Spatial and Spatio-Temporal Data Mining

Vania Bogorny
Universidade Federal do Rio Grande do Sul
www.inf.ufrgs.br/~vbogorny
vbogorny@inf.ufrgs.br

Shashi Shekhar
University of Minnesota
www.cs.umn.edu/~shekhar
shekhar@cs.umn.edu

Tutorial - Outline

- Part I
  - Introduction to Spatial Databases
  - Spatial Data Mining Methods
- Part II
  - Introduction to Moving Object Databases
  - Spatio-Temporal Data Mining

Outline

Part I

- Spatial/Geographic Databases
  - Geographic Data
  - Spatial Relationships / Spatial Operations
  - Spatial Query Language
- Introduction to Spatial Data Mining
- Spatial Data Mining Methods
  - Co-location
  - Outliers
  - Spatial Association Rules
  - Classification
  - Clustering

Introduction to Spatial Databases

What is a Spatial/Geographic Database?

Database that:

- Stores spatial objects
- Provides data types for spatial objects
- Provides operations to manipulate spatial objects
- Manipulates spatial objects just like other objects in the database

Spatial Databases

- Oracle Spatial
- IBM DB2 Spatial Extender
- Informix Spatial DataBlade
- MS SQL Server (with ESRI SDE)
- Geomedia on MS Access
- PostGIS / PostgreSQL

- TerraView (Spatial Data Visualization and Analysis)

What is a Spatial/Geographic Data?

- Data which describe a location or a shape
  - e.g. House, Hospital, Road, River, Forests, Parks, Soil
- *Is something that describes objects or phenomena that happen on the Earth and that have associated a geographic position*

What is a Spatial/Geographic Data?

- Three main characteristics describe a geographic object:
  - ***Non-spatial attributes (what):*** describe either quantitatively or qualitatively a geographic entity.
    - These attributes may be treated by non-spatial databases;
  - ***Spatial attribute (where):*** describe the spatial location and representation of the geographic object, considering the geometry and a coordinate system.
    - This aspect requires a specific data type not available in conventional DBMS;
  - ***Spatial relationships (how):*** neighbourhood relationships (e.g. topology, distance).
    - Requires special operations that are not available in conventional DBMS;

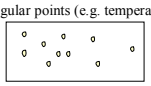
Spatial Representation: OBJECT

- Discrete objects (Features): well defined border/limit
  - *0-dimensional*
    - representation: point
    - E.g.: school, hospital
  - *Uni-dimensional*
    - representation: line
    - E.g.: river, road
  - *Bi-dimensional*
    - representation: polygon
    - E.g.: state, city
  - *Tri-dimensional*
    - representation: surface

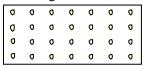
Spatial Representation: FIELD

Continuous Data

Irregular points (e.g. temperature)

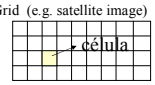


Regular Points

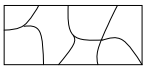


Isoline (e.g. relief)

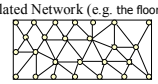
Grid (e.g. satellite image)



Adjacent polygons (e.g. soil)



Triangulated Network (e.g. the floor of a valley)



Geographic Data and Geographic Databases: an example

Street

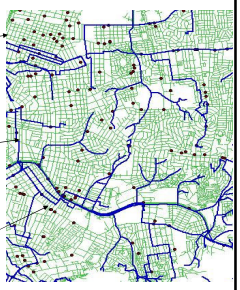
Gid	Name	Shape
1	Ijui	Multiline [(x1,y1),(x2,y2)...]
2	Lavras	Multiline [(x1,y1),(x2,y2)...]

WaterResource

Gid	Name	Shape
1	Jacui	Multiline [(x1,y1),(x2,y2)...]
2	Guaiba	Multiline [(x1,y1),(x2,y2)...]
3	Uruguai	Multiline [(x1,y1),(x2,y2)...]

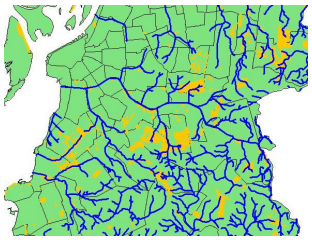
GasStation

Gid	Name	VolDiesel	VolGas	Shape
1	BR	20000	85000	Point[(x1,y1)]
2	IPF	30000	95000	Point[(x1,y1)]
3	Esso	25000	120000	Point[(x1,y1)]

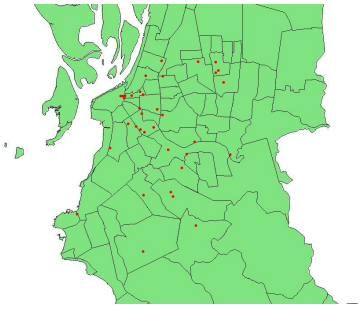


Example of Geographic Data

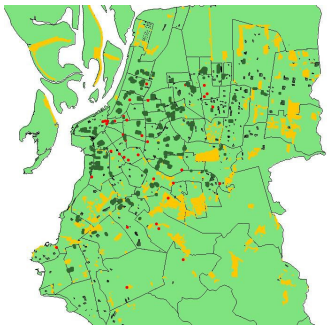
Districts
Slums
Water bodies



Example of Geographic Data



Overlay of Geographic Data



Combination of some previous data

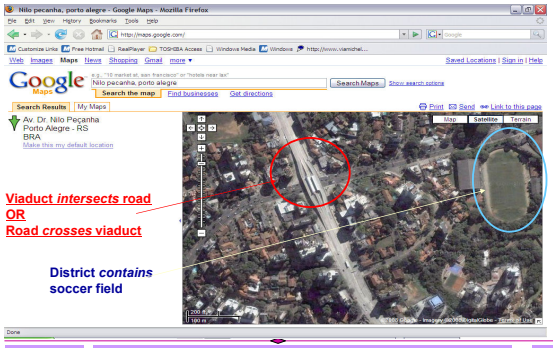
⇩

SPATIAL RELATIONSHIPS

Spatial Relationships

- Main characteristic which differs spatial data from non-spatial data
- Main aspect to be considered in SPATIAL DATA MINING

Spatial Relationships: an example

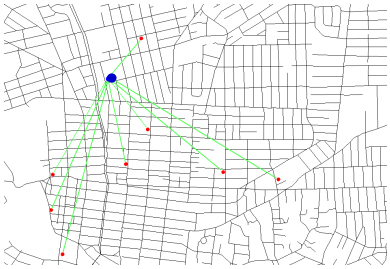


Spatial Relationships

What is the average distance between industrial repositories and water collecting points?

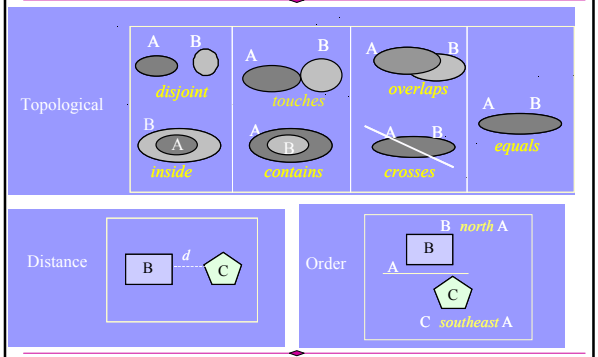
- Water collecting points
- Districts
- Industrial Repositories
- Water Body

Spatial Relationships

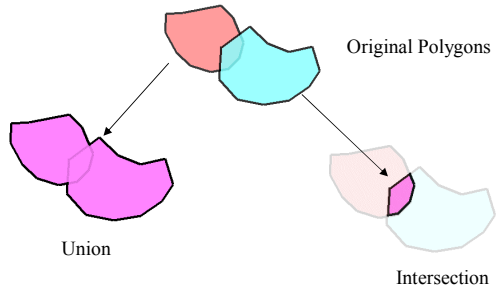


Which is the closest gas station to my house?

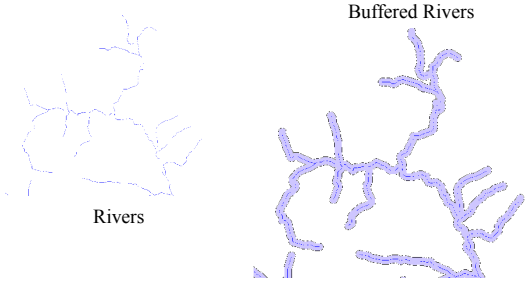
Main Spatial Relationships



Examples of Spatial Operations



Examples of Spatial Operations



Advantages of Spatial Databases

Spatial querying using SQL

- SQL expressions to determine *spatial relationships*
  - distance
  - order
  - topology
- SQL expressions to perform *spatial operations*
  - area
  - length
  - union
  - buffer

Spatial Query Language

- Spatial query language provides:
  - spatial data types, e.g. point, line, polygon, …
  - spatial operations, e.g. overlap, distance, buffer, …
  - example:

```
SELECT S.name
  FROM State S
WHERE area (s.the_geom) > 300
```

- Standards
  - OGIS is a standard for spatial data types and operators

Spatial Query Example

Q1: Retrieve the rivers and countries that have the relationship "crosses".

Query

```
SELECT r.name, c.cntry_name
   FROM river r, country c
 WHERE crosses (r.the_geom,c.the_geom) = 'True'
```

Answer

```
         name          |  cntry_name
-----------------------+--------------
Pembina                | United States
Pembina                | Canada
Rainy                  | United States
Rainy                  | Canada
Souris                 | United States
Souris                 | Canada
Red River of the North | United States
Red River of the North | Canada
(8 rows)
```

Spatial Query Example

Q2: Retrieve the states adjacent to the state "Santa Catarina".

Query

```
SELECT s1.name as state_Neighbors
  FROM state s1, state s2
 WHERE touches(s1.the_geom, s2.the_geom)=TRUE
         and  s1.state_name='Santa Catarina';
```

Answer

```
State_Neighbors
--------------------
Rio Grande do Sul
Paraná
(2 rows)
```


References

- GUTING, R. H. An Introduction to Spatial Database Systems. The International Journal on Very Large Data Bases, [S.l.], v.3, n.4, p. 357 – 399, Oct. 1994.
- RIGAUX, P.; SCHOLL, M.; VOISARD, A. Spatial Databases: With Application to GIS. San Francisco: Morgan Kaufmann, 2002.
- SHEKHAR, S., CHAWLA, S. Spatial databases: a tour. Upper Saddle River, NJ: Prentice Hall, 2003.
- OPEN GIS CONSORTIUM. Topic 5, the OpenGIS abstract specification–OpenGIS features–Version 4. 1999a. Available at <http://www.OpenGIS.org/techno/specs.htm>.
- OPEN GIS CONSORTIUM. OpenGIS simple features specification for SQL. 1999b. Available at <http://www.opengeospatial.org/docs/99-054.pdf>.
- OPEN GIS CONSORTIUM. Feature Geometry. 2001. Available at <http://www.opengeospatial.org/specs>.