CS 6630: SPATIAL AND MULTIDIMENSIONAL DATABASES

Semester Hours: 3.0
Coordinator: Ray Kresman
Text: Spatial databases- a tour
Authors: Shekhar and Chawla
Year: 2003

SPECIFIC COURSE INFORMATION

Catalog Description:

Introduction to advanced database structures and large datasets. Efficient data structures and related algorithms for spatial, streaming and multi-dimensional and semi-structured datasets. Employs concepts from databases, algorithms, computer graphics and computational geometry.

Prerequisites: CS 5620 or permission of instructor.

Course type: ELECTIVE

SPECIFIC COURSE GOALS

- I am able to store, retrieve and manipulate multidimensional data using advanced data structures such as MX-quad tree, BBD-tree, R-tree, and others.
- I am able to formulate spatial queries that permit efficient data.
- I am able to distinguish between various spatial distance metrics.
- I am able to explain the mechanics of certain algorithms for similarity searching.
- I am able to use advanced SQL operations to query data warehouses.
- I am able to explain the nature of streaming data and algorithms for certain problems.
- I am able to critically evaluate a research literature in the realm of multidimensional, spatial or streaming data.
LIST OF TOPICS COVERED

1. Introduction
   - Large datasets
   - Spatial data & GIS
   - Streaming data

2. Graph Theory
   - Elementary graphs
   - Computational geometry

3. Multidimensional Datasets
   - Transactional data and relational schemas
   - Dimensional models
   - Snowflake schemas
   - Data warehousing & SQL

4. Spatial Datasets
   - Representation
   - Access methods
   - Trees: R-tree, Kd-tree, quad-tree, etc.
   - Performance tradeoffs

5. Data Storage and Manipulation
   - Spatial Object types
   - Spatial queries & operations
   - Similarity search/methods
   - Spatial algebra

6. Streaming Data
   - Sample problem: sampling, cardinality/moments estimation
   - Clustering & space filling curves
   - Approximation algorithms

7. Performance
   - Spatial indices
   - Clustering & space filling curves
   - Data quality and metrics

8. Mining
   - Association rules
   - Continuous space and spatial co-location
   - Spatial autocorrelation