Title
SIP5: Understanding and Controlling Instrumented Physical Systems: Modeling is Complex, but Optimization is Easy

Permalink
https://escholarship.org/uc/item/2p13n8md

Authors
Bradley Bennett
Alberto Cerpa
Jessica Feng
et al.

Publication Date
2005
Understanding and Controlling Instrumented Physical Systems:
Modeling is Complex, but Optimization is Easy

Bradley Bennett, Alberto Cerpa, Jessica Feng, Farinaz Koushanfar, Richard Park, Thomas Schoelhammer, Jennifer L. Wong, and Miodrag Potkonjak
Group for Stochastic Algorithmics  http://www.cs.ucla.edu/~miodrag

Introduction: Data Integrity and Computational Sensing

We address two of the canonical problems in sensor networks:
• Data integrity
• Computational sensing.

Due to the large scale and distributed nature of sensor networks, their heterogeneous node structure, cost and power constraints, operation in unpredictable and unconditioned (and often harsh) environmental surroundings and inherent unreliability of sensors sensor networks often collect data with errors, faults and missing samples.

Problem Description: Inter-sensor Modeling and Prediction

• Scientists and application engineers
  - Analysis and synthesis, application
  - Complete, accurate and fault-free datasets
• Sensor network engineers
  - Modeling and management
  - Energy and cost efficient and secure sensor networks
  - Ensure scientists requirements satisfied

We have developed a generic approach for both tasks (i.e. data integrity and computational sensing) that has three phases:
(i) statistical modeling, (ii) prediction, and (iii) fusion and analysis.

Key conceptual novelties include new techniques for capturing of physical hidden covariates, mapping of statistical problem into combinatorial domain, prediction using constraint manipulation, and introduction of class-based networks for analysis of computational complexity of associated optimization problems. We demonstrate that commonly used modeling techniques are inherently inaccurate and that majority of the optimization problems are surprisingly easy to solve.

Experimental Results: Modeling, Prediction, and Optimization

Three phase approach

- Phase-1: Modeling
  - Statistical data-driven models
  - Unique features of computational sensor modeling
- Phase-2: Prediction
  - Constraint manipulation
- Phase-3: Fusion and analysis
  - Formulate data integrity as an optimization problem
  - Objective function: Minimize the discrepancies between the sensor readings and the models
  - Constraints: Model constraints and user’s specified constraints

Inter-sensor Modeling

• Data from sensors X and Y, find the model ŷ=f(x)
• Hidden covariate problem captured by isotonicity constraint:
  - For model f: x_i < x_j ⇒ f(x_i) ≤ f(x_j)
• Univariate CIR (Combinatorial Isotonic Regression):
  - Given data (x_i, ŷ_i), i=1,…,K
  - Given an error measure ε_i and x_i < x_j < x_j < … < x_K
  - ŷ_i isotonic regression is set (x_i, ŷ_i), i=1,…,K s.t.
  - Objective function: min ε_i(x_i, ŷ_i)
  - Constraints: ŷ_i ≤ ŷ_j ≤ … ≤ ŷ_K

Combinatorial Domain-based Statistical Modeling

- Flexibility of the combinatorial modeling
  - Nodes: Different error norms, outlier elimination, robust regression, etc.
  - Edges: Maximum/minimum slope, adding and removing ordering constraints, etc.
  - Paths: Unimodular, convex, locally monotonic, number of break points, etc.
  - Optimization-friendly
- Density estimation, consistency-based techniques

<table>
<thead>
<tr>
<th>Mean Error Bound</th>
<th>Linear</th>
<th>Nonpar</th>
<th>CIR</th>
<th>Lim CIR</th>
<th>Average CPLEX Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0.07</td>
</tr>
<tr>
<td>2%</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>0.31</td>
</tr>
<tr>
<td>3%</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>6</td>
<td>1.16</td>
</tr>
<tr>
<td>4%</td>
<td>4</td>
<td>4</td>
<td>10</td>
<td>9</td>
<td>1.16</td>
</tr>
<tr>
<td>5%</td>
<td>6</td>
<td>8</td>
<td>13</td>
<td>12</td>
<td>0.98</td>
</tr>
<tr>
<td>6%</td>
<td>7</td>
<td>12</td>
<td>24</td>
<td>17</td>
<td>3.22</td>
</tr>
</tbody>
</table>

Sleeping Coordination

- Power law
- Uniform
- Power law

- Constant degree networks
- Random graph networks
  - Suitable for probabilistic analysis, difficult for optimization
- Small world networks
  - Models different phenomena: e.g. social networks, Internet
  - Less suitable for probabilistic analysis
  - Better for optimization that random graphs, but can be also difficult
- Class-based networks
  - Explain a group of phenomena: e.g. genome functional interaction, WADN
  - Less suitable for probabilistic analysis
  - Optimization-friendly