Cartography Using Data Warehousing and Mining

Narinderjit Kaur, Dr. Vivek Chandra
Head of Department, G.G.S. Khalsa College, Kamalpur (Ludhiana)
Head of Department (I.T.) , M.P. East Zone DIS.COM H14, Jabalpur (M.P.)
narinder.rajal@gmail.com

Introduction

Cartography is the art and science of making maps For preparing a map one has to collect data on the basis of original ground survey of a part of the earth on given scale. Making of globes charts and relief models is also included in cartography; different scholars have defined cartography in their own way.

Raisz Eruein said Cartography is" the Art and Science of making maps, charts, globes and relief Models". According to Buchanan, R.O Cartography is " The making of maps and charts in a wider sense the terms Covers the process of map making, from field survey or compilation to printed form" and Smith Jackie said Cartography is " the Science of constructing maps and charts. It includes the making of original surveys, the selection of suitable map projection and decisions on colors, layer tinting and other visual representations".

The above definitions clear that cartography is an art and science of map making and map is a carefully designed instrument for recording , displaying , analyzing and calculating and in general understanding the interrelation of things in their spatial relationship.

In starting years, cartographers collect the data and create the graphs, maps using their skills and methods. When Photography and Plastic films were introduced in cartography, then the work of mapping specialists were changed, that was the time of analog cartography. Analog cartography is useful for technical people not for society .Public did not have ready to access analog mapping technologies such as mechanical devices like Printing Press, telescope and photochemical technology etc. Because only technical people give the best result using these devices . The revolutionary change in 20th century is the use of computer . Society and cartographers use computers for several purposes. The use of computer in cartography technology is known as Digital Cartography. Today computers are individualized not institutionalized .The availability of computers and peripherals such as monitor , Printers, scanners along with computer programs for visualization, image processing and database management have democratized and greatly expanded the making of maps. With the help of Digital Cartography cartographers as well as society easily use and access Cartography.

Digital cartography provides two distinct product, one to satisfy each of the former function that maps alone once served

1. Digital database
2. Cartographic visulization.

Data Warehouses

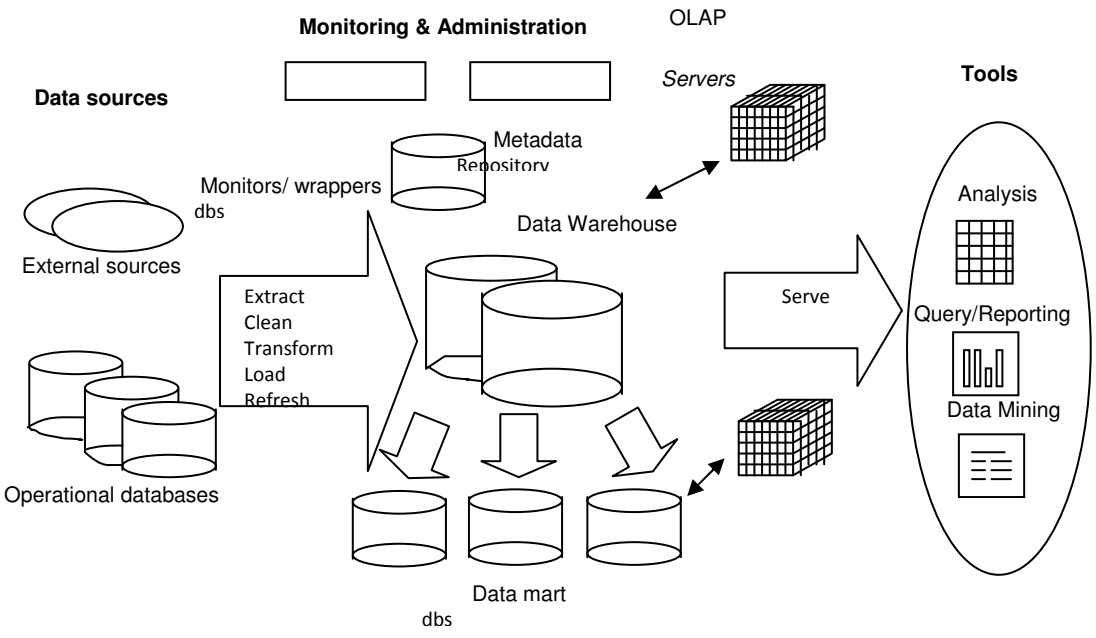
Data warehousing is defined as a process of centralized data management and retrieval. Data warehousing, like data mining, is a relatively new term although the concept itself has been around for years. Data warehousing represents an ideal vision of maintaining organizational data. For data warehouse centralization of data is needed to maximize user access and analysis. Data warehouses generalize and consolidate data in multidimensional space. The construction of data warehouses involves data cleaning, data integration and data transformation and can be viewed as important pre processes step for data mining

Data warehouses can also be referred as a data base that is maintained separately from an organization operational databases. A Data warehouse is a subject oriented , integrated time variant non volatile collection of data in support of management decision process. A data warehouse is organized around major subjects such as sales, customers, suppliers and product. Data Warehouse provide a simple and concise view around particular subject issue, a data warehouse is usually constructed by integrating multiple heterogeneous sources. Data cleaning and data integration techniques are applied to ensure consistency and encoding structures and so on .data Warehouse contains either implicitly and explicitly and element of time . It is always physically separate store of data transform from the application data found in the operational environment. Two operation that requires in data accessing : initial loading of data and access of data

OLAP in Data Warehouse

Data warehousing addresses a broad range of decision support requirements that are served by a variety of personal-productivity, query and reporting tools, as well as OLAP servers. OLAP servers fulfill a very wide set of user-driven functional requirements for reporting, Analysis, modeling and planning applications.
OLAP applications share a set of user and functional requirements that cannot be met by query or personal-productivity tools that work directly against the historical data maintained in the warehouse relational database. An OLAP server provides functionality and performance that leverages the data warehouse for reporting, analysis, modeling and planning requirements. These processes mandate that the organization look not only at past performance, but also more importantly, at the future performance of the business. It is essential to create operational scenarios that are shaped by the past yet also include planned and potential changes that will impact tomorrow's corporate performance. The combined analysis of historical data with future projections is critical to the success of today's corporation.
OLAP servers deliver warehouse applications such as performance reporting, sales forecasting, product line and customer profitability, sales analysis, marketing analysis, what-if analysis and manufacturing mix analysis — applications that require historical, projected and derived data. With OLAP servers' robust calculation engines, historical data is made vastly more useful by transforming it into derived and projected data. Users gain broader insights by combining standard access tools with a powerful analytic engine.
Requirements for the OLAP component of a data warehouse include:
• The ability to scale to large volumes of data and large numbers of concurrent users
• Consistent, fast query response times that allow for iterative speed-of-thought analysis
• Integrated metadata that seamlessly links the OLAP server and the data warehouse
Relational database
• The ability to automatically drill from summary and calculated data, which is managed by the OLAP server, to detail data stored in the data warehouse relational database
• A calculation engine that includes robust mathematical functions for computing derived data (aggregations, matrix calculations, cross-dimensional calculations, OLAP-aware formulas and procedural calculations)
• Seamless integration of historical, projected and derived data
• A multi-user read/write environment to support users what-if analysis, modeling and
Planning requirements

Possible Role of OLAP in Cartography
• The ability to be deployed quickly, adopted easily and maintained cost-effectively
• Robust data-access security and user management
• Availability of a wide variety of viewing and analysis tools to support different user communities
Cartography is used for designing a maps. For making maps Cartographers needs large amount of data. OLAP servers fulfilled very large amount of data requirements. Collected data is well managed by OLAP servers and cartographers use that data for making maps.



DATA MINING

Data mining is the process of discovering interesting knowledge such as patterns, associations, changes and significant structures from large amount of data stored in databases, data warehouses or other information repositories. it refers to the mining or discovery of new information in terms of patterns or rules from vast amounts of data. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified.

Data Mining is extracting the knowledge from large amount of data. Data Mining is a technique to analyze data from different sources and summarize that data into useful information. Data Mining is a process of extracting patterns or co-relations in large relational data bases . Data Mining is based on data, information and knowledge

DATA

Data is raw facts and figures .It has no significance beyond its existence. It can exist in any form usable or not. It does not have meaning of itself and it represent a facts or statement of event without relation to other things. Data is very useful for generating reports, graphs and statistics.

For e.g. student fill an admission form when they get admission in college. The form consists of raw facts about the student. These facts are student name, father name and address etc.

INFORMATION

The manipulated and processed form of data is called information. It is more meaningful than data. It is used for making decision. Information is data that are processed to be useful ,provides answers to who, what, where and whom questions .Information is in finished form.

For e.g. Data collected from census is used to generate different types of information. The Government can use it to determine the literacy rate in country. Government can use the information in important decision to improve literacy rate.

KNOWLEDGE

Knowledge is outcome of the information. It consists of data items and information organized and processed to convey understanding, experience, and expertise as they apply to the current problem or activity. Knowledge can be application of data and information in making a decision.

Data mining consists these major elements:

- Extract the data from data warehouse system.
- Transform, and load transaction data onto the data warehouse system.
- Store and manage the data in a multidimensional database system.
- Provide data access to business analysts and information technology professionals.
- Analyze the data by application software.
- Present the data in a useful format, such as a graph or table.

References

1. Raisz Eruein (1962): Principles of Cartography, Mc.Grawhill Book Co. Ing New York ,p-293
2. Buchanan R.O(1974): an Illustrated Dictionary of Geography,fep international ltd,P-38
3. Wrembel Robert & Koncilia Christian (2007), Data Warehouses and OLAP: Concepts, Architectures and Solutions
4 Jiawei Han & Micheline Kamber (2006), Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers.
5. Cheung DW, Han J, Ngv, WongCy (1996), "Maintenance of Discovered association rules in large databases: an Incremental updating technique" In Proceedings of the ACM SIGMOD international Conference on Management of data