CS 3550

Data Stream Management

Prof. Panos K. Chrysanthis
Prof. Alexandros Labrinidis

Fall Term 2008 (09-1 or 2091)

Logistics

• When: Mon & Wed 1:00-2:15pm
• Where: 6516 SENSQ
• What:
  – Project-oriented investigation
  – Study state of the art-papers

Goals

• Understand the state-of-the-art in data streams, related applications
• Discover unsolved problems and challenges
• Learn (practice) how to give a good presentation
• Learn (practice) how to review papers
• Learn (practice) how to write a good technical paper
• Produce a publishable paper

Course Requirements

• Participation: 20%
• Presentations: 25%
• Paper Reviews: 15%
• Term Project & Report: 40%
**Course Structure**

Choose a topic from the suggested ones
- group of 2 students is permitted (FCFS)
1. Each group will
   - Come with bibliography; initial list of 2-3 other papers
   - Present the papers (1-2 talks)
2. Each group will
   - Execute a project
   - Do a project presentation at the end of each month and at the end of the term
   - Write a project report

**Preparation of your Talk**

- Reading: Read the papers but read others as well:
  - Citers and Cited, follow-ups by the same author, etc.
- Assume that the average reader has understood the easiest 2/3 of the paper.
- You, the expert on the papers, need to supply the rest.

**Talk Outline**

- Categorize issues and solutions in your topic
  - those that are unique to the new environments and systems
  - those that are shared with any distributed system
- Broad-brush sketch of important results
  - give outline of talk in this context
- Postpone discussion of things you are going to treat in detail later
- Details of 2-3 chosen issues/solutions
- Summary of solved problems, unsolved problems, non-problems
  - A peek into your paper & project ideas

**Administrative**

  - check often!
- use keyword cs3550 in all emails to instructor (as part of the subject line)
- class mailing list: You would be signed up.
# Topics & Projects

<table>
<thead>
<tr>
<th>Topics</th>
<th>Metrics</th>
<th>Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduling</td>
<td>QoD, QoS</td>
<td>QoD Scheduling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Class-based Scheduling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aggregate CQ Scheduling</td>
</tr>
<tr>
<td>Load Shedding</td>
<td>Energy</td>
<td>Semantic / Syntactic / Interaction with Scheduler</td>
</tr>
<tr>
<td>Implementation</td>
<td>Time</td>
<td></td>
</tr>
<tr>
<td>Challenges</td>
<td>Space</td>
<td>Impact of Scheduling / Query Optimization</td>
</tr>
</tbody>
</table>