The Third International Conference "Problems of Cybernetics and Informatics", September 6-8, 2010, Baku, Azerbaijan. Section #7 "Numerical Methods and Computational Technology"
www.pci2010.science.az/7/13.pdf

REAL-TIME METEOROLOGICAL DATA ANALYSIS AND MAPPING

Refik Samet[1] and Serhat Tural[2]

Ankara University, Ankara, Turkey
[1]*samet@eng.ankara.edu.tr*, [2]*serhattural@hotmail.com*

Information technologies have gained importance today. Also Geographical Information Systems (GIS) are becoming increasingly popular. GIS are utilized to improve efficiency, decision making and communication by integrating various multiple and complex sets of information. These systems provide a framework for management, analysis and representation of geographical information. There are three major components of GIS: the data sets and models which represent the raw information, the maps and globes in which this information is placed, and the processing and manipulation techniques that can be applied [1]. GIS can be divided into five components: hardware, software, data, people, and methods. All of these components need to be in balance for the system to be successful. No one part can run without the other [2].

Meteorological information is important to everyone, each minute of the day. Independently on the location, terrain and climate, particular sections of the country, and regions of the world, people have a more immediate need for meteorological data. Damage to life and property from severe meteorological conditions is all too common, but with improved decision support systems that incorporate the latest meteorological information technology including real-time meteorological data observations and GIS, notification, planning and recovery can be dramatically improved [3].

In the new millennium the need for instant meteorological information with the click of a mouse is not only desired, it is expected. Information renders knowledge; and knowledge spurs prudent decision making. Since the meteorological information affects every facet of our real-times from food production to energy consumption, public safety and travel, its impact is extensive [3]. The aim of this study is to help people to make more accurate and immediate decisions on the base of real-time meteorological information.

The easiest way to be informed about the current meteorological information for end users can be a web based information system through internet. But most systems are not real-time and do not have interactive map features. Analysis is commonly done before representation and the results of analysis are stored as an image in offline manner. When analysis results are represented as image, it loses interactive map features like vector data, zoom-in, zoom-out, selecting, querying, editing on map. Analysis period can be 1 day, 12 hours, 6 hours, 1 hour etc. While these systems are working offline, meteorological data changes can be sharply and the end users cannot reach latest useful information. It is possible to encounter these problems in offline systems. To solve these problems a real-time information system, which performs an analysis whenever the new raw data come, is proposed. So, end users can benefit with more accurate information and make better decisions about their related works. Proposed system can analyze various meteorological data like temperature, humidity, wind, rain, etc.

One of contributions of proposed real-time meteorological data analysis and mapping information system is the providing the real-time analysis and mapping whenever the latest real-time meteorological data arrives to the server via a determined protocol from meteorological stations. This contribution allows the system to store, analyze and map new meteorological data once. Existing information systems use the exact time intervals that are defined by system previously to initiate the analysis and mapping operations [4], [5]. According to existing methods the same meteorological data may be stored, analyzed and mapped repeatedly. Assume that an analysis period of the existing systems is 1 hour and the data of parameter **X** are changed once per 6 hours. In this case, during 6 hours the data storing, analyzing and mapping operations will be executed 6 times. The first time will be executed on the base of the new data of **X**, but the others will be executed by using the same data of **X** repeatedly. As a result the same data and results of analysis may be duplicated in the database which may cause chaos during queering to analyze and map old data and results of analysis. Other deficiency of this method is unproductive use of server processor time to analyze and map the same data repeatedly.

Another contribution is related to mapping operations. Existing information systems create a map once in the exact time intervals in different formats like jpeg, png, etc. and publish it through web as picture. So, existing systems do not take into consideration are there the new data or not at the input of system, they simply repeat the same analysis routines in the exact time intervals on the base last received data. As an example, let us consider an assumption in previous paragraph. In that case, the existing systems will create and publish the same map 6 times, although the data will be changed only one time during 6 hours. Proposed method creates the map of data and analysis results and renews each time when the data are changed.

This study aims to fill the vacuum in the area of decision making on bases of the real-time meteorological data and the results of analysis while system is collecting the data measured continuously by the meteorological stations. Proposed architecture to capture, preprocess, analyze, publish and represent the real-time meteorological data is shown in Fig.1.

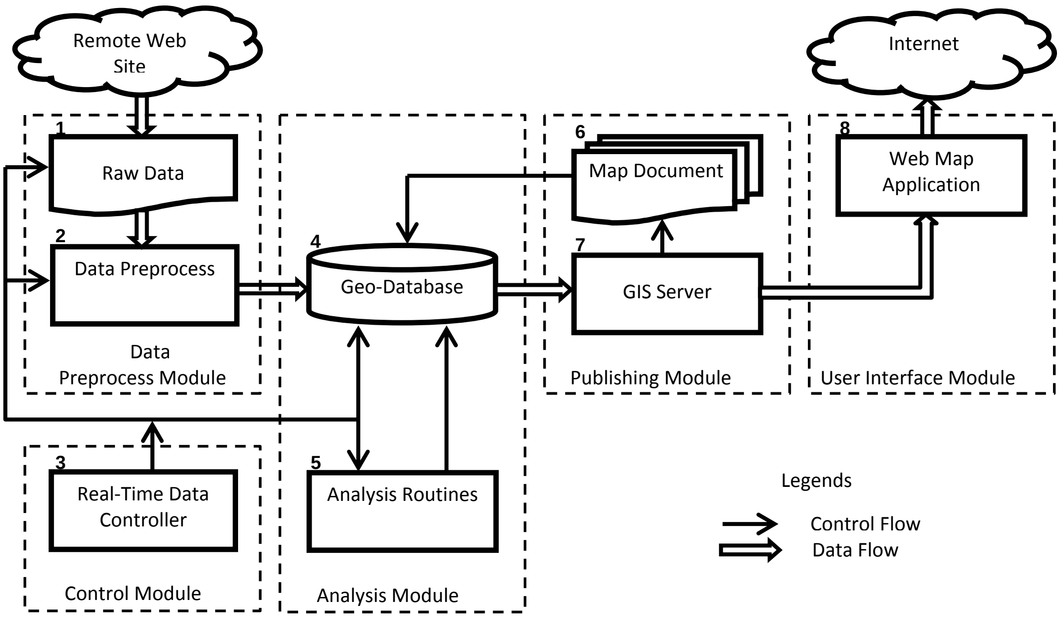

Fig.1. Architecture of real-time meteorological data analysis and mapping information system

Proposed architecture consists of five modules:

1) Data Preprocess Module;
2) Analysis Module;
3) Publishing Module;
4) User Interface Module;
5) Control Module.

Proposed architecture realizes the procedure which captures the new raw data come from meteorological stations and initiates preprocess, analysis, publish and representation operations. Mentioned procedure consists of four steps:

1) Transform, preprocess and store raw data (Data Preprocess Module);
2) Analysis routines (Analysis Module);
3) Publishing a map document as XML services (Publishing Module);
4) Creating the web application to represent a map for end users (User Interface Module).

Control Module consists of only Real-Time Data Controller (Block 3) which organizes above mentioned procedure to work whole system in synchronization and represents the raw data and results of analysis to end users in real-time through internet. Control Module controls the raw data (Block 1) which come from stations. If data are changed or the new meteorological data are come, Block 3 realizes this and initiates the transform, preprocess and store raw data step (Blocks 2 and 4). Simultaneously this module is responsible for analysis routines and initiates analysis step (Blocks 4 and 5). So, the whole system reacts to new meteorological data whenever data are changed with the help of Real-Time Data Controller.

Main architecture is a classic server client. Server includes web services, database and application server. Clients are end users which can access real-time map via their web browsers.

Meteorological stations continuously produce meteorological data. These raw data can be transferred from stations to server in several ways. For example, they can be transmitted as file via web service, FTP, socket and transmitted file can be in XML, text, excel, and other formats (Block 1). This file includes latest metrological data which will be analyzed and represented to end users. The fields of the record are separated with a special character.

First, the records of the latest file are transformed to store in spatial database. After reading the raw data from a file, it is need to treat these data as a geographical object. Transform raw data to a geographical feature is need for mapping, analyzing and all other operations in spatial domain (Block 2). All raw data are transformed to the point features (because all stations are as point in spatial domain) and are stored in Geodatabase (Block 4).

Geodatabase is a storage mechanism for spatial and attribute data that contains specific storage structures for features, collection of features, attributes, relationships between attributes and relationships between features. Received raw data are stored in a Geodatabase (Block 4) to query, analyze and benefit its advantages [7].

The geodatabase structure includes two tables. One of them is used for historical purpose to store all received data. This table may also be used for statistical purpose. Second table is used for temporary storing the last arrived data. Two tables are used to increase the performance of the system. Besides, there is also a background layer for the map that will be served to end users in geodatabase. For example, a layer may include a background with country boundaries.

The objectives of data analysis routines are to allow the end users to set queries and to retrieve useful information to satisfy the specific requirements of decision makers, and an important function of the analysis is the ability to predict what will occur at a location, at another point in time and under certain conditions. The most important analytical process of the GIS is the provision of capabilities for spatial analysis functions that are responsible for the manipulation and analysis of the spatial data.

Currently, the analytical capabilities of GIS related to the structure of the database (raster or vector) are used, and the proposed prototype uses the raster GIS structure because that raster family is determined to have greater analytical power. System provides the user with two kinds of analytical capabilities as introduced in the following sections: statistical analysis and geostatistical (spatial) analysis routines [8].

Analysis module is initiated by event of new arrived data. This module analyses the last arrived point features and its value on the geodatabase. Meteorological data needs interpolation for suitable analysis. The input of the analysis module is the point data or layer which is converted to the last read spatial objects on the geodatabase. The output of the analysis module is a raster file. This file will be last layer which will be added to the map. In order to represent the map in understandable form raster file is colored.

Statistical analysis can be useful for end users (Block 5). Daily maximum, minimum, average values are statistical analysis samples. It can easily calculate without need of spatial information of stations. Almost all web sites about weather give information to end users like current temperature, average temperature, minimal and maximal temperature.

Geostatistical Analysis (Block 5) provides a cost-effective, logical solution for analyzing a variety of data sets that would otherwise cost an enormous amount of time and money to accomplish. Identifying variation in natural phenomena to assessing possible environmental risks is very important for end users.

In natural resources the point field sampling is often used for spatially oriented projects and interpolation methods are implemented to predict the values in an unsampled location and to generate maps. The mapping and spatial analysis often requires converting the field measurements into continuous space. Therefore the point data sets must be converted to a continuous form using an interpolation method [9], [10].

This step is the most time consuming step. The interpolation is needed to calculate the values for all coordinates on the base of the raw data getting from specific coordinates. There are a lot of interpolation methods for geostatistical analysis.

SOA (Service-Oriented Architecture) based architecture is used to integrate other platforms like mobile, desktop and other web platforms. This map is also serviced for applications corresponding to standards of Open Geospatial Consortium, Inc (OGC).

A map document, which includes a country layer for background, latest received data layer and analysis results layer is created. Map document file (in mxd format) will be served as representation and it will contain the references about what kinds of data will be represented. Three layers will be represented for end users in proposed system. The first of them is a layer pointed to the meteorological stations from where the points and their values were arrived. The second is the background layer used to allow the end users to see and understand the map in more clear and visible form. This layer contains the information like the countries and their boundaries, cities and their boundaries, etc. And the last layer is a raster data which are created in the result of geostatistical analysis and will be used by end users for decision making. This layer represents the interpolation at the result of geostatistical analysis.

Representation the map as service: In order to be accessible from web and other sources and also from mobile, desktop and other web environments the map document it is necessary to convert it as service form. For this purpose GIS server products are used. As input the map document file (in mxd format) is used. As output the different standard formats like XML, SOAP, REST, etc. are offered.

As final step it is necessary to serve and represent a map prepared in previous step. To this end, a web application was developed which takes map service and converts it as raster images. Web application represents the maps and layers, handles error management, prints maps and does other functionalities. As a result, the end users (clients) can use real-time map.

In this study, we have realized a dynamic architecture that executes the full procedure to get raw data, to analyze these data statistically, to analyze the raw data and results of statistical analysis geostatistically and publish the results as map in real-time.

Ease of use and efficiency were achieved using ArcGIS products. Although a virtual machine is used for development and testing, its performance acceptable for a real-time map system. Whole system can react to the new data during 10 seconds. The geotatistical analysis routines cover a large portion of this time. Krigging interpolation method gives the best result. The speed of proposed real-time analysis automation system is acceptable according to real-time data analysis applications.

During each analysis period, ArcGIS Server needs to be stopped for a short time and started again for writing the analysis results. Although a user do not aware of this situation, it can be thought as a disadvantage of this system. Solving this situation is the subject of future works.

Creating the user interactive version of proposed system, where the analysis and mapping operations will be done according to the user's requests, is the subject of another future work.

References

[1] M.P. Harding, "GIS Representation and Assessment of Water Distribution System for Mae La Temporary Shelteri Thailand", *MIT Civil Engineering Master Thesis,* 2007.

[2] Spatial Information Clearinghouse, http://maic.jmu.edu/sic/gis/components.htm, [Access: April, 2010]

[3] R. Copley, "Real-time Weather in GIS", [Online]. Available: http://proceedings.esri.com/library/userconf/proc00/professional/papers/pap711/p711.htm, [Access: Feb. 2, 2010].

[4] O. Ozdilek and D. Z. Seker, "A Web-Based Application for Real-Time GIS", *XXth ISPRS Congress*, Istanbul, Turkey, 12-23 July, 2004, pp. 1-5.
[5] R. A. Mcpherson, C. A. Fiebrich and others, "Statewide Monitoring of the Mesoscale Environment: A Technical Update on the Oklahoma Mesonet", *Journal of Atmospherıc and Oceanıc Technology,* March 2007, pp. 301-321.
[6] D. Burke and P. Dodd, "Easy Access to Near Real-Time Data", *ArcUser Magazine Spring 2009*, March 2009, pp. 36-38.
[7] Working with the Geodatabase: Powerful Multiuser Editing and Sophisticated Data Integrity, ESRI White Paper, February 2002
[8] Al-Sabhan W, Mulligan M, Blackburn GA. "A real-time hydrological model for flood prediction using GIS and the WWW". *Computers, Environment and Urban Systems,* 2003 pp. 9-32.
[9] Siska P P, Hung I K , Assessment of kriging accuracy in the GIS environment [online], 2005.
http://gis.esri.com/library/userconf/proc01/professional/papers/pap280/p280.htm
[10] Heywood, H., Cornelius, S. and Carver, S. "An introduction to geographical information systems", 1998, Longman, Harlow.
[11] M. A. Oreal-Timer and R. Webster, "Kriging: a method of interpolation for geographical information system", INT. J. Geographical Information Systems, 1990, Vol. 4, No. 3, pp.313-332.
[12] S. Sensoy, Y. Ulupinar, M. Demircan, I. Balta, T. Tastekin and I.Alan, "Using ArcGIS in Climatological Applications", *12th ESRI User Conferences,* 2007, Ankara, pp. 1-11.
[13] Konstantin Krivoruchko, "Introduction to Modeling Spatial Processes Using Geostatistical Analyst", http://www.esri.com/library/whitepapers/pdfs/introFmodeling.pdf, 2006.