A

Seminar report

On

Geographic Information System (GIS)

Submitted in partial fulfillment of the requirement for the award of degree of CSE

SUBMITTED TO:
www.studymafia.org

SUBMITTED BY:
www.studymafia.org

Acknowledgement

I would like to thank respected Mr…….. and Mr. ……..for giving me such a wonderful opportunity to expand my knowledge for my own branch and giving me guidelines to present a seminar report. It helped me a lot to realize of what we study for.

Secondly, I would like to thank my parents who patiently helped me as i went through my work and helped to modify and eliminate some of the irrelevant or un-necessary stuffs.

Thirdly, I would like to thank my friends who helped me to make my work more organized and well-stacked till the end.

Next, I would thank Microsoft for developing such a wonderful tool like MS Word. It helped my work a lot to remain error-free.

Last but clearly not the least, I would thank The Almighty for giving me strength to complete my report on time.

Preface

I have made this report file on the topic **Geographic Information System** (**GIS**); I have tried my best to elucidate all the relevant detail to the topic to be included in the report. While in the beginning I have tried to give a general view about this topic.

My efforts and wholehearted co-corporation of each and everyone has ended on a successful note. I express my sincere gratitude to …………..who assisting me throughout the preparation of this topic. I thank him for providing me the reinforcement, confidence and most importantly the track for the topic whenever I needed it.

What Is GIS?

A **geographic information system** (**GIS**) is a system designed to capture, store, manipulate, analyze, manage, and present spatial or geographic data. The acronym GIS is sometimes used for geographic information science (GIScience) to refer to the academic discipline that studies geographic information systems and is a large domain within the broader academic discipline of geoinformatics. What goes beyond a GIS is a spatial data infrastructure, a concept that has no such restrictive boundaries.

In general, the term describes any information system that integrates, stores, edits, analyzes, shares, and displays geographic information. GIS applications are tools that allow users to create interactive queries (user-created searches), analyze spatial information, edit data in maps, and present the results of all these operations. Geographic information science is the science underlying geographic concepts, applications, and systems.

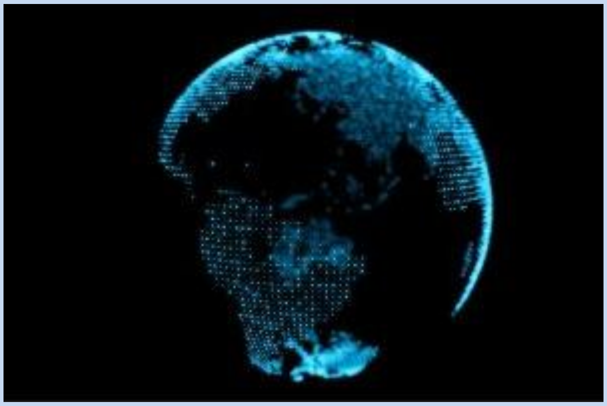
GIS can refer to a number of different technologies, processes, and methods. It is attached to many operations and has many applications related to engineering, planning, management, transport/logistics, insurance, telecommunications, and business. For that reason, GIS and location intelligence applications can be the foundation for many location-enabled services that rely on analysis and visualization.

GIS can relate unrelated information by using location as the key index variable. Locations or extents in the Earth space–time may be recorded as dates/times of occurrence, and x, y, and z coordinates representing, longitude, latitude, and elevation, respectively. All Earth-based spatial–temporal location and extent references should be relatable to one another and ultimately to a "real" physical location or extent. This key characteristic of GIS has begun to open new avenues of scientific inquiry.

History of GIS

There is some debate over when true GIS really started, thanks to the disparate technologies that came together to give us GIS as we understand it today, but effectively it has been around since the early 1960s. Advocates claim that it was truly born in 1962 with the first CLI conference (Canadian Land Inventory) that set out to produce masses of data of maps of Canada covering a large number of potential uses and data sets. The conference produced these maps using the old methods but it was first theorised here that in future, such data could be produced using the developing computing technology as data got bigger and potential to explore it became more and more complex.

The next few decades saw the technology strictly limited to those who had the resources: the hardware and the software were both expenses that had to be justified for the business or the industry, hence that fossil fuel companies represent some of the earliest groups to take up the technology. The 1980s when home computing was increasingly the norm and IT technology began to spread its wings, the business ESRI was formed. Today they are famous for the popular package ARCGIS; the second most popular package in the world today is MapInfo and the corporation was also formed in the same decade. But it would be some time before their software - as popular and as useful as it was - would be anything more than a niche interest.



GIS would not begin to grow until the birth of the internet; even when uptake was relatively slow, those for whom the internet was proving a useful business tool finally saw the potential for the affordability of GIS. Historians of technology believe that the increasing portable nature of the internet of the last ten years has really aided GIS technology to experience its growth in the same period. Why? Because data collection and transmission became so much easier as well as less costly than it had been before. Suddenly, GIS was no longer just available to the big companies who could afford to invest in the technology to gather and manipulate it, but to everyone. Charities could collect data relatively cheaply, as could conservation organisations and land authorities, town councils and urban planners. None of these groups need rely on older and slower methods of data collection when they could do so far more cheaply. Now, non-profits and even the general public had access to the same data previously only afforded to private industry and governments. A great example of this is the UK Environment Agency website that has collated data from a number of government agencies and NGOs and made it freely available for public use.

Today, most organisations that collect and use data make information available to almost anyone. Though confidential data may require that the user register on a site and sign a non-disclosure agreement, the nature of Big Data coupled with the exponential growth of Web 2.0 and how that data is used means that anyone can produce useful maps. In the early days, GIS was concerned with how the world looks but less concerned with how it works. This is a result of the growth of technology, decreasing price of the technology and the fact that we can do more things with more data.

Functions of a Geographic Information System

Data Capture

Data input to a geographical information system can be best broken into three categories: entering the spatial data, entering non-spatial data, and linking the two together. Entering the spatial data can be done numerous ways. Spatial data can be acquired from existing data in digital or paper form, or it can be collected from scratch.

Finding already mapped data in a paper format for an area can be accomplished in several ways. Paper map collections can usually be found within large libraries or universities. Libraries often times will also contain books with maps for international and domestic data. Another good resource for geographic data is local, state, or national government. Many countries have a wide range of data available at their country mapping agencies. If the data is to be more localized to a specific area, the local governments such as planning departments should contain the information. In addition, there are many commercial mapping companies that will sell data world wide for certain countries. The Internet is a good resource to search for data either from a vendor or a site offering free data. (Clarke, 2001)

There are two methods of getting paper maps into the computer: digitizing and scanning. Geocoding is the term used for the conversion of analog spatial information into digital form. Digitizing on a tablet captures map data by tracing lines by hand, using a cursor and an electronically sensitive tablet, resulting in a string of points with (x,y) values. Scanning involves placing a map on a glass plate while a light beam passes over it, measuring the reflected light intensity. The result is a grid of pixels. Image size and resolution are important to scanning. Small features on the map can drop out if the pixels are too big. (Clarke, 2001)

Finding data via the Internet can be done by performing a basic search. There are several sources for downloadable data such as:

- The Geography Network
- Data Depot
- Spatial Information Clearinghouse

Finally, if the data available does not meet the needs of the user, it can created by use of GPS, Remote Sensing, Aerial Photography, and field collection techniques.

Projection and Rectification

In order for the spatial data of a 3-dimensional earth to be represented in a 2-dimensional GIS, the data must make use of one of the various projection methods (See Remote Sensing Section for further detail on projections). Because different projections place the same special entities on different coordinates on the flat surface, it is vital that a projection be set for the specific data set being used. One of the main features of a GIS is the ability to overlap different data layers for better analysis. These different layers must have the same projection, datum, and reference ellipsoid so that all coordinates are lined up correctly.

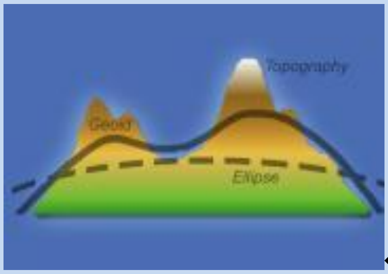


Figure 1, Reference Ellipsoid and Geoid. (SIC, 2002)

Data Modeling

Spatial modeling represents the structure and distribution of features in geographical space. In order to model spatial processes, the interaction between these features must be considered. There are several types of spatial data models including: vector, raster, surface, and network (Burrough, 1998).

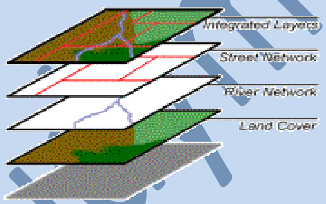


Figure 2, Integrated Layers of GIS Model (SIC, 2002)

Vector Data Model

The vector data model is a method of storing and representing data on an X,Y Cartesian plane. A coordinate and an equation defining the curvature of each feature is stored for both the beginning and the end point of each feature. The building block of the vector structure is the point; lines and areas are composed of a series of points in a specific order that gives the object direction (Clarke, 2001). The attribute data in the vector model is stored in a separate table that can be linked to the map. Because every item on the map has its own separate attribute data, analysis can be very easy. For example, if a vector road network is being used to analyze the amount of carbon monoxide produced by cars per year in both rural and urban communities, each road would be capable of having separate attributes, thus allowing the GIS user to view or select each road and access information associated with just that road.

Vector data entities in a GIS hold individual values, for example, if two lines overlap, unique values are recorded for each line in the database (spaghetti model). Selecting an appropriate number of points is another consideration to be made with vector data; if too few points are chosen, the shape and properties of the entity will be compromised and if too many points are used, duplicated information can be stored resulting in data overload (Burrough, 1998)

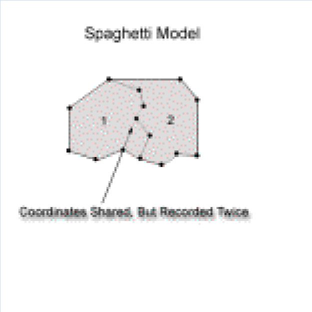


Figure 3, Vector Spaghetti Model (SIC, 2002)

Raster Data Model

The raster data model uses a grid composed of rows and columns to display map entities. Each cell in the grid is equivalent to one map unit or one pixel. Spatial resolution determines the precision of spatial representation by raster data. The smaller the size of the pixel, the higher the resolution and the better the precision of spatial representation (Lo, 2002). An entity code is assigned to each cell that is connected to a separate attribute table, which provides information to the user as to what entity is present in what cell.

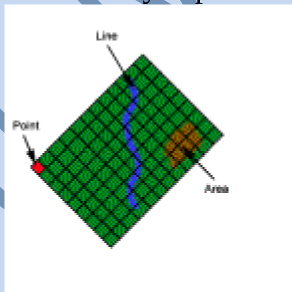


Figure 4, Raster Representation(SIC, 2002)

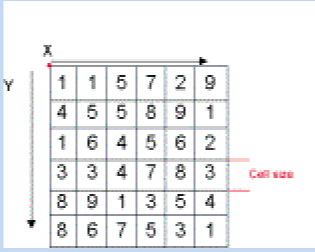


Figure 5, Raster Attribute Table(SIC, 2002)

The term raster data when applied to GIS and mapping includes scanned monochrome and color printing separates, scanned black and white and color aerial photographs, remote sensing images, digital elevation models, as well as thematic spatial data created by manual and computer-based methods (Lo, 2002). These methods of storing one or more values for each grid location in the data drastically increase the file size. Several methods have been developed to compact the size of raster files. The first is *run length encoding* which reduces data on a row-by-row basis. If an entity occupies a large number of cellsin a row, a single value is stored representing the object followed by the number of cells in that row, rather than recording each individual value. Another compaction technique is called the *quadtree data model*. In this model, instead of dividing the entire area into cells of equal size, only areas with specific details are broken down into smaller cells. For example, if a land-use map had only one land use type, one cell would represent the entire area. If there were 4 classes, 4 cells would be used, and quadrant that had more than one land use type would be broken down until it only contained one type (Lo, 2002).

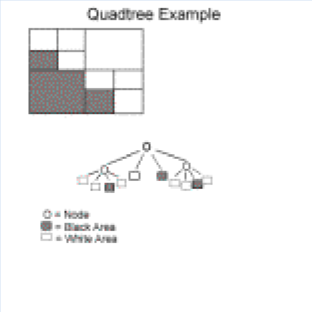


Figure 6, Quadtree Compaction (SIC, 2002)

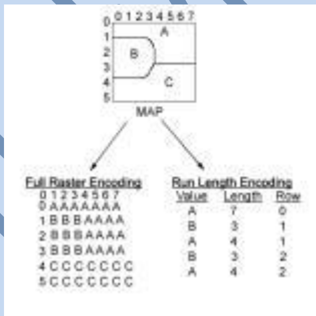


Figure 7, Run Length Encoding(SIC, 2002)

The raster data model represents spatial phenomenon such as topography, land use cover, and air quality as categorical or continuous surfaces. This makes raster-based methods particularly suitable for spatial modeling that involves multiple surface data sets. However, this method is not suitable for applications that rely on individual spatial features represented by points, lines, and polygons (Lo, 2002).

Tabular Data

Tabular data, also called attribute or descriptive data, is one of the most important elements in a GIS. It is statistical, numerical, or characteristic information that can be attributed to spatial features. Similar to spatial data the tabular data is stored by the GIS software in a method that allows it to be accessed and viewed, usually in a relational database format. Depending on the application, attributes that may be useful to assign to a feature would be population of an area, traffic measurement of a road, or types of landmines in a particular area. The GIS software allows the attribute data to be linked to the spatial data in such a way that it gives the attributes a location. A GIS package knows a specific location geographically from the storage of spatial data. By linking attribute data to the spatial data, the GIS package knows some of the characteristics of a feature in the spatial data set.

Two or more tabular databases can be linked when there is a common data filed. This allows the GIS to become a powerful spatial analysis tool. A GIS user, after integrating both spatial and attribute data, has the capability to learn a great deal about the defined study area.

Components of a GIS

A GIS can be divided into five components: People, Data, Hardware, Software, and Procedures. All of these components need to be in balance for the system to be successful. No one part can run without the other.

People

The people are the component who actually makes the GIS work. They include a plethora of positions including GIS managers, database administrators, application specialists, systems analysts, and programmers. They are responsible for maintenance of the geographic database and provide technical support. People also need to be educated to make decisions on what type of system to use. People associated with a GIS can be categorized into: viewers, general users, and GIS specialists.

- Viewers are the public at large whose only need is to browse a geographic database for referential material. These constitute the largest class of users.

- General Users are people who use GIS to conducting business, performing professional services, and making decisions. They include facility managers, resource managers, planners, scientists, engineers, lawyers, business entrepreneurs, etc.

- GIS specialists are the people who make the GIS work. They include GIS managers, database administrators, application specialists, systems analysts, and programmers. They are responsible for the maintenance of the geographic database and the provision of technical support to the other two classes of users. (Lo, 2002)

Procedures

Procedures include how the data will be retrieved, input into the system, stored, managed, transformed, analyzed, and finally presented in a final output. The procedures are the steps taken to answer the question needs to be resolved. The ability of a GIS to perform spatial analysis and answer these questions is what differentiates this type of system from any other information systems.

The transformation processes includes such tasks as adjusting the coordinate system, setting a projection, correcting any digitized errors in a data set, and converting data from vector to raster or raster to vector. (Carver, 1998)

Hardware

Hardware consists of the technical equipment needed to run a GIS including a computer system with enough power to run the software, enough memory to store large amounts of data, and input and output devices such as scanners, digitizers, GPS data loggers, media disks, and printers. (Carver, 1998)

Software

There are many different GIS software packages available today. All packages must be capable of data input, storage, management, transformation, analysis, and output, but the appearance, methods, resources, and ease of use of the various systems may be very different. Today's software packages are capable of allowing both graphical and descriptive data to be stored in a single database, known as the object-relational model. Before this innovation, the geo-relational model was used. In this model, graphical and descriptive data sets were handled separately. The modern packages usually come with a set of tools that can be customized to the users needs (Lo, 2002).

Data

Perhaps the most time consuming and costly aspect of initiating a GIS is creating a database. There are several things to consider before acquiring geographic data. It is crucial to check the quality of the data before obtaining it. Errors in the data set can add many unpleasant and costly hours to implementing a GIS and the results and conclusions of the GIS analysis most likely will be wrong. Several guidelines to look at include:

- Lineage – This is a description of the source material from which the data were derived, and the methods of derivation, including all transformations involved in producing the final digital files. This should include all dates of the source material and updates and changes made to it. (Guptill, 1995)

- Positional Accuracy – This is the closeness of an entity in an appropriate coordinate system to that entity's true position in the system. The positional accuracy includes measures of the horizontal and vertical accuracy of the features in the data set. (Guptill, 1995)

- Attribute Accuracy – An attribute is a fact about some location, set of locations, or features on the surface of the earth. This information often includes measurements of some sort, such as temperature or elevation or a label of a place name. The source of error usually lies within the collection of these facts. It is vital to the analysis aspects of a GIS that this information be accurate.

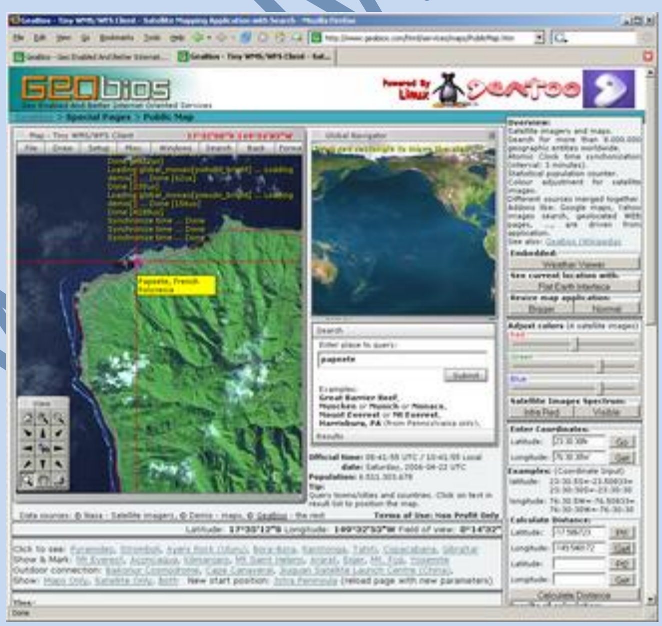
- Logical Consistency - Deals with the logical rules of structure and attribute rules for spatial data and describes the compatibility of a datum with other data in a data set. There are several different mathematical theories and models used to test logical consistency such as metric and incidence tests, topological and order related tests. These consistency checks should be run at different stages in the handling of spatial data. (Guptill, 1995)

- Completeness – This is a check to see if relevant data is missing with regards to the features and the attributes. This could deal with either omission errors or spatial rules such as minimum width or area that may limit the information. (Guptill, 1995) (Chrisman,1999).

Applications Of GIS

The implementation of a GIS is often driven by jurisdictional (such as a city), purpose, or application requirements. Generally, a GIS implementation may be custom-designed for an organization. Hence, a GIS deployment developed for an application, jurisdiction, enterprise, or purpose may not be necessarily interoperable or compatible with a GIS that has been developed for some other application, jurisdiction, enterprise, or purpose.

GIS provides, for every kind of location-based organization, a platform to update geographical data without wasting time to visit the field and update a database manually. GIS when integrated with other powerful enterprise solutions like SAP and the Wolfram Language helps creating powerful decision support system at enterprise level.

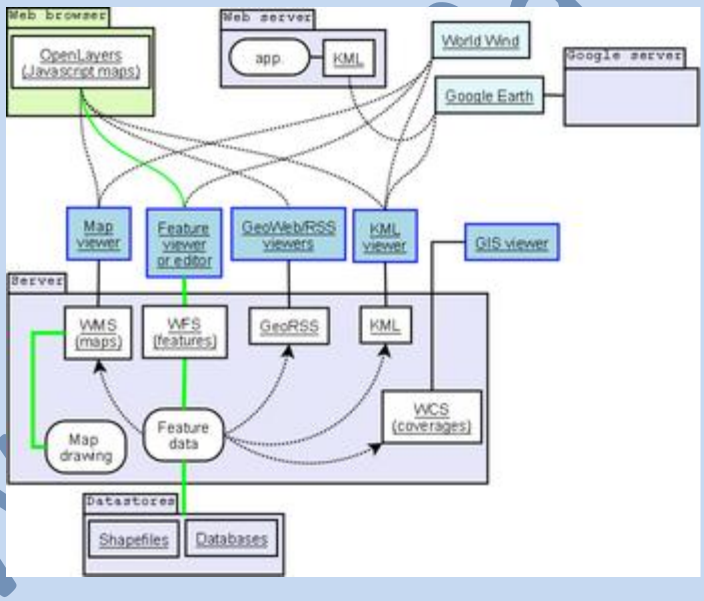


Many disciplines can benefit from GIS technology. An active GIS market has resulted in lower costs and continual improvements in the hardware and software components of GIS, and usage in the fields of science, government, business, and industry, with applications including real estate, public health, crime mapping, national defense, sustainable development, natural resources, climatology, landscape architecture, archaeology, regional and community planning, transportation and logistics. GIS is also diverging into location-based services, which allows GPS-enabled mobile devices to display their location in relation to fixed objects (nearest restaurant, gas station, fire hydrant) or mobile objects (friends, children, police car), or to relay their position back to a central server for display or other processing.

Open Geospatial Consortium standards

The Open Geospatial Consortium (OGC) is an international industry consortium of 384 companies, government agencies, universities, and individuals participating in a consensus process to develop publicly available geoprocessing specifications. Open interfaces and protocols defined by OpenGIS Specifications support interoperable solutions that "geo-enable" the Web, wireless and location-based services, and mainstream IT, and empower technology developers to make complex spatial information and services accessible and useful with all kinds of applications. Open Geospatial Consortium protocols include Web Map Service, and Web Feature Service.

GIS products are broken down by the OGC into two categories, based on how completely and accurately the software follows the OGC specifications.



OGC standards help GIS tools communicate.

Compliant Products are software products that comply to OGC's OpenGIS Specifications. When a product has been tested and certified as compliant through the OGC Testing Program, the product is automatically registered as "compliant" on this site.

Implementing Products are software products that implement OpenGIS Specifications but have not yet passed a compliance test. Compliance tests are not available for all specifications. Developers can register their products as implementing draft or approved specifications, though OGC reserves the right to review and verify each entry.

Web mapping

In recent years there has been a proliferation of free-to-use and easily accessible mapping software such as the proprietary web applications Google Maps and Bing Maps, as well as the free and open-source alternative OpenStreetMap. These services give the public access to huge amounts of geographic data.

Some of them, like Google Maps and OpenLayers, expose an API that enable users to create custom applications. These toolkits commonly offer street maps, aerial/satellite imagery, geocoding, searches, and routing functionality. Web mapping has also uncovered the potential of crowdsourcing geodata in projects like OpenStreetMap, which is a collaborative project to create a free editable map of the world.

Adding the dimension of time

The condition of the Earth's surface, atmosphere, and subsurface can be examined by feeding satellite data into a GIS. GIS technology gives researchers the ability to examine the variations in Earth processes over days, months, and years. As an example, the changes in vegetation vigor through a growing season can be animated to determine when drought was most extensive in a particular region. The resulting graphic represents a rough measure of plant health. Working with two variables over time would then allow researchers to detect regional differences in the lag between a decline in rainfall and its effect on vegetation.

GIS technology and the availability of digital data on regional and global scales enable such analyses. The satellite sensor output used to generate a vegetation graphic is produced for example by the Advanced Very High Resolution Radiometer (AVHRR). This sensor system detects the amounts of energy reflected from the Earth's surface across various bands of the spectrum for surface areas of about 1 square kilometer. The satellite sensor produces images of a particular location on the Earth twice a day. AVHRR and more recently the Moderate-Resolution Imaging Spectroradiometer (MODIS) are only two of many sensor systems used for Earth surface analysis. More sensors will follow, generating ever greater amounts of data.

In addition to the integration of time in environmental studies, GIS is also being explored for its ability to track and model the progress of humans throughout their daily routines. A concrete example of progress in this area is the recent release of time-specific population data by the U.S. Census. In this data set, the populations of cities are shown for daytime and evening hours highlighting the pattern of concentration and dispersion generated by North American commuting patterns. The manipulation and generation of data required to produce this data would not have been possible without GIS.

Using models to project the data held by a GIS forward in time have enabled planners to test policy decisions using spatial decision support systems.

Advantages of GIS

The advantages of using a geographic information system include:

- Improved decision making – decisions are made easier because specific and detailed information is presented about one or more locations.
- Reduce costs and increase efficiency – especially regarding maintenance schedules, fleet movements or scheduling timetables.
- Improved communication between any involved organisations or departments as the visual format is easily understood by all.
- Easy recordkeeping – geographical changes are easily recorded by GIS for those responsible of recording the changes.
- Managing geographically – knowing what is and will be occurring in a geographic space in order to plan a course of action.

Disadvantages of GIS

- Very expensive
- Requires enourmous amount of date: makes it prone for error
- Geographical error increases with larger scale
- Relative loss of resolution
- Violation of privacy

References

1. www.google.com
2. www.wikipedia.org
3. www.studymafia.org