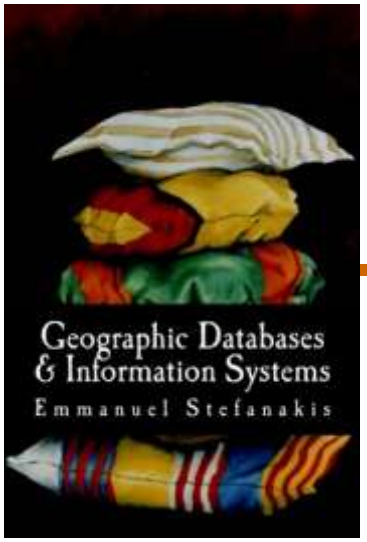


Stefanakis, E., 2014. *Geographic Databases and Information Systems*. CreateSpace Independent Publ. [In English], pp.386.

Get a copy from Amazon

Chapter 14

Geographic Operation Execution Strategies

Emmanuel Stefanakis

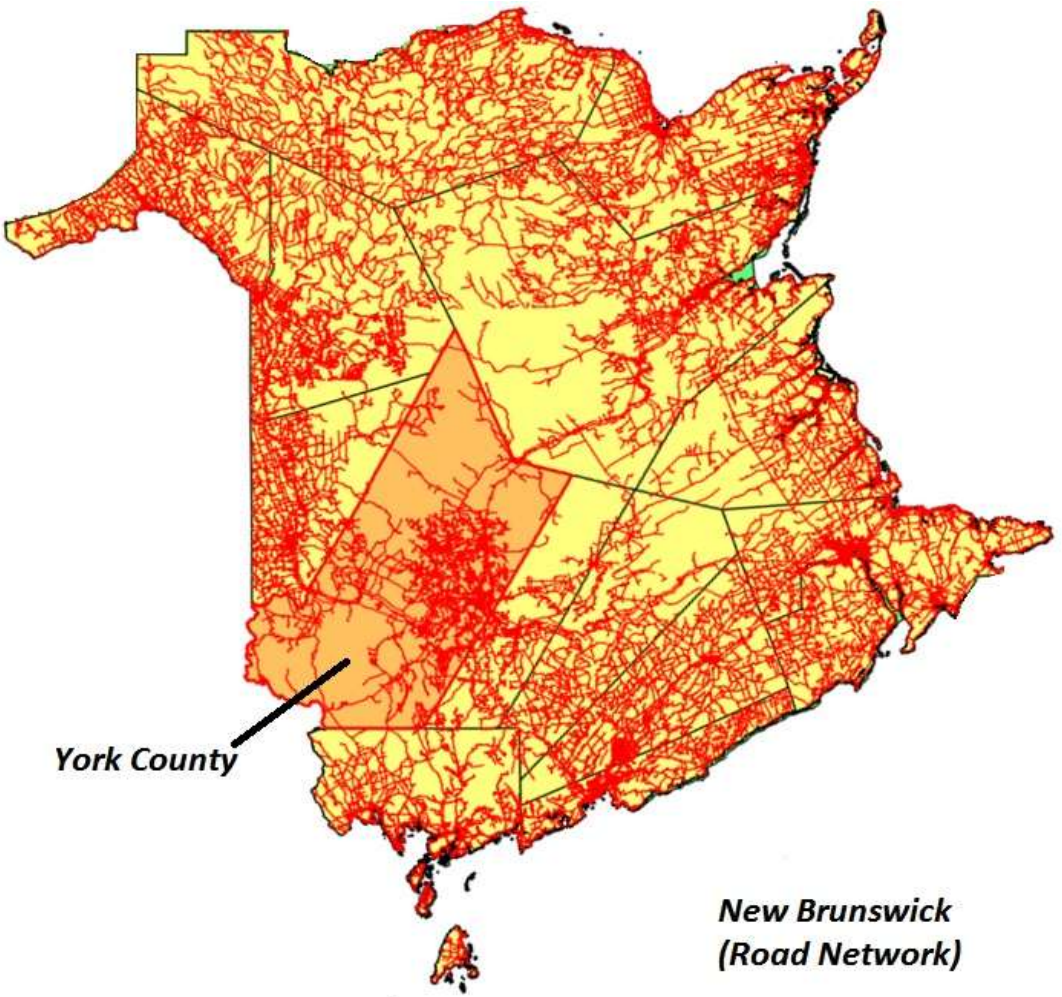
http://www2.unb.ca/~estef/

Example

- Road network
- Counties

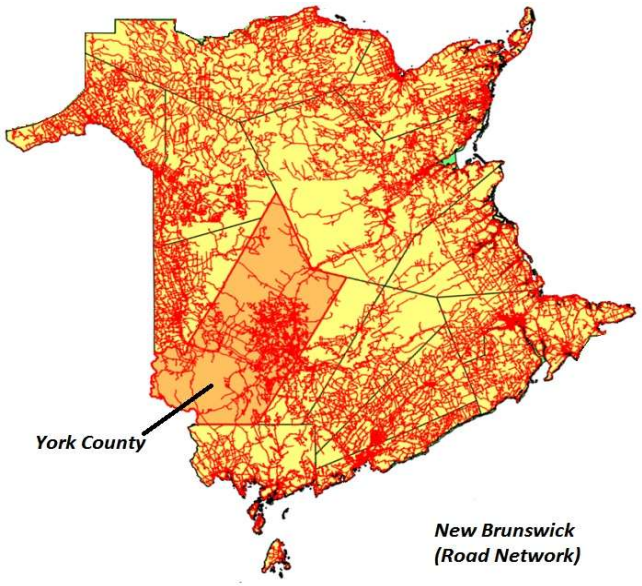
What highways cross York County?

(spatial join operation)

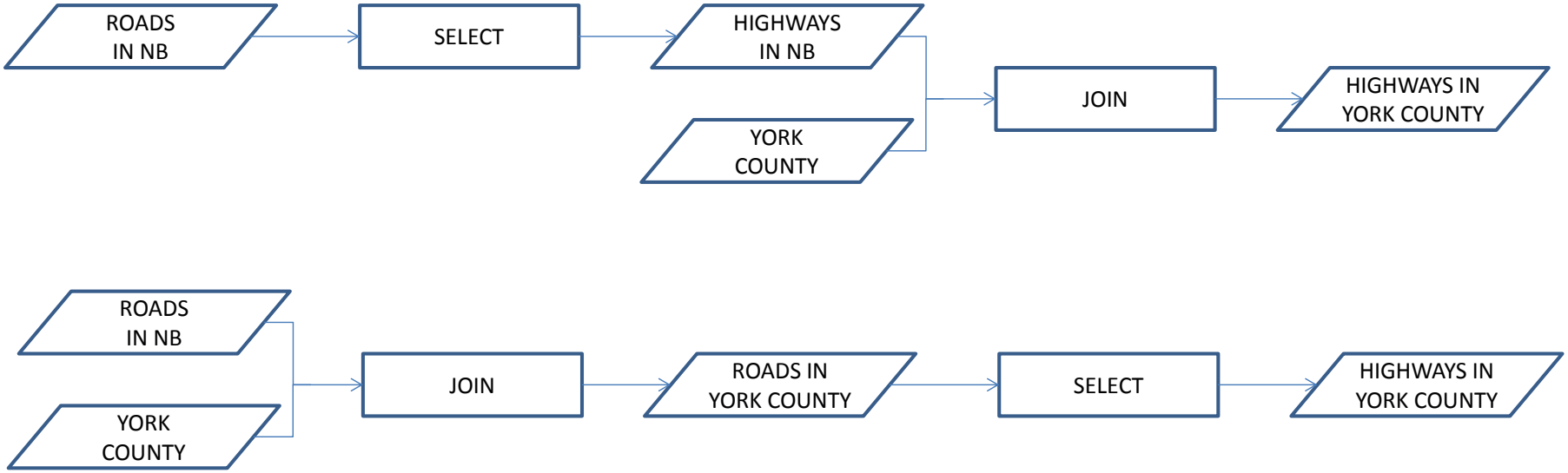


Example

What highways cross York County?



Two scenarios – which one is executed faster?

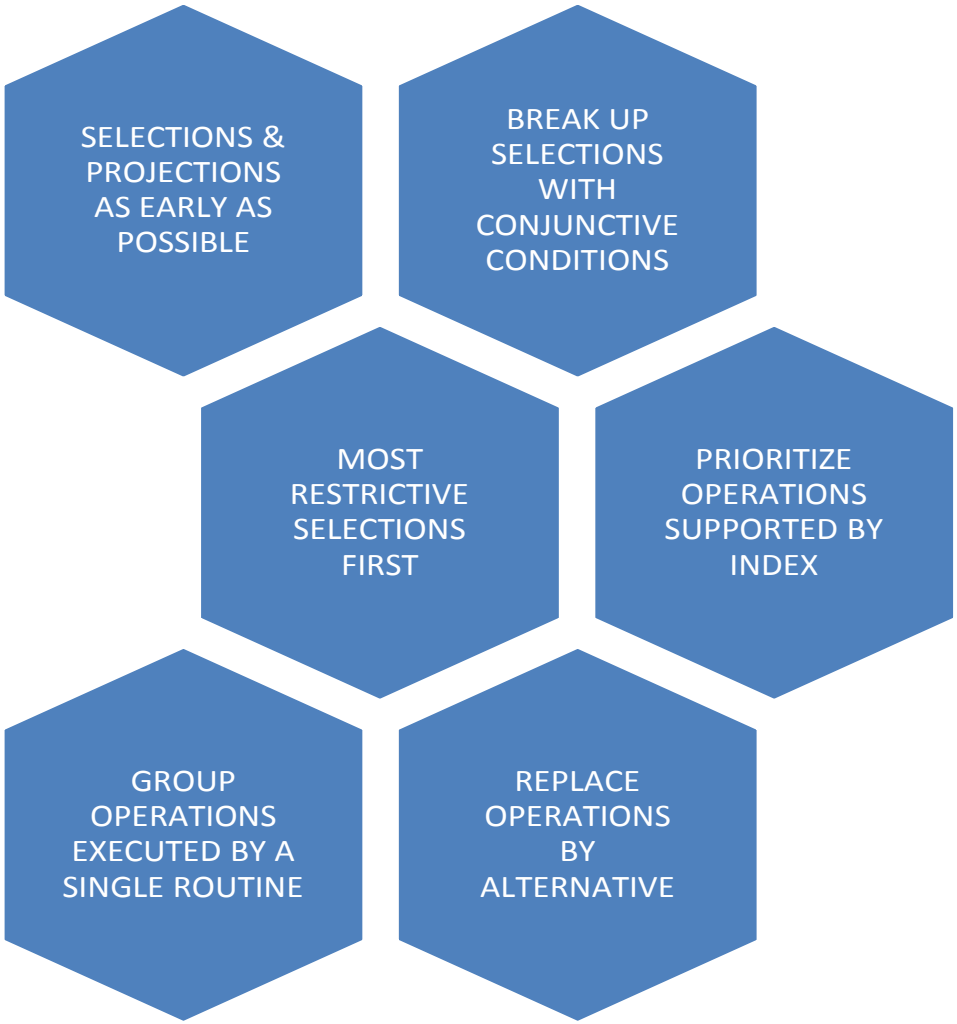


Optimization techniques

- Applied by the **DBS Optimizer**

1. Heuristic rules
2. Systematic cost estimate (cost models)
3. Pre-compute and use ancillary spatial attributes and relationships

1. Heuristic rules



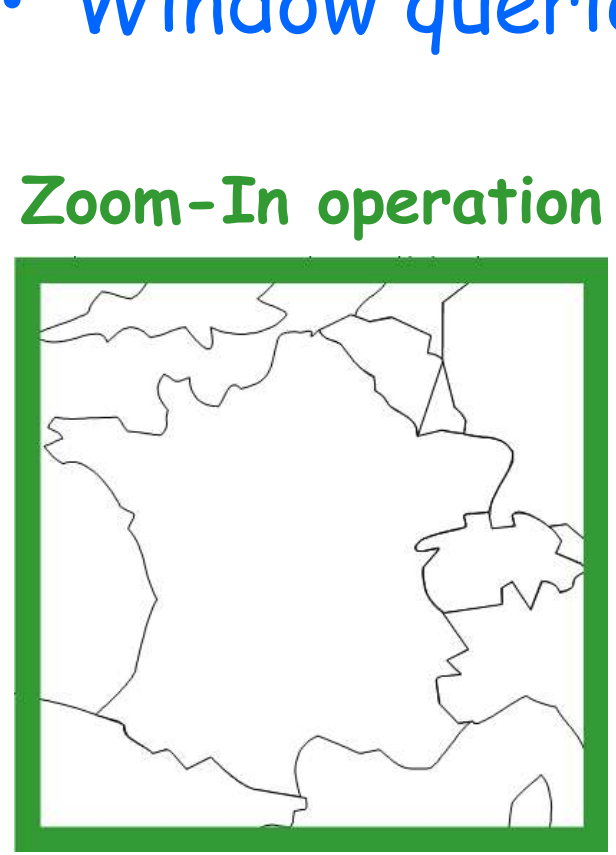
2. Systematic Cost Estimate

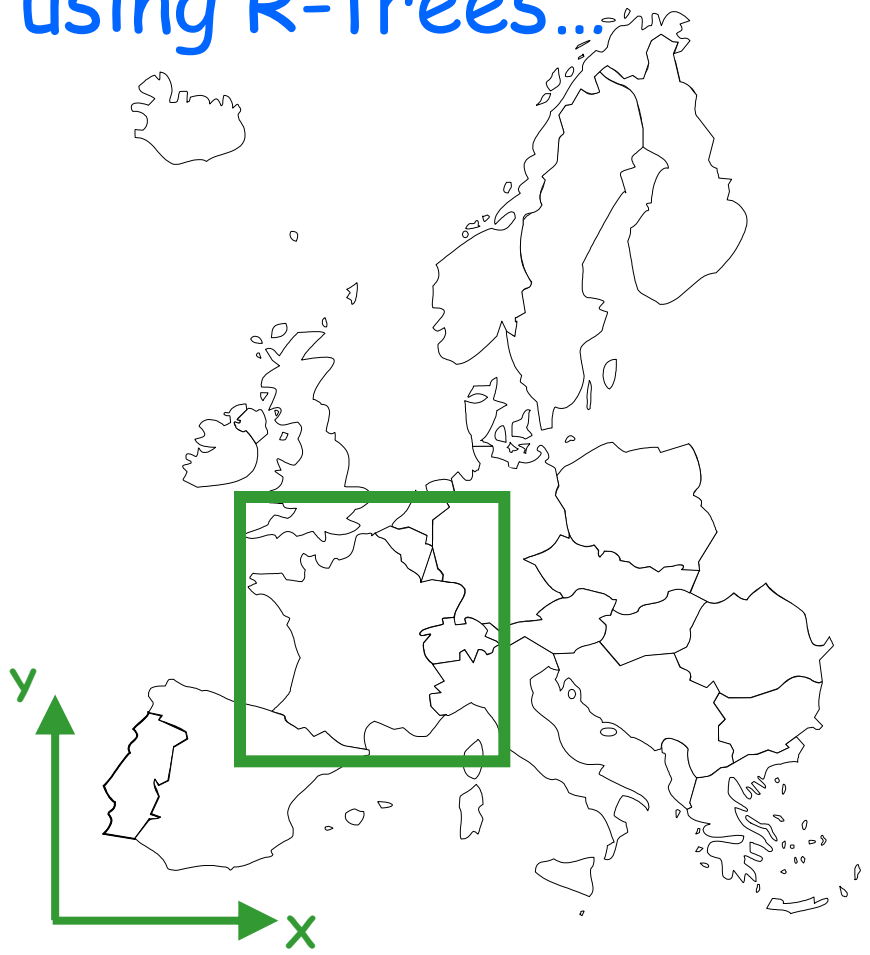
- Analytical models…
  - that can predict the cost of spatial queries
    - Point queries
    - Window queries
    - Spatial joins
    - Topological queries
  - Supported by spatial indices (e.g., R-trees)

2. Systematic Cost Estimate

- Window queries using R-trees…

Zoom-In operation





2. Systematic Cost Estimate

- Window queries using R-trees…
  - n-dimensional query window Q with extents (Q1 , Q2 , … , Qn), along each dimension
  - h the height of the tree structure;
  - Nj the expected number of nodes in the tree; and
  - sj the average node extent along each dimension at level j of the tree

$$C(Q) = 1 + \sum_{j=1}^{h-1} \left\{ N_j \cdot \prod_{i=1}^{n} \left(s_j + Q_i \right) \right\}$$

2. Systematic Cost Estimate

- Window queries using R-trees…

$$C(Q) = 1 + \sum_{j=1}^{h-1} \left\{ N_j \cdot \prod_{i=1}^{n} \left(s_j + Q_i \right) \right\}$$

$$h = 1 + \left\lceil \log_{c \cdot M} \frac{N}{c \cdot M} \right\rceil$$

$$s_j = \left(D_j / N_j \right)^{1/n}$$

$$N_j = \frac{N}{(c \cdot M)^j}$$

$$D_j = \left\{ 1 + \frac{\left(D_{j-1} \right)^{1/n} - 1}{(c \cdot M)^{1/n}} \right\}^n$$

2. Systematic Cost Estimate

- Spatial Joins using R-trees…

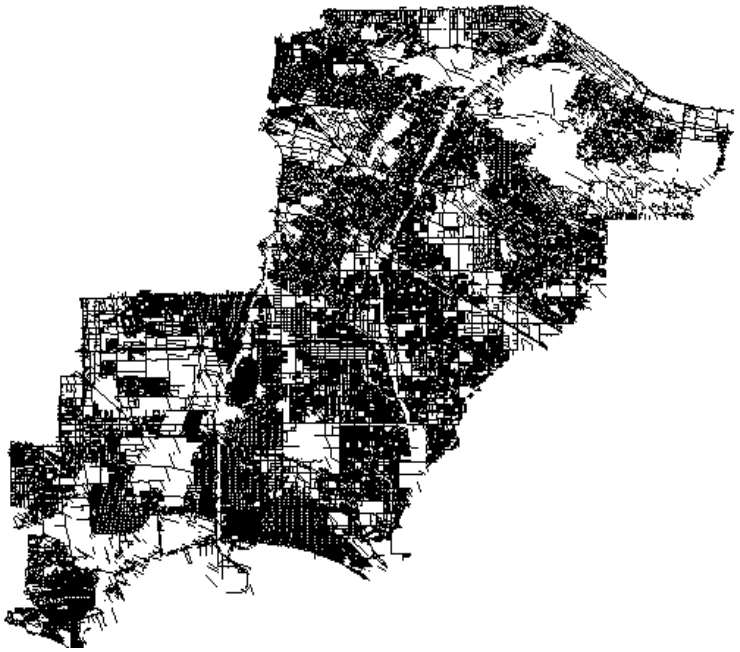
$$C(R_1, R_2) = \sum_{j=1}^{h-1} \left\{ N_{R_2,j} \cdot N_{R_1,j} \cdot \prod_{k=1}^{n} \left(s_{R_1,j,k} + s_{R_2,j,k} \right) + N_{R_2,j} \cdot N_{R_1,j+1} \cdot \prod_{k=1}^{n} \left(s_{R_1,j+1,k} + s_{R_2,j,k} \right) \right\}$$

$$N_{Ri,j} = \frac{N_{Ri}}{(c \cdot M)^j} \qquad s_{Ri,j,k} = \left(D_{Ri,j} \Big/ N_{Ri,j} \right)^{1/n}$$

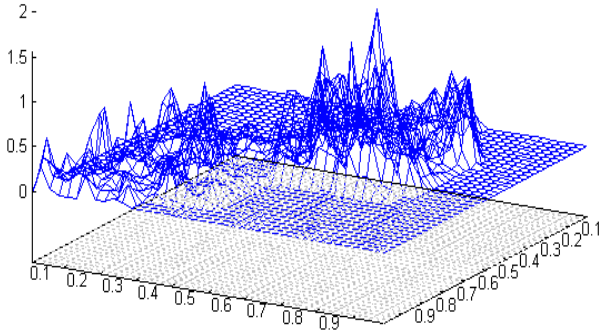
$$h = 1 + \left\lceil \log_{c \cdot M} \frac{N_{Ri}}{(c \cdot M)^i} \right\rceil \qquad D_{Ri,j} = \left\{ 1 + \frac{\left(D_{Ri,j-1} \right)^{1/n} - 1}{(c \cdot M)^{1/n}} \right\}^n$$

2. Systematic Cost Estimate

- Evaluation of cost models…



(a)



(b)

A real data set (a) and the corresponding density surface (b).

2. Systematic Cost Estimate

- Evaluation of cost models…

Data sets	Relative Error		
	point queries	range queries	join queries
Random data	0%-10%	0%-5%	0% - 15%
Skewed data	0%-15%	0%-10%	0%-20%
Real data	0%-15%	0%-20%	0%-15%

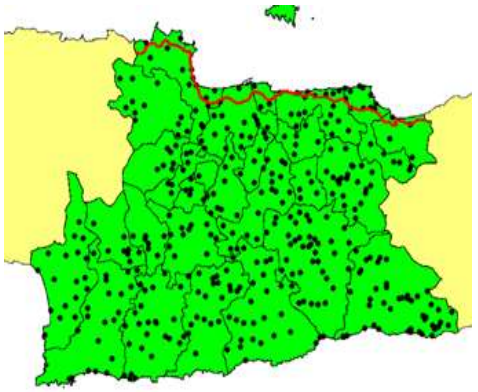
3. Pre-compute spatial attributes and relationships

- pre-compute and store certain spatial attribute values…
  - such as the area and perimeter of a region

Store area explicitly

Table: municipalities

gid integer	ID double precision	NAME character varying(60)	POP_01 double precision	the_geom geometry	area double precision
1	841	Irakleio	137711	…	108856910.092
2	842	Agia Varvara	5310	…	98473803.8593
3	843	Arkalochori	10897	…	238720576.640
4	844	Archanes	4548	…	31845436.375
5	845	Asterousia	6303	…	204604555.984



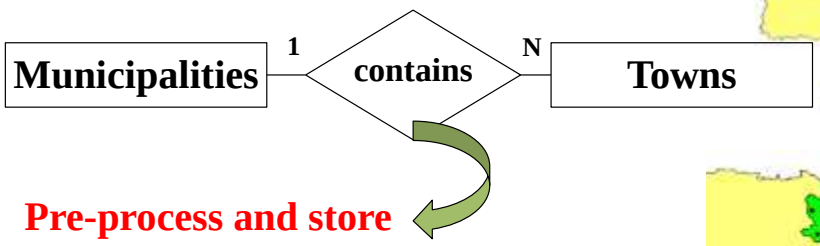
3. Pre-compute spatial attributes and relationships

- pre-compute and store certain spatial attribute values…
  - such as the area and perimeter of a region

- Pros: no need to compute all the time
- Cons: waste of disk; need to re-compute after updates

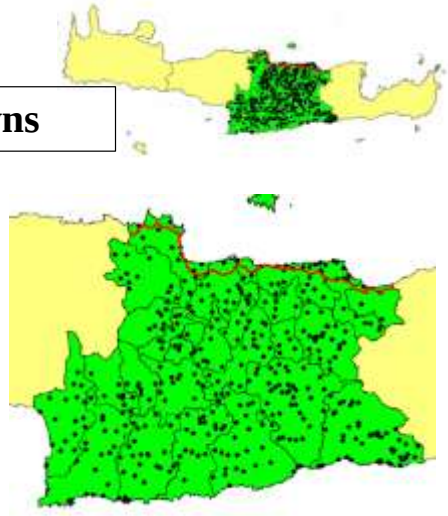
3. Pre-compute spatial attributes and relationships

- spatial relationships…
  - are seldom stored explicitly due to their large cardinality

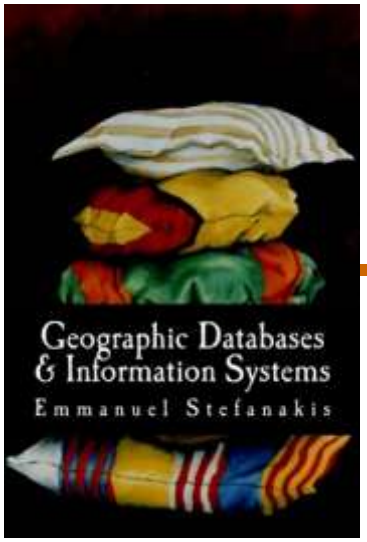


Pre-process and store or extract on demand?

Topological relationships



	NAME (Municipality) character varying(60)	name (Town) character varying(60)
1	Agia Varvara	Prinias
2	Agia Varvara	Doulio
3	Agia Varvara	Agios Thomas
4	Agia Varvara	Peirouniana
5	Agia Varvara	Preveliana
6	Agia Varvara	Megali Vrysi
7	Agia Varvara	Kolena
8	Agia Varvara	Agia Varvara
9	Agia Varvara	Genna
10	Agia Varvara	Ano Moulia
11	Agia Varvara	Larani
12	Agia Varvara	Kato Moulia



Stefanakis, E., 2014. *Geographic Databases and Information Systems*. CreateSpace Independent Publ. [In English], pp.386.

Get a copy from Amazon

Chapter 14

Geographic Operation Execution Strategies

Emmanuel Stefanakis

http://www2.unb.ca/~estef/