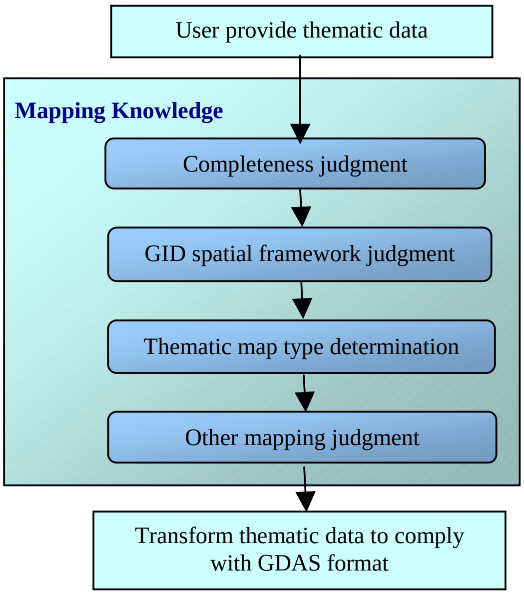

Fig. 4 Mapping knowledge in the thematic map server.

2) Mapping process

A typical creation procedure of thematic maps involves the interactions among all of the servers discussed in section 2. The following is a step-by-step description starting from users' map request till the completion of JPEG format thematic maps sent back to users (illustrated in Fig. 5):

Step 1: Users send mapping request and upload Excel format file to Thematic Map Servers via internet
Step 2: Thematic Map Server judges if maps can be correctly created, transform Excel file to XML format according to GDAS standards and send XML file to GDAS servers.
Step 3: Thematic Map Server sends GeoLink request to GLS servers via HTTP protocol
Step 4: GLS server sends getData request to GDAS servers via HTTP protocol to access thematic data (created in step 2)
Step 5: GDAS servers respond getData request and send back the XML format thematic data
Step 6: GLS servers link thematic data and corresponding spatial data via common GID, and register a new layer in the WMS servers
Step 7: WMS servers send the URL of the newly created map layer back to the GLS servers
Step 8: GLS servers respond a XML file to the Thematic Servers to indicate the Geolink process is completed
Step 9: Thematic Map Servers send a getMap request to WMS servers via HTTP protocol
Step 10: WMS servers response with the JPEG image created by Geolink process (created in Step 6)
Step 11: Thematic Map Servers make necessary adjustment and send the JPEG file back to the users.

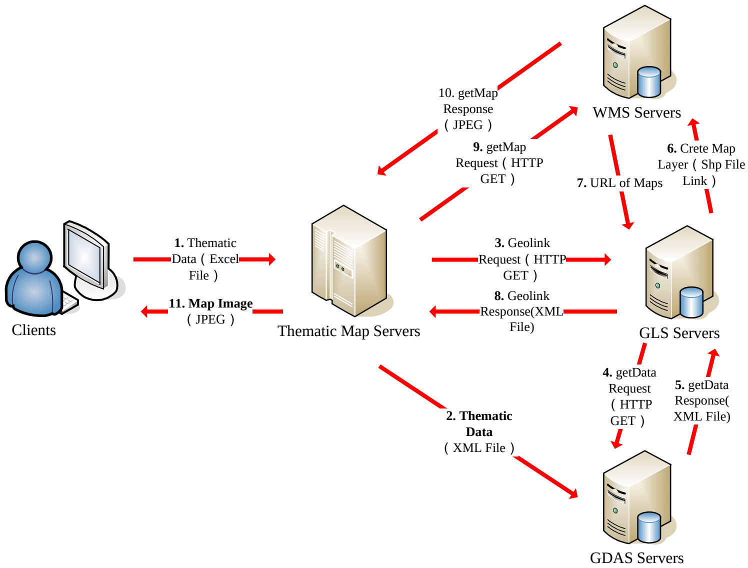

Fig. 5 Web-based thematic map creation procedures.

Fig. 6 demonstrates an example of thematic map servers' Geolink request (step 3). Note it includes the information about the requested URL of the GDAS server, the GID framework, attribute data, and the geolink GID.

http://041.no-ip.com:8080/gls/gls.jsp?Service=gls&Request=Geolink&Version=0.9.1&GDAS=http://gdas.041.no-ip.com:8080/gdas/gdas.jsp&GDASVersion=2004&FrameworkDomain=www.tncg.gov.tw &FrameworkName=台南市行政區域名稱 &FrameworkVersion=1945&DatasetDomain=www.healthycities.ncku.edu.tw&DatasetName=2003 年台南市各行政區各類病床數統計&Attribute=醫院一般病床,醫院特殊病床,診所病床&Geolinkids=東區,中區,西區,安南區,安平區,南區,北區

Fig. 6 Geolink request example issued by the thematic map server

4. Result Analysis

Table 1 shows the statistics of different types of hospital beds of the Tainan city in the year of 2003. The statistics is created on the basis of the 7 political regions, therefore the "Tainan city political district framework" should be chosen as the default GID framework. With the attribute data in table 1, several different types of thematic maps can be created. When the selected attribute is a certain type of beds, a graduated symbol map can be created. If more than two attributes are selected, a bar chart map or a pie chart map can be created to illustrate not only the spatial distribution, but also the difference in different categories.

Table 1. Test data of hospital bed statistics in Tainan city

	A	B	C	D
1	GID	Normal bed (Hospital)	Special Bed(Hospital)	Bed(Clinic)
2	East	825	342	223
3	South	289	72	138
4	West	386	243	87
5	North	999	294	252
6	Central	627	249	152
7	An-Nan	0	0	159
8	An-Ping	0	0	23

Fig. 7 shows the homepage the thematic map servers prompt to clients. Clients have to select the corresponding GID framework of the uploaded file (Step 1), fill in metadata and selected attributes (Step 2) and specify the pathname of the uploaded file (Step 3). The GID framework is designed as a "selected" pull down menu, meaning if no appropriate framework is available, then the thematic map service cannot satisfy clients' mapping request. In this example, all three attributes in Table 1 are selected and clients have to provide necessary metadata for each attribute. After clients submit their data, the thematic map servers analyze uploaded data and response analyzed results (Fig. 8). Clients are prompted with illustrations to select preferred types of thematic maps and reference map data (roads and political districts).

Fig 7. Clients' homepage interface

Fig. 8 Returned analysis result.

When more than one attribute is chosen, bar chart and pie chart thematic maps can be used to depict the difference among different categories or period of time（Fig. 9）. Unfortunately current WMS server does not provide such capability to draw bar or chart symbol. The current version of the thematic map service superimposes designed symbols to the output maps according to the specific location of GID, but does not have sufficient cartographic knowledge to automatically adjust their sizes and labeling. This may need additional works in the future.

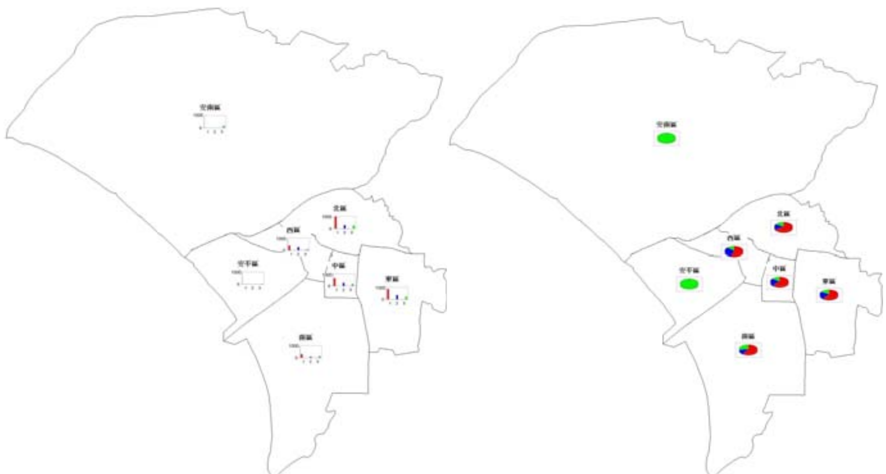

Fig. 9. Bar chart and Pie chart thematic map from Table 1.

Fig. 10 shows the thematic maps of the population density in the Tainan city. Although these two maps refer to the same original data source, but the making of graduate symbol maps and choropleth maps are actually based on different dimensionality of symbols (point vs. polygons). After clients select the preferred type of thematic maps, the thematic map servers apply built-in knowledge to select appropriate GID framework. If graduate symbol map is chosen, point-type spatial data in GLS server is chosen and the boundary of political district is superimposed. If choropleth map is chosen, polygon-type spatial data is chosen instead.

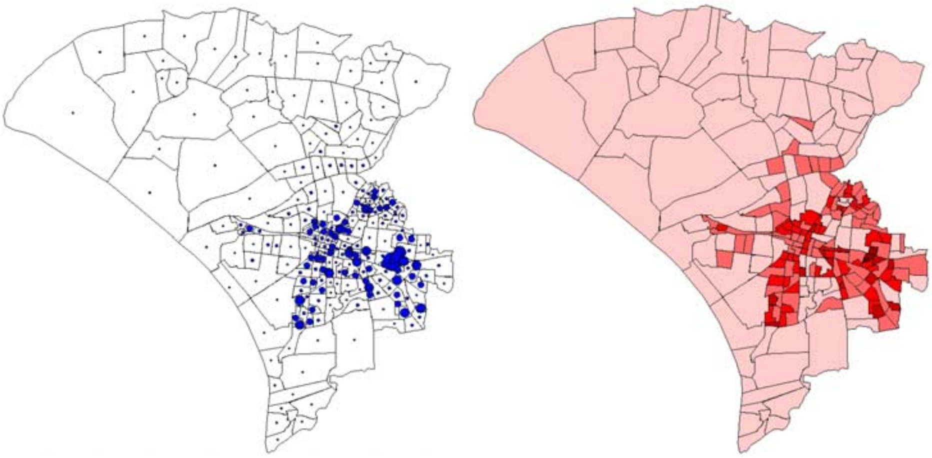

Fig. 10 Graduate symbol map and choropleth map for the population density in Tainan city

In the next example, because only the statistic data of four political districts is uploaded, a warning message indicating only part of the GID in the framework have been successfully linked is sent to clients (Fig. 11), but the mapping process continues and only political districts with statistic data are shown in the map.(Fig. 12)

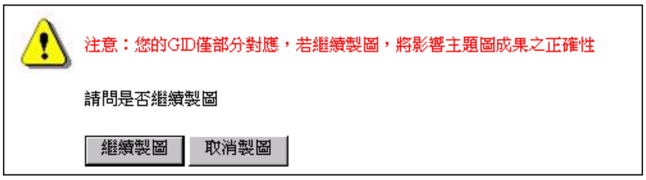

Fig. 11 Partial linkage warning message

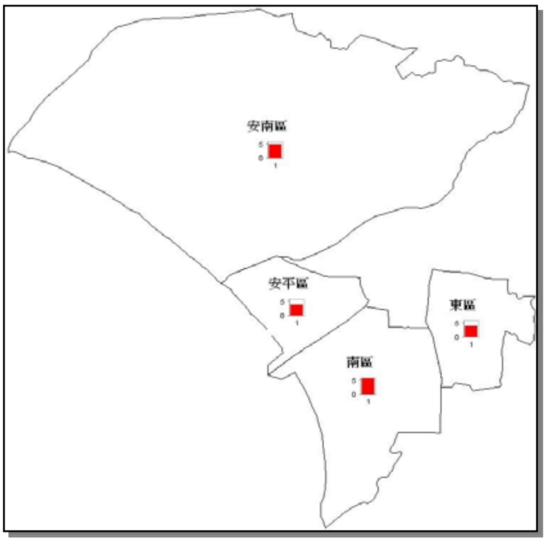

Fig. 12 Partial mapping result with only districts with statistic data

5. Conclusions

The OpenGIS technology opens a new dimension for the creation of thematic maps. In the past, the closed data format and insufficient professional skill impede many non-spatial domain users to create thematic maps fulfilling their needs. While most of the current WebGISs still only allow users to display and create thematic maps based on data already stored in their databases, we proposed a web-based thematic map service based on a number of OGC standards like GDAS, GLS, and WMS, which will allow naïve users to submit their thematic data via internet and receive thematic maps based on the data they sent from the service. The proposed service architecture works on the basis of standardized interface among different servers and allow easier data integration and interoperability. With built-in cartographic knowledge, the major breakthrough of this research is even the users have only little professional knowledge and thematic data without explicit spatial description, the thematic map can still be created accordingly. Furthermore, the separated nature of aspatial and spatial data (corresponds to GDAS and GLS servers, respectively) in the system architecture makes it possible for these two types of data to be managed separately in the National Spatial Data Infrastructure environment. Domains without spatial knowledge can maintain their own GDAS servers and take advantage of the proposed service to create their own thematic maps. In the mean time, the thematic map service and a variety of GLS servers can be maintained by the NSDI authority organizations, who often has superior ability to handle spatial data. Under this circumstances, naïve users no longer have to deal with unfamiliar data format and worry about the barrier of cartographic and information knowledge.

Acknowledgement

Part of the content of this paper was sponsored by the research grant of the National Science Committee (Project NSC93-2211-E-006-065 and NSC94-2211-E-006-071)

References

[1]Dent, B., 1996, *Cartography: Thematic Map Design*, 5th edition, William C Brown Pub.
[2]Dangermond, J. 1986, The software toolbox approach to meeting the user's need for GIS analysis, *Geographic Information Systems Workshop Proceedings*, pp. 66-75,.
[3]Vitek, J. D., Giardino, J. R. and Fitzgerald, J. W., 1996, “Mapping Geomorphology: A Journey from Paper Maps, through Computer Mapping to GIS and Virtual Reality”, *Geomorphology*, Vol. 16, 231-249.
[4]Peng, Zhong-Ren and Ming-Hsiang Tsou, 2003, *Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks,* John Wiley and Son Inc.
[5]Gardels, K., 1997, Open GIS and On-Line Environmental Libraries, *ACM SIGMOD Record archive*, Volume 26 , Issue 1, pp. 32 - 38.
[6]Kresse, W. and K. Fadiaie, 2004, *ISO Standards for Geographic Information*, pp. 55-96.
[7]Ummadi, Praveen, 2004, A White Paper on Standards & Interoperability in GIS, *Michigan State University Press*.
[8]Open Geospatial Consortium, 2004a, Geolinked Data Access Service, Version 0.9.1.
[9]Open Geospatial Consortium, 2004b , Geolinking Service, Version 0.9.1,.
[10]Open Geospatial Consortium, 2004c, Web Map Service interface, Version 1.3.
[11]Robinson, Arthur H., J. L. Morrison, P. C. Muehrcke, A. J. Kimerling, S. C. Guptill, 1995, *Elements of Cartography*, 6th Edition, Wiely.
[12]Allen, J,F,, 1983, Maintaining Knowledge about Temporal Intervals, *Communications of the ACM*, Vol.26, No.11, pp. 832-843.