Visualizing and Understanding Large-Scale Bayesian Networks

Michele Cossalter, Ole Mengshoel, Ted Selker

Problem Domain

LARGE HETEROGENEOUS DATASETS

- Some information can be encoded in a node-edge structure, other dimensions of data need a different representation
- Bayesian networks [1]:
  - A directed acyclic graph where nodes represent variables and edges represent dependencies
  - Conditional probability tables (CPTs) describing the probability for the value of each node given the value of its parents
- Other examples: e-mail collections, electrical networks, oil and gas plants, manufacturing units

VISUALIZATION GOALS

- Improved interactive and visual support for debugging the network structure and tuning CPT parameters
- Structure, states, and CPTs are all important
- Scalability as the size and connectivity of the network increase

Current Tools

- CPT BUBBLE ANCHORS
  - Cartoon-inspired bubble lines (B) facilitate association of CPTs (A) with respective nodes
- INTEGRATED TIME SERIES DISPLAY
  - Integration of time series sensor readings together with the CPTs (C), enriched with horizontal lines showing the thresholds used to discretize the states
- CPT MERGING
  - Facilitated comparison of different nodes by allowing data from two nodes to be merged into a single box by drag and drop
- NETWORK OVERVIEW
  - Overview+detail [5]: simultaneously displaying both an overview (D) and detailed view (E) of the network helps orienting

Related NetEx Results

- USER STUDY
  - 25 students with some knowledge of electrical circuits tried to detect faults in the ADAPT network
  - Subjects viewed plots to find which and when components failed in a simple and complex task
  - 12 subjects used NetEx for Task 1 (simple) and a baseline tool (Intelliviz) for Task 2 (complex), 13 subjects did the opposite
  - For both tasks, NetEx users were more accurate than Intelliviz users, with ANOVA p=.061 for Task 1 and p=.075 for Task 2
  - On the complex task Intelliviz is faster (p=.018), but this is not the case for the simple task
  - NetEx users perceived the difficult task to be harder (p=.060) but were more confident about their performance (p=.053) and were more satisfied with the tool (p=.019)

References


Sponsor: NSF, Cylab

Acknowledgements:
This material is based, in part, upon work by Michele Cossalter, Ole J. Mengshoel and Ted Selker supported by NSF grants CCF-0937044 and EEC-0939178.