Multidimensional Analysis and Location Intelligence Application for Spatial Data Warehouse Hotspot in Indonesia using SpagoBI

This content has been downloaded from IOPscience. Please scroll down to see the full text.
2016 IOP Conf. Ser.: Earth Environ. Sci. 31 012011
(http://iopscience.iop.org/1755-1315/31/1/012011)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 37.187.190.247
This content was downloaded on 01/07/2016 at 20:29

Please note that terms and conditions apply.
Multidimensional Analysis and Location Intelligence Application for Spatial Data Warehouse Hotspot in Indonesia using SpagoBI

Gamma Uswatun Hasanah¹, Rina Trisminingsih²
Department of Computer Science, Bogor Agricultural University
Jalan Meranti Wing 20 level 5 Kampus IPB, Bogor, West Java, Indonesia

E-mail: gammauswatun@gmail.com

Abstract. Spatial data warehouse refers to data warehouse which has a spatial component that represents the geographic location of the position or an object on the Earth's surface. Spatial data warehouse can be visualized in the form of a crosstab tables, graphs, and maps. Spatial data warehouse of hotspot in Indonesia has been constructed by researchers from FIRM NASA 2006-2015. This research develops multidimensional analysis module and location intelligence module using SpagoBI. The multidimensional analysis module is able to visualize online analytical processing (OLAP). The location intelligence module creates dynamic map visualization in map zone and map point. Map zone can display the different colors based on the number of hotspot in each region and map point can display different sizes of the point to represent the number of hotspots in each region. This research is expected to facilitate users in the presentation of hotspot data as needed.

1. Introduction
A spatial data warehouse is a collection of spatial data that is subject-oriented, integrated, time variant, non-volatile and used in decision-making process [1]. Spatial data is data that represents the geographic location of position or an object at the Earth's surface. The source of spatial data includes graphic data of analog map, satellite imagery, field survey results data, and measurements result in data using global positioning system [2].

Spatial data warehouse development by integrating thematic and geographical data to produce spatial data cube for multidimensional spatial data analysis [1]. Some of the research that utilizes spatial data warehouse include [3], [4], [5], and [6]. The studies did construction of spatial data warehouse for distribution in the hotspot in Indonesia that could perform a multidimensional analysis and show cartographic visualization which tabular charts features. The system was built using Geomondrian framework. The Geomondrian framework has OLAP visualization features, spatial query function with MultiDimensional eXpression (MDX) and visualization map supports using Geoserver as server folder. But in 2011, the Geomondrian framework stopped doing updates version so that the...
development of spatial data warehouse system to detected hotspot with Geomondrian difficult to proceed.

This research did spatial data warehouse visualization with multidimensional analysis module and location intelligence module using SpagoBI. SpagoBI is one of business intelligence tools that are complete, flexible, open source and has 16 features so that find the best many solutions who able to save time and economic resources [7]. According [8], benefit from SpagoBI is an addition facilitates integration system of new modules that appear to do the development and repair of deficient products. The multidimensional analysis module SpagoBI is a module that provides the OLAP engine is flexible and user friendly which can monitor data from different level – level through the drill down, drill up, slice, dice, and so on. On the location intelligence module, SpagoBI provides two geographical (GIS and GEO) that can make the relationship between geographic data and business data in the data warehouse.

In this paper, spatial data warehouse hotspot used are obtained from previous research [9]. The location intelligence machine feature that was selected is GIS that was able to display dynamic map. This research used SpagoBI application so that users expect to more quickly displayed hotspot summary data.

2. Data and Method

2.1. Data
The spatial data warehouse had been done in the previous research [9] which is sourced by Fire Information for Resource Management (FIRM) National Aeronautics and Space Administration (NASA) in year 2006 to 2015.

2.2. The analysis of spatial data warehouse hotspot
The early stages in this research were spatial data warehouse hotspot analysis from previous research [9]. The purpose spatial data warehouse hotspot analysis was to know the characteristics of the used data. The results is used to know the scheme, table dimensions, and table of fact. Therefore, the result is used to design of development system.

2.3. Design and implementation of multidimensional analysis module
OLAP visualization is aimed to spread hotspot data presentation. Design of multidimensional analyis module was make the multidimensional data cube using SpagoBI Studio. The steps in implementation of multidimensional analysis was made hierarchy of each dimension table, made server SpagoBI, made and loaded off the Mondrian template. The result is OLAP operations featuring analysis either a drill down, roll up, slice, dice, and pivot. The OLAP operations are displayed in tabular chart and crosstab table.

2.4. Implementation of dynamic map location intelligence
Spatial data warehouse hotspot from previous researched [9] visualized dynamic maps using SpagoBI applications. The steps in implementation of dynamic map location intelligence on SpagoBI i.e. saved Indonesia map in Geoserver, made SpagoBI Dataset, and made dynamic map SpagoBI. Indonesia map in Geoserver needed a worksheet (workspace) as a space to accommodate a map. Dataset SpagoBI can be CSV file, database queries, Web service Class, Java, scripts, QB, Custom, and Flat. Dynamic map SpagoBI was divide into 2 types map i.e. map zone (different color areas) and map point (point size). The result of this stage was the dynamic map showing the difference in the color and size of points based on the number of hotspots each year.
2.5. Testing multidimensional analysis module and location intelligence module

This stage was tested the main functions on the application using Blackbox method. The multidimensional analysis module has 4 main functions such as select data for the dimension and measure, display data in a graph, display data in a crosstab table, and display OLAP operations. The location intelligence module has 5 main functions such as map zone visualizations, map point, zoom in and zoom out, displays indicators map, and do filter map. Every function was tested to make sure the system could run and produce the expected output. The creation of OLAP visualization system and dynamic map using spatial data warehouse hotspot had been completed when the entire test was declared successful.

3. Result and Discussion

3.1. The analysis of spatial data warehouse hotspot

The data used was hotspot data of forest fire in year 2006 to 2015, and administrative map data of Indonesia by 2014 from Geospatial Information Agency (BIG). Spatial data warehouse design by [9] generated star schema with one table of fact as a center that contain a foreign key for each dimension and number of hotspot for measure. The results of the star schema can be seen in Figure 1. Analysis data was based on hotspot of the year, quarter, month, satellite name, island, provinces, and municipalities.

![Figure 1. The result of star schema.](image)

After created spatial data warehouse design, [9] did implementation ETL model using Geokettle. The results of the ETL was table of fact and dimension tables was loaded into DBMS PostgreSQL. Based on the ETL process performed [9], the time dimension table has four columns with the number of rows attribute as many as 120. Satellite dimension table contains two columns with the number of lines by two lines. Location dimension has five columns with a number of lines as many as 505 where there are 6 Islands, 34 provinces, and 505 districts. Table of Fact consists of four columns attribute with the number of lines as much as 32937.

3.2. Design and implementation of multidimensional analysis module

This stage was designed OLAP with build multidimensional data cube using SpagoBI Studio 5.1.0. The results of ETL implementation [9] in PostgreSQL was linked to SpagoBI Studio using PostgreSQL JDBC Driver 8. Type of fact table was changed to cube and jumlah_hotspot selected as a function of measure. Table type in the time dimension, satellite dimension and location dimension was converted into a type of dimension. Then every dimension was created a hierarchy.

OLAP implementation was made the server SpagoBI and made Mondrian OLAP templates with XML format. The result of the implementation spatial data warehouse OLAP was showed hotspot in tabular chart and crosstab table. There were OLAP operations features (roll-up/drill-up, drill down, slice and
dice, pivot) featuring graphs, data storage, print files etc. Figure 2 is an example of a crosstab table results and graphs are formed.

![Figure 2. Representation hotspot data in the form of crosstab tables and graph.](image)

### 3.3. Implementation of dynamic map location intelligence

Attribute on Indonesia map was modified by adding a new string type attribute columns using Quantum GIS application (QGIS). Indonesia map was modified because the existing of id_lokasi type was integer while id_lokasi can be shown on SpagoBI must of string type. The Indonesia map has been modified and then saved in Geoserver so that you can perform on SpagoBI. Geoserver can display large map with the URL to access it. Worksheet (workspace) and storage areas (stores) were created as a space to accommodate a map. Created workspace done in order the hotspot area only displays the region of Indonesia. At created workspace, configuration system did reference coordinates and bounding box. The coordinate reference system is a reference point to define the coordinates of a point vertically or horizontally. The bounding box represents the limits of the Earth's surface in the area that would later be shown. In this research, a coordinate reference system used WGS-84 (EPSG: 4326) and on the bounding box selected reproject native to declare because the value of native SRS unknown so it can't be matched with the EPSG.

Dataset serves as definition media of the number of hotspots data that appears on the map. This research used the dataset query in the form of a database showing the columns id_lokasi, district, province, and the number of hotspots each year. The addition of new string typed id_lokasi column at lokasi dimension table was done because SpagoBI read string typed column. Any program code requires union to combine a number of hotspots from 2006 – 2015.

Map visualization type used in this research is a dynamic map that was able to display a map in the shape of polygon on map zone and in the form of point on map point. The selection of module for visualization of map on SpagoBI was location intelligence. JSON formatted dynamic map template was needed to making dynamic map.

Dynamic map SpagoBI has 4 sections. In figure 3, the biggest section was the Indonesia map that displayed the different of colors or the sizes of the number of hotspots. The top right section was the part to select map zone appearance which displayed the map in the form of polygon and map point that displayed the map in the form of point. Indicator section displayed selection of hotspot each year (2006 – 2015) in Indonesia. The bottom right section is the filter menu based on districts, provinces and the number of hotspots. The number of filter shown based on datasets that have been previously entered.
Dynamic maps visualization displayed a map of the hotspot in 2006 in the form of polygons by default. The differences colors from light blue to red indicated changes the number of hotspots. The more red color was displayed, then the larger the number of hotspots in the district. The number of hotspot analysis was represented by number of hotspots in 2006 with the largest number of hotspot dominated by Sumatra and Borneo. The largest number of hotspot in Sumatra were located in the province of Riau and South Sumatra Province. The largest number of hotspots in Borneo were in the province of West Kalimantan, Central Kalimantan, and some territory in East Kalimantan.

Figure 4 is point-shaped visualization of dynamic map SpagoBI. The sizes of each point represents the frequency of the number of hotspots in each district. The point location that appears to represent hotspot for one district which are not necessarily representative of regions with specific longitude latitude has a particular hotspot. When zoom in map, the map did not show the visible area to the level of districts because the maps are stored on the Geoserver is an administrative district map and the scope of the data used [9] only up to the level of the district.

Each point describes the location of the hotspot in the area that contains the code of the district, the name of district, the code of province, and the name of the province. If seen in use terms of map visualization, map zone used when users want to know the overall area that has one of the largest known hotspots of red on the map. Map point used when the user wants to see a more detailed areas that have the largest number of hotspots. That is because the size of the point represents the frequency of the number of hotspots so that users know which areas that became the top priority in handling the emergence of hotspots.
3.4. Testing multidimensional analysis module and location intelligence module

Testing system was performed with scenario test in functions for each module. The black box testing method is used in this research. Testing was conducted to find the function of multidimensional analysis module and location intelligence module was already running. Table 1 shows the scenario of multidimensional analysis module and table 2 shows the scenario for location intelligence module.

**Table 1. Scenario of multidimensional analysis module.**

| No | Testing Scenario | Expected Result | Test Result |
|----|------------------|-----------------|-------------|
| 1  | Select data for a dimension and measure | Choose one of the dimensions (time, satellite, location) | The data is displayed according to the selected dimensions | Successful |
| 2  | Display the data in a graph | Select the dimension, and then select the 'graph'. | The data represented in a graph | Successful |
| 3  | Display data in a crosstab table | Select the dimension, then by default the crosstab table appears. | The data appear according to the selected OLAP operations | Successful |
| 4  | Featuring OLAP operations | Select the dimension, the OLAP crosstab will display operation, select the desired OLAP operations | The data appear according to the selected OLAP operations | Successful |

**Table 2. Scenario of location intelligence module.**

| No | Testing Scenario | Expected Result | Test Result |
|----|------------------|-----------------|-------------|
| 1  | Display visualization map zone | Choosing the zone map | Differences in color maps for each district | Successful |
| 2  | Display visualization map point | Choosing the point map | Appears spheres of different sizes corresponding number of hotspots for each district | Successful |
| 3  | Displays zoom in and zoom out map | Selecting enlarge or reduce the map view | A map is displayed to be larger or smaller | Successful |

4. Summary

This work is able to make multidimensional analysis module and the location intelligence module for spatial data warehouse of hotspots in Indonesia from the years 2006 to 2015 using SpagoBI. Multidimensional analysis module is able to display OLAP visualization in crosstab table and graph, therefore it can help to present the hotspot summary. Location intelligence module uses GIS engine that is able to display a dynamic map in the form of map zone and map point that capable of running its functions properly. Visualization of the map zone can display different colors based on the number of hotspots in each district and visualization of map point can display different sizes of point that represent the number of hotspots in each district.

References

[1] Han J, Kamber M, Pei J. 2012. *Data Mining: Concept and Techniques. Ed ke-3*. Amsterdam, Belanda: Elsevier, Morgan Kaufmann.
[2] Ekadinata A, Dewi S, Hadi DP, Nugroho DK, Johana F. 2008. *Sistem Informasi Geografis untuk Pengelolaan Bentang Lahan Bebas Sumber Daya Alam*. Bogor, Indonesia: World Agroforestry Centre.
[3] Trisminingsih R. 2010. *Pembangunan spatial data warehouse berbasis web untuk persebaran hotspot di wilayah Indonesia*. Bogor, Indonesia: Institut Pertanian Bogor.
[4] Fadli MH. 2011. *Data warehouse spatio-temporal kebakaran hutan menggunakan Geomondrian dan Geoserver*. Bogor, Indonesia: Institut Pertanian Bogor.
[5] Imaduddin A. 2012. *Sinkronisasi antara visualisasi peta dan query olap pada spatial data warehouse kebakaran hutan di Indonesia*. Bogor, Indonesia: Institut Pertanian Bogor.
[6] Kusumah B. 2014. *Pengembangan modul update data pada sistem spatial data warehouse hotspot*. Bogor, Indonesia: Institut Pertanian Bogor.
[7] Cazzin G. 2012. *Business Intelligence with SpagoBI*. Padua, Italy: SpagoBI Competency
[8] Golfarelli M. 2009. “Open source BI platforms: a functional and architectural comparison” in 11th International Conference DaWak. Bologna, Italy, pp. 287-297

[9] Astriani W. 2015. Modul extract, transform, dan load untuk spatial data warehouse titik panas menggunakan Geokettle. Bogor, Indonesia: Institut Pertanian Bogor.