The recent development of innovative technologies related to mobile computing combined with smart city infrastructures is generating massive, heterogeneous data and creating opportunities for novel applications in transportation computation science. The heterogeneous data sources provide streams of information that can be used to create smart cities. The knowledge on stream analysis is thus crucial and requires collaboration of people working in logistics, city planning, transportation engineering and data science.

We provide a list of materials for a course on stream processing for computational transportation science. The objectives of the course are:

- Motivate data stream and event processing, its model and challenges.
- Acquire basic knowledge about data stream processing systems.
- Understand and analyze their application in the transportation domain.

Since the subject is large and comprises many aspects, we propose that the course should start with an examplary application which is familiar to the audience. The chosen example expands through the whole course and illustrates a particular aspect in each section.

Topics to be covered:

1. Introduction
   - Literature:
     - Models and Issues in Data Stream Systems [4, 13]
     - Data Stream Management: [21]
     - Transportation and Data Streams: [40]
     - Smart Cities and Heterogeneous Data Streams: [36]
   - Event stream examples in transportation:
     - linear ordered sequence events, e.g., bus arrival times
     - an event cloud consists of many event streams, e.g., traveler arrival time and bus arrival time at interchange
     - moving car trajectories
   - Challenges in stream processing
   - OGC standards and interfaces [15]

2. Data Stream Management Systems (DSMSs)
   - Lambda Architecture for Stream Processing [31]
   - Speed Layer: STREAM [1], Aurora [6], Borealis [39], Storm [38], streams [9], S4 [32], Kafka [19]
   - Batch Layer: MapReduce [14], Hadoop [41], Spark [42], Disco [16]
   - Distributed NoSQL Databases: Cassandra [24], MongoDB [35]

3. Data Analysis
   - Query Languages: Esper [30], NiagaraCQ [10], and others [5, 22, 25]
   - Complex Event Processing (CEP): [7, 12, 27]
   - Learning: streams [9], Mahout [33], MOA [8]
   - Distributed streams: [37]
   - Sketches: [11, 18] privacy with sketches [23]

4. Example applications in the transportation domain: [2, 3, 17, 20, 26, 28, 29, 34, 43]
Social Issues in Computational Transportation Science

Possible home assignment:
- Study a certain DSMS and summarize its features in a report.

References


Social Issues in Computational Transportation Science


