OVERVIEW OF GIS: FUNDAMENTALS, DESIGN CONCEPTS AND FUNCTIONALITIES

Introduction

Archeological evidences have unearthed the fact that the history of map making exists since ages. Human have broadened, over the years, their understanding about size, shape and processes associated with Earth which in turn contributed in making sophisticated and accurate representation of the Globe and its phenomena. Advancements in information and communication technology, space technology have stimulated the growth of geography oriented information science/system, in short GIS, as which helps in modeling and representing earth's phenomena in an efficient way. GIS has started only after 1950s, in real sense. Many people who were involved in the development of very first versions of GIS are still alive today. In order to have a proper record of these historical developments an initiative named "The GIS History Project" was launched at a workshop held in Santa Barbara, California, in September of 1996 (Anonymous, 1998). The major objective of this project was "to uncover and document the technical, social, and institutional factors that shaped the early development of GIS, and thus have had a major influence on GIS as we know it today" (David et al. 1997). Another interesting source of detailed information about the history of map making, GIS and respective time line of developments is available in the websites of GIS@Development (2004 a, b).

David et al (1997) points to a special issue of the American Cartographer in 1988 entitled "Reflections on a Revolution: The Transition from Analogue to Digital Representations of Space 1958-1988", edited by Roger Tomlinson and Barbara Petchenik, which refers to many papers that review the history of GIS developments. This paper gives interesting references to the initial efforts of the Canada Geographic Information Systems (CGIS) (Tomlinson, 1988), early days of Environmental Systems Research Institute (Dangermond and Smith's 1988), experiments at the Harvard Laboratory for Computer Graphics and Spatial Analysis (Chrisman, 1988), and activities at the Experimental Cartographic Unit in the United Kingdom (Rhind, 1988). Work of Coppock and Rhind (1991) also provides a good coverage of institutions and individuals involved in the early developments of GIS. 1991 issue of the cartographic journal also had a special issue entitled "British Cartography 1987-1991: An Overview", but it stresses mainly the developments in cartography. Roger Tomlinson is considered as the father of GIS because of his initial contribution in making Canada GIS (Anonymous, 2006). According to him GIS is not a field by itself but rather the common ground between information processing and the many fields utilizing spatial analysis techniques.

Developments during 1960s were snagged with many technical problems like converting analogue map into computer compatible form, format for storage, display techniques and more. 1970s saw interests and participation of universities and need for topology (spatial relation) was felt. 1980s contributed for the major growth of GIS due to advancements in personal computers, cheaper hardware and efficient software. This lead to new initiatives and progress in spatial modeling, data structure issues and remote sensing linkages. This period also saw successful and reliable systems and government interests and investments. Further major leap was seen during 1990s as more and more PCs, object oriented architectures, networks, internet and mobile technology. Recently internet has become a major medium of communication and data dissemination. Over the years, Geographical information System has evolved into Geographical information Science and at present it is a billion dollar market because it is leading to Geographical Information Services. Many software companies have also evolved with GIS and they have been playing a crucial role in making GIS a commercially viable solution providing mechanism for various domains of human activity.

Elements of a GIS

The GIS comprises of four elements. They are hardware, software, dataware, and humanware. The following table gives details of different elements: -

Table 2.1: The elements of GIS

Elements of GIS	Details
Hardware	Type of Computer Platforms Modern Personnel Computers High performance workstations Minicomputers Mainframe computers Input Devices Scanners Digitizers Tape drivers CDs Keyboard Graphic Monitor Output Devices Plotters Printers
Software	Input Modules Editing Manipulation/ Analysis Modules Modeling Capability
Dataware	Attribute Data Spatial Data Remote Sensing Data Global Database
Humanware	Trained People responsible for database preparation, modeling and implementing using GIS

Geographic Phenomena, types and its representation

GIS is a computer based tool which helps in storing, retrieving, manipulating, analyzing and producing maps about information related to geographic phenomena with the help of a human expert. Geographic phenomenon refers to a process associated with the Earth. In order to represent Geographic phenomenon in GIS, its position, its property and time of occurrence must be known so that one can retrieve information about ***what*** has happened, ***where*** it has occurred and ***when*** it has happened. In simple words, Geographic phenomenon is nothing but what we see or observer about our earth, e.g.., observation of daily temperature over a city, weather pattern, crop cycle, human settlement mapping etc,. Some of the Geographic phenomena are easy to explain verbally, some are easy to represent through drawings. Some are easy to draw, like buildings and some are difficult to draw like temperature or elevation. So in order to represent such diverse phenomena some framework needs to be followed so that everybody represents the same thing in a same manner and hence it will be easy to understand by all globally. In this regard the geographic phenomena is divided into 2 major groups; ***Objects*** *and* ***Fields*** (Rolf de By et al., 2004).

Objects refer to the phenomena which are having definable crisp boundaries i.e., discrete existence. Fields refers to the phenomena which do not have sharp boundary, but rather fuzzy in their presence and occurs at all places i.e., continuous existence. Field can be again categorized into 2 more category as ***Continuous field*** *and* ***Discrete Field*** according to their fuzziness in representation. Examples of Object like phenomena are rivers, buildings, volcanoes, islands, etc,. Examples of *discrete field* are Landuse Map, Soil Map, Geology map. Soil & Geology occurs every where and we cannot exactly see their starting and ending points on the ground, unless until some sharp natural barrier occurs. But for representing them we have to consider some probable end point and we introduce artificial discreteness, hence it is called discrete fields. Examples of *continuous fields* are temperature, elevation, humidity etc,. It is generally observed that man made things are generally objects and natural things are fields, with exceptions like river, volcano, islands are natural things but they are objects as it does not occur everywhere. So we can say that “all manmade things are objects and all objects are not man made”.

Point, Line and Polygon are the basic building blocks for representing any phenomena in the computer. But representation of any phenomena in computer depends upon the scale of mapping because at 1:1Million scale towns will become points, but at 1:10,000 scale they become polygons. Generally landuse, administrative boundary and thematic maps are mapped through polygon. Roads, rivers, pipe lines, electricity line are mapped as Lines. Village locations, Utility locations, Field observations are represented as Points.

Apart from understanding their locational and conformal property, one must also understand their attribute properties as every element, whether it is a point or line or polygon, needs to have some description about what it is representing. The domain of attribute data is classified into 4 categories as ***Nominal, Ordinal, Rational and Interval***. Qualitative data are represented through Nominal data, which cannot provide any quantitative meaning. Nominal refers to the data which are generally used for identification purposes. For example, name of a person, name of the road, telephone number, house number, etc, are nominal kind of data. Ordinal data refers to ordered data, in which we can infer order of importance, for example, if we rank the people as per their exam score then it is an ordinal data. Ratio data is the actual fact/data measured on the ground, which is a quantitative measurement which has actual origin at zero and which has same meaning all over; for example 0mm is same as 0Km, hence distance or length is a "Ratio" data. The name "ratio" do not make any sense about ratioing, actually the name "rational" has become "ratio". Interval data is a data for which actual origin differs at different places and in which zero do mean same thing, for example, 0 degree Kelvin is different from 0 degree Celsius. Hence temperature is an interval data. Also when we measure earthquake in Richter scale the energy difference of two earthquakes of magnitude 5.1 and 5.2 is different than the energy released by 7.1 and 7.2. Though the difference in magnitude is 0.1 in both the cases the actual quantitative meaning is completely different, hence earthquake measurement is an Interval data. User must also be aware that one has to be clear about these "types of data" while doing spatial operations as all the operations/analysis are not possible with all the data type.

Spatial Data Structure

There are two broad types of data structure generally adopted to represent all geographic phenomena in computer under GIS. The types are Vector and Raster. Vector in mathematical sense reveals a 'quantity and direction'. But in GIS, Vector is used for referring to the basic mathematical elements or building blocks; point, line and polygon. Any map which is made using these building blocks are called Vector map. It is possible to represent roads, buildings, administrative units, landuse, geology, and more using vector concepts, but there are many phenomena which cannot be represented in vector. For example how can we record temperature or elevation or humidity or remote sensing reflectance value. GIS has been molded to represent this type of phenomena using ***tessellation*** concept. Tessellation is nothing but dividing the space/area into uniform grid and finding the dominant phenomenon occurring within each grid. The space can be divided using uniform grid also called regular grids like squares or rectangles or triangles pentagons or hexagon. Normally square grid is adopted in the tessellation process due to its simplicity and many hidden advantages like it is easy to find the location of any grid if we know the origin grid's location and grid size, easy to store, retrieve and analyze. If the size of the grid is smaller then it occupies lot of storage space and if we increase the size of the grid then we loose some information variability within that grid. So one has to come to a compromise is selecting size of grid, which is called ***resolution***.

In general the tessellated space looks like a two dimensional matrix of grids/cells, which is called ***raster.*** Raster has lot of advantage over vector especially for spatial analysis and modeling, as we can deal with any part of the study area as every pixel is explicitly represented, but in vector only we represent the boundary of the phenomena. So we can do only logical operations like AND, OR, NOT with vector layers and we cannot do arithmetic operations like addition, subtraction, multiplication or division. But in raster we can do all kind of operations like comparison, logical operations, arithmetic operations and trigonometric operations and dynamic simulations. Remote Sensing data is a raster data and hence we can directly adopt remote sensing into our GIS models, if our data is in raster structure. Vector data is good for printing accurate representation.

If we adopt the regular tessellation in the case of phenomenon which occur over very large spatial area, then we will end up with having the same cell value occurring repeatedly and hence occupying lot of storage space. In order to avoid such redundant storage, a new approach was adopted which is called irregular tessellation. In irregular tessellation the space is divided only if it has diverse features. By this approach lot of storage space can be saved. One of such irregular tessellation is Quad-tree data storage structure. There are many data storage formats which adopt various algorithms and some of them are even proprietary. Main objective behind such algorithms are lossless compression, faster retrieval and efficient manipulation. In order to represent elevation a new kind of storage technique using an approach called TIN, Triangulated Irregular Network, has been used. TIN is a irregular tessellation in which a triangle is a basic building block, which made using three input reference height points. From the network of triangles it would be easy to calculate elevation, slope and aspect for any point in the study area.

Spatial Layers or Geodatabase

The common requirement to access data on the basis of one or more classes has resulted in several GIS employing organizational schemes in which all data of a particular level of classification, such as roads, rivers or vegetation types are grouped into so called layers or maps. The concept of layers is to be found in both vector and raster models. The layers can be combined with each other in various ways to create new layers that are a function of the individual ones. The characteristic of each layer within a layer-based GIS is that all locations with each layer may be said to belong to a single aerial region or cell, whether it be a polygon bounded by lines in vector system, or a grid cell in a raster system. Such database bundle is called Geodatabase. But it is possible for each region to have multiple attributes. The following figure shows layers and coverage concept in GIS (Figure 1).

Planimetric Requirements

All the maps are used for some kind of measurement purposes in real life applications. Hence the measurement made on the maps have to be accurate. Since our earth is a 3D body and layers have to be in 2D mode, there is a need for a mechanism which can help in achieving this transition between 3D to 2D (Figure 2). This mechanism is called Projection. All the geospatial database generated under GIS must be in a planimetric coordinate system, means it must be represented in a 2-dimensional reference frame so that we can find out area and length correctly. There are many ways by which a 3D globe can be converted into 2D. Cylinder, Cone and Plane are simple mathematical figures which can be utilized for this conversion. Conversion is done by wrapping the earth with the paper of these shapes and then making the imprint of global feature on the paper and then unwrap the paper. Converting from 3D to 2D introduces some loss in either area or shape or direction measurement. Based on the property it preserves the projection receives its name like equal area projection, conformal projection etc.

Another major hurdle in the projection process is that the earth does not have smooth surface. It has gravitational undulations which is visible from mean sea level plots. Because of the non-linear complex undulations it is very difficult to replicate exact position and height of a location accurately on a map. So assumptions about the shape of our earth have to be made as either ellipsoid or spheroid so that mathematically it would lead to projection. Every country has adopted their local mean sea level surface reference for height measurement, which is called vertical datum and adopted a mathematical surface (ellipsoid) that fits better for their portion of the globe which is called horizontal datum. If we want to convert from one projection to other project we have to have 7 parameters related to translation effect (dx, dy, dz), rotation effect (rx, ry, rz) and a scale factor.

Spatial Analysis

Whether it is for natural resources or sustainable development, or natural disaster management, selecting best site for waste disposal, optimum route alignment or local problems have a geographical component; GIS has power to create maps, integrate information, visualize scenarios, solve complicated problems, present powerful ideas and develop effective solutions like never before. In brief it can be said as supporting tool for decision making process. Only when all the maps are projected to an agreed same coordinate system, it would be possible to do spatial analysis.

GIS is used to perform a variety of spatial analysis, including overlaying combinations of features and recording resultant conditions, analyzing flows or other characteristics of networks; proximity analysis (*i.e.* buffet zoning) and defining districts in terms of spatial criteria. GIS can interrogate geographic features and retrieve associated attribute information, called identification. It can generate new set of maps by query and analysis. It also evolves new information by spatial operations. Following are the analytical procedures applied with a GIS. GIS operational procedure and analytical tasks that are particularly useful for spatial analysis include:

- ❖ Single layer operations
- ❖ Multi layer operations
- ❖ Spatial & non-spatial query
- ❖ Measurement operations
- ❖ Neighborhood Analysis
- ❖ Network analysis
- ❖ Surface analysis
- ❖ Predictive & Simulation analysis

There are many applications of Geoinformatics *viz*: facility management, planning, environmental monitoring, population census analysis, insurance assessment and health service provision, hazard mapping and many other applications. The following list shows few applications in natural resource management;

- Agricultural development
- Land evaluation analysis
- Change detection analysis
- Analysis of deforestation and associated environmental hazards
- Monitoring vegetation health
- Crop acreage and production estimation
- Wasteland mapping
- Soil resources mapping
- Groundwater potential mapping
- Geological and mineral exploration
- Snow-melt run-off forecasting
- Monitoring forest fire
- Monitoring ocean productivity *etc*.

Future Trend in GIS

With the rise of World Wide Web (WWW), new Internet protocol such as the hypertext Transfer Protocol (HTTP), new markup languages like HTML, DHTML, XML, GML, as well as easy to use interfaces (browsers), tools (Flash) and languages (.NET, Java, scripts), the internet becomes a powerful media for future. Internet applications and Internet GIS are the driving forces of future. Google Earth and Wikimapia are wonderful examples of such internet based geo-information services. Furthermore, with the advancements in mobile communication technology the world is moving towards having GIS in the mobile phones, by which one can find out nearest ATMs, nearest theatres, nearest hospitals, optimal route between source & destination etc,. Such applications are called Location Based Services (LBS) (Figure 3), which will help in better e-governance and also during disaster relief operations.

References:

Anonymous, 1998. *The GIS History Project*. Web source: http://www.ncgia.buffalo.edu/gishist/

Anonymous, 2006. *GIS World Interview*: Roger Tomlinson. Web source: http://www.geoplace.com/gw/1996/0496/0496feat2.asp

Chrisman, N. 1988. *The risks of software innovation: a case study of the Harvard Lab*. The American Cartographer, 15 (3), 291-300.

Coppock, J. T.; and Rhind, D.W. 1991. *The History of GIS*. In Maguire, D. J.; Goodchild, M. F.; and Rhind, D. W. (editors). *Geographical Information Systems: Principles and Applications*. London: Longmans Publishers, 1, 21-43

Dangermond, J.; and Smith, L. K. 1988. *Geographic information systems and the revolution in cartography*: *The nature of the role played by a commercial organization*. The American Cartographer, 15 (3), 301.

David M. Mark; Nicholas Chrisman; Andrew U. Frank; Patrick H. McHaffie; John Pickles. 1997. *The GIS History Project*. With contributions from Michael Curry; Jon Goss; Francis Harvey; Ken Hillis; Roger Miller, Eric Sheppard; Dalia Varanka. Web Source: http://www.ncgia.buffalo.edu/gishist/bar_harbor.html

GIS@Development (2004 a). *Mapping GIS Milestones*. Websource: http://www.gisdevelopment.net/history/1950-1960.htm

GIS@Development (2004 b). Mapping the Maps. Websource: http://www.gisdevelopment.net/history/pre200ad.htm

Rolf A. de By; Richard A. Knippers; Michael J. C. Weir; Yola Georgiadou; Menno-Jan Kraak; Cees J. van Westen; Yuxian Sun. 2004. *Principles of Geographic Information System: An Introductory Textbook*. ITC Educational Textbook Series. International Institute for Geo-Information Science and Earth Observation. Enschede, The Netherlands.

Rhind, D. 1988. *Personality as a factor in the development of a discipline: The example of computer-assisted cartography*. The American Cartographer, 15 (3), 277-289.

Tomlinson, R. F. 1988. *The impact of the transition from analogue to digital cartographic representation*. The American Cartographer, 15 (3), 249-261.

Valerie Bennett and Andrew Capella. 2006. *Location-based services: wherever you are, wherever you go, get the information you want to know*. Web source: http://www-128.ibm.com/developerworks/ibm/library/i-lbs/

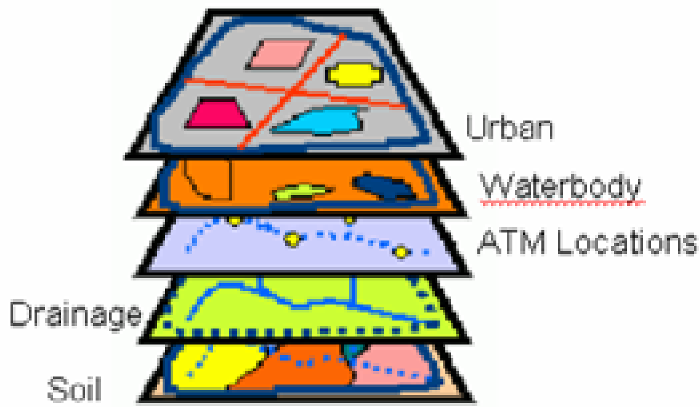

Figure 1: Spatial Layers / Geodatabase

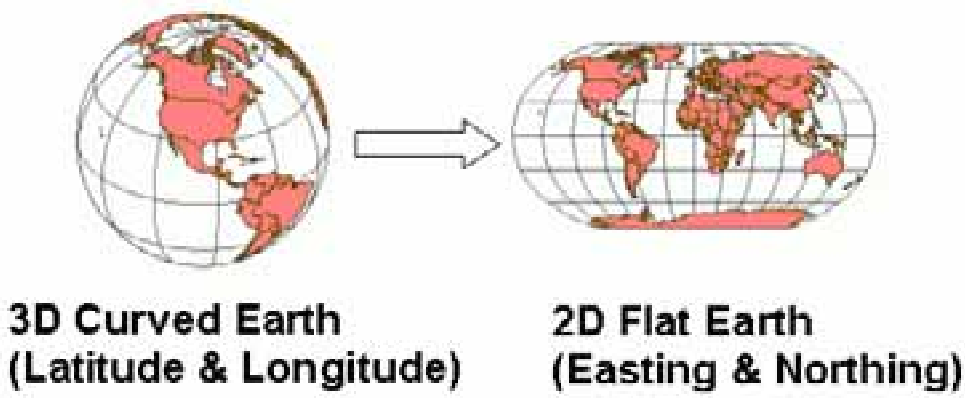

Figure 2: Map Projection

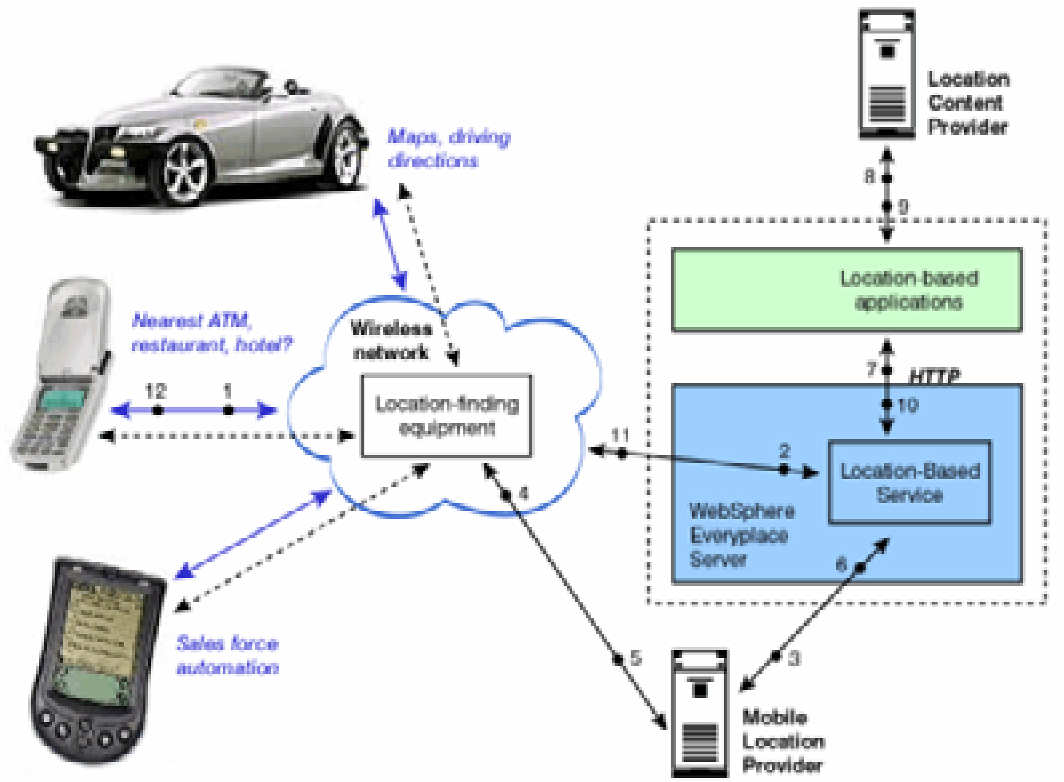

Figure 3: Concept of Location Based Service (LBS)

(Source: Valerie Bennett and Andrew Capella, 2006)