TERM PAPER
DATABASE MANAGEMENT SYSTEMS

11 February 2022

DBMS Requirements for Geographic Information System (GIS)

Sahitya Charchit(19111050)
BME/6th Sem/B.Tech

---

1 Abstract

In geographic information systems (GIS) large amounts of data are stored and must be made available to multiple users. Database management systems (DBMS) were designed to facilitate storage and retrieval of large data collections. They include facilities to protect and secure data, enforce consistency of the data stored, and make data available to multiple users at the same time. These services are necessary for GIS, and GIS should therefore be built using database management systems. However, geographic information systems demand high performance and pose some very special requirements for database management. DBMS designed for commercial usage are not well suited for GIS because they cannot accommodate spatial data and cope with retrieval of map graphics. An overview of the architecture of a DBMS especially suited for spatial data handling is presented.

2 Introduction

GEOGRAPHIC INFORMATION SYSTEMS (GIS) must store large mounts of data and make them available on demand. Users have learned from their personal computer experience to demand nearly instantaneous responses even for relatively complex requests. Traditional solutions in which data are stored on disk or on magnetic tape and must be searched sequentially cannot respond fast enough to user queries and are no longer sufficient to accommodate frequent changes in the users needs. A modern GIS is expected to be able to integrate data for different topics and from different sources. The integration of multiple data sets, often visualized as multiple data layers, is expected to produce a synergistic effect and yield better information for decision making. Traditional file oriented storage cannot easily respond to this requirement either. Geographic information systems are comprised of a complex of several parts that interact. In order to build computerized GIS, we have to deal with organizational, software, and hardware problems. It must be noted that organizing the cooperation of different groups to collect data and to share the results is an especially difficult task, for which few guidelines and rules are available. Many projects fail not for technical

reasons, but for lack of organizational arrangements or because of a poor understanding of social or economic implications. Hardware problems are more easily resolved the components for storage and processing of very large amounts of data are available from various manufacturers. Prices are increasingly reasonable and the general trend is toward "zero cost hardware" (Dangermond and Morehouse, 1987). GIS software, on the other hand, is much more difficult to build than many had previously thought. The software system to manage GIS data must contain a module that provides database management system functionality. This paper deals primarily with this software component and the requirements placed on it by GIS applications.

Database management systems (DBMS) are appropriate tools for GIS. Fast access to spatial data out of a large data collection is difficult to achieve. Many current GIS store data as a collection of map sheets (or similar spatial partitions) which are then handled as units. This requires all users to understand their structure and hinders access by postal addresses or other logical concepts, for example. To achieve the desired "seamless" database where objects (i.e., map features) are not arbitrarily divided by map boundaries and where users can freely move or zoom over the map, special methods and optimizations are necessary DBMS software provides the services needed to integrate and protect the data. But, the conventional DBMS does not deliver the performance and cannot retrieve map data quickly enough for interactive work. Not all GIS software packages currently on the market contain a DBMS or include all the services necessary for data protection. In this paper, we detail these necessary DBMS services and show in an architectural overview how they interact. We use modern software engineering concepts to organize the discussion. Particular attention is given to the integration of database management systems with other software specifically written for spatial data processing. Emphasis is placed on data storage and retrieval functions, including the protection of the data in a GIS. Equally important problems of adequate modeling of reality and the data model support necessary for GIS are excluded and treated elsewhere in order to conserve space (Egenhofer and Frank, 1988). The discussion of access methods and, especially query languages is therefore intentionally limited. Many of the ideas reported here are based on experience with the PANDA database management system (Frank, 1982a, 1984b, 1986a; Egenhofer and Frank, 1987a). We identify methods successfully implemented, and include a critique of methods which have not worked as well and will be replaced in the future.

3 Spatial Information Systems

The use of computers for "batch" processing, where all the input data are collected and an output with the result is delivered later, has been largely replaced by interactive information systems, where the system maintains a collection of data which is then interrogated by users as they need the information. In general terms, an information system contains an image or model of reality, which we can use to make decisions and need not reinvestigate the facts each time. This is extremely important in all situations where data collection is expensive, cumbersome, or slow, and is one of the major forces behind GIS: substantial savings by sharing the cost of data collection and at the same time improved usage of the

data and higher quality information output is expected. Geographic information systems deal with data related to location in real world space - here referred to as spatial data. Many operations of government at all levels, as well as planning and research, exploit data which have a spatial component. Such systems are referred to by various other names (e.g., land information system, AM/FM, multi-purpose cadastre). We will concentrate on general aspects of systems dealing with spatial data referred to as "spatial information systems", without consideration of differences between systems designed for specific tasks. We will concentrate on systems which store data with an exact reference to location and which describe geometry using points and vectors. This is not to exclude systems of other types there are obvious advantages in the use of raster operations for certain tasks, but they seem to have substantially different requirements for data storage and warrant a separate discussion. A GIS is a model of reality and not just a repository of cartographic data necessary to draw maps. Methods to represent complex aspects of reality in a computer system therefore become important. Only if the structure of reality is appropriately modeled in the data stored can we expect that the combination of multiple data sources and the extraction of complex information will produce results that are meaningful. In such situations we encounter relations between the data elements, e.g., a building is at the same time related to a lot on which it is built, to a street it is on, and to persons who are living in it. A method to store and retrieve the data, using and preserving these multiple relations, is necessary.