Mining Spatial Associations

Spatial Frequent Patterns and Associations

- Spatial frequent patterns and association rule: $A \Rightarrow B$ [*s%, c%*]
  - A and B are sets of spatial or non-spatial predicates, e.g.,
    - Topological relations: *intersects, overlaps, disjoint,* etc.
    - Spatial orientations: *left_of, west_of, under,* etc.
    - Distance information: *close_to, within_distance,* etc.
  - Measures: *s%*: support, and *c%*: confidence of the rule
- Example: Rules likely to be found
  - *is_a(x, large_town) ^ intersect(x, highway) → adjacent_to(x, water) [7%, 85%]*
- Explore *spatial autocorrelation*: Spatial data tends to be highly self-correlated (*nearby things are more related than distant ones)*
  - E.g., neighborhood, temperature

Mining Spatial Associations: Progressive Refinement

- Hierarchy of spatial relationship:
  - *close_to* is a generation of *near_by*, *touch*, *intersect*, *contain*, ...
  - Progressive refinement: First search for rough relationship and then refine it
- Two-step mining of spatial association:
  - Step 1: Rough spatial computation (as a filter)
    - Using MBR (Minimum Bounding Rectangle) or R-tree for rough estimation
  - Step2: Detailed spatial algorithm (as refinement)
    - Apply only to those objects which have passed the rough spatial association test (no less than *min_support*)