International Journal of Information Technology and Knowledge Management
July-December 2010, Volume 2, No. 2, pp. 255-257

GREY INFORMATION IN SPATIAL DATABASES

Surya Prakash Tripathi[1], Praveen Kumar Shukla[2] & Shyam Kishore Bajpai[3]

Geographical Information Systems (GIS) and spatial databases are rapidly growing field in the database research. These databases provide the efficient handling of spatial data for storage, querying and indexing. Spatial data may be uncertain (grey) in its nature. In this paper, different issues regarding grey spatial information has been discussed with a brief introduction to spatial data and databases.

Keywords: Geographical Information Systems (GIS), Spatial Databases, Grey Information, Fuzzy Techniques.

1. Introduction

Geometric, geographic and spatial data are the data related to the space. This space may be a two dimensional abstractions of the earth surface called geographic space, a space created by a man like a layout of VLSI design, a model of human brain and 3-D space representing the architecture of protein molecules. Several names have been used for the databases supporting such type of data like pictorial, image, geometric, geographic and spatial database systems. Pictorial and image words are initiated from the fact that they also manage the data in the form of raster images, captured from remote sensing by satellite or by computer tomography (CT) scan in medical applications.

A spatial database offers spatial data types (SDTs) in its data model and query language and also supports spatial data types in its implementation, providing at least spatial indexing and efficient algorithms for spatial join.

A GIS (Geographic Information System) integrates the approach of digital mapping with database technology to support and provide a wide range of applications [1]. Spatial databases are the heart of GIS applications. The information stored in spatial database gives the description of location and shape of geographic properties in terms of points, lines and areas. Also, four type of attributes are associated with spatial objects. These includes spatial data (location, where an object is), temporal data (when an object is), thematic (attribute) data (when an object is) and scale data (how an object is) [2].

Since, the spatial data related to the GIS are fuzzy and uncertain (grey) [3, 4] in its nature. In this paper, different issues regarding uncertain information in spatial databases have been discussed.

2. Data types in Spatial Databases

This subsection describes four different data types that are common in two-dimensional spatial databases.

2.1. Point Data

Point data is data for which the position, but not the shape, size or other spatial properties are of interest. Examples of such features might be buildings in rural areas, mountain tops, and view points. Many point features arise because of generalization. Such point features should be modelled as the feature types they have on high scales and then generalized into points.

Points are often used in the representation of other kinds of features. A line, for instance, is often stored as a set of points with straight lines between them. Therefore, storing points is important in a spatial database even if there are not a lot of point features.

2.2. Line Data

Line data consists of objects which have a length and a shape, but no discernible area. One good example of a line feature would be a network of power lines. Other features which are often represented by lines are roads and rivers, but these are really long, narrow areas rather than lines. They may be stored as lines if the database is only going to be used to generate small-scale maps, or they can be stored as areas and generalized into lines for small scales.

[1]Asst. Professor & Head, Department of Computer Science & Engineering, Institute of Engineering & Technology, Lucknow, Uttar Pradesh, INDIA

[2]Senior Lecturer, Department of Information Technology & Engineering, Northern India Engineering College, Lucknow, Uttar Pradesh, INDIA

[3]Ex-Professor & Head, Department of Computer Science & Engineering, Institute of Engineering & Technology, Lucknow, Uttar Pradesh, INDIA

Email: tripathee_sp@yahoo.co.in[1], praveenshuklak@rediffmail.com[2], skbajpaiiet@gmail.com[3]

2.3. Field data

Field data is data which varies continually across the area of interest rather than being divided into discrete areas. Terrain models, soil types, pollution levels, and some forms of geological data are all good examples of such data. On traditional paper maps such data are either presented using contour lines, or by color coding the points in the map. Fields can have any number of dimensions. (Terrain models are two-dimensional, while geological data are three-dimensional. Meteorological data might be considered to be a four-dimensional field with time as the fourth dimension).

2.4. Region Features

A region is a geographical object where the shape and size are of interest, such as a plot of land or a country. Field features are sometimes represented as region features, but are a class of their own and should be represented differently in most cases. The main difference is that a region represents strictly bounded, discrete areas such as countries, while field data represent phenomena which vary continuously in the area of interest.

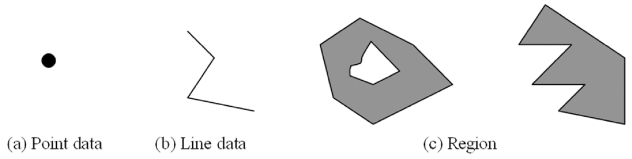

Fig. 1: Different Spatial Data

3. FUZZY TECHNIQUES: AN INTRODUCTION

In 1965, L.A. Zadeh developed the concept of fuzzy set theory [5, 6]. The basic idea behind this theory was existence of sets, in which it is not totally clear, whether an element belongs to the set or not. Each member is associated with a degree of membership that shows the possibility of a member to be present in the set. For example, the set of young people is a fuzzy set because there is an uncertainty with the decision of ranges of age that clearly specify the set young. Fuzzy logic can be integrated with the different approaches like, control systems, embedded systems, simulation, databases, modeling, pattern recognition, knowledge based systems and different Artificial Intelligence approaches.

Let U be a universe of discourse. A fuzzy value on U is characterized by a fuzzy set F in U. A membership function $\mu_F : U \rightarrow [0,1]$ is defined for the fuzzy set F, where $\mu_F(u)$ for each u ∈ U, denotes the degree of membership of u in the fuzzy set F. Thus, fuzzy set F is described as follows: $F = \{\mu_F(u_1)/u_1, \mu_F(u_2)/u_2, ..., \mu_F(u_n)/u_n\}$ or $F = \int_{u \in U} \mu_F(u)/u$ with u ∈ U, where μ_F(u) is a measure of possibility that a variable x has the value u in this approach, where x takes a value in U.

4. UNCERTAINTY ISSUES IN SPATIAL DATABASES

Several types of uncertainty have been found in spatial databases. These are related to spatial objects, spatial querying, regions and boundaries etc.

1. In spatial GIS, the information associated with objects may be uncertain (grey) in its nature. Few of these uncertainties are due to: 1. Missing data, 2. Uncertain Data, 3. Geo statistic, 4. Multi dimensional uncertainties, 5. Uncertain boundaries, 6. Uncertain object grouping.
2. A vague region is defined as a pair of disjoint, crisp regions. The first region is called Kernel and this describes the area which definitely and always belongs to the vague region. The second region is called the boundary describes the area for which we cannot say with any certainty whether it or its parts belong to the vague region or not.
3. Boundaries are considered as a sharp line that represent abrupt changes of spatial phenomenon and discriminate the regions with different characteristics features. These boundaries may be vague in certain fields, like 'boundary of the Indian Ocean' or a boundary between a mountain and a valley.
4. Different kinds of uncertainties have been found in querying, like 'Found all the high buildings near the airport'. In the above query, 'high' and 'near' are the linguistic variables which are uncertain in its nature and may be represented by fuzzy set theory.

Several models have been proposed to model the uncertainty in spatial data.

Table – 1 Spatial Database Models

S. No.	Model	Description	References
1	Exact Model	Spatial objects with sharp boundaries and with restricted support of uncertainties	[7, 8]
2	Probabilistic Model	Based on probability theory	[9]
3	Fuzzy Model	Based on Fuzzy Sets Theory	[10]

Different issues regarding uncertain spatial data has been discussed in [11, 12]. A discrete model for uncertain spatial data has been proposed in [13].

5. CONCLUSION

One of the important features of the spatial data related to GIS applications is that they are imprecise and uncertain inherently. The accurate representation of this uncertainty is an important research issues. Several approaches have been introduced to handle this uncertainty, like Fuzzy Logic, Probability Theory, Possibility Theory, etc. This paper introduces different places of imprecise information in spatial databases. In future, our efforts are oriented towards the development of a database model that represents imprecise (grey) information in spatial databases, precisely.

REFERENCES

[1] Maguire, D. J., Goodchild, M., Rhind, D. W., Geographical Information Systems: Principles and Applications, Longman, 1991.

[2] Feuchtwanger, M., Towards a Geographic Semantic Database Model. Ph.D. Thesis, Simon Fraser University, Vancouver, 1993.

[3] Burrough, P. A. “Natural Objects with Indeterminate Boundaries”, Geographic Objects with Indeterminate Boundaries (Burrough, P., and Frank, A., eds.), 3-28, Taylor and Francis, London, 1996.

[4] Frank, A., “The Prevalence of Objects with Sharp Boundaries in GIs”, Geographic Objects with Indeterminate Boundaries (Burrough, P., and Frank, A., eds.), 29-39, Taylor and Francis, London, 1996.

[5] L.A. Zadeh, "Fuzzy Sets", Information and Control, 8, No. 3, 1965, pp. 338-353.

[6] L.A. Zadeh, "Fuzzy Sets as a Basis for a Theory of Possibility", Fuzzy Sets and Systems, 1, 1978, pp. 3-28.

[7] E. Clementini, P. di Felice, An Algebraic Model for Spatial Objects within Indeterminate Boundaries, in [BF96], 1996, pp. 171-187.

[8] M. Erwig, M. Schneider, Vague Regions, 5th International Symposium on Advances in Spatial Databases, Springer Verlag, LNCS, 1997, pp. 298-320.

[9] P. A. Burrough, Natural Objects with Indeterminate Boundaries, in [BF96], pp. 3-28, 1996.

[10] D. Altman, "Fuzzy Set Theoretic Approach for Handling Imprecision in Spatial Analysis", International Journal of Geographic Information Systems, 8, No. 3, pp. 271-289, 1994.

[11] E. Tøssebro and R. H. Güting: Creating Representations for Continuously Moving Regions from Observations. In Proc. 7th Int. Symp. on Spatial and Temporal Databases, July 2001, pp. 321-344.

[12] E. Tøssebro and M. Nygård: Uncertainty in Spatiotemporal Databases. In Proc. 2nd Int. Conf. on Advances in Information Systems (ADVIS), October 2002.

[13] E. Tøssebro and M. Nygård: An Advanced Discrete Model for Uncertain Spatial Data, in Proc. 3rd Int. Conf on Web-Age Information Management (WAIM), August 2002.