Inference method for Bayesian networks with imprecise datasets

Presenter: Hector Diego Estrada Lugo
Second year PhD student

Dr E. Patelli, Dr M. de Angelis
Institute for Risk and Uncertainty
University of Liverpool, U.K.

18/07/2018
Motivation
Bayesian nets methodology
  Different data sets implemented
Bayesian Update (Inference)
  Method 1: Application to an OWC system
  Method 2: Application Railway system
Comparison of inference methods
Conclusions
Motivation

• Risk factors representation and uncertainty quantification is complicated in large infrastructure projects.

• Multidisciplinary nature needs a standard tool to facilitate risk communication.

• Risk management must take into consideration the uncertainty factors in the system.
Motivation

• Probabilistic graphical models (like Bayes nets), effective mathematical tool for uncertainty quantification and system modelling.

• Allows to capture variable dependencies of complex systems.

• Inference computation is a key method to update outcomes in Bayesian networks.

• Reliable method of inference computation in Credal networks is necessary.


H.D. Estrada-Lugo
Bayesian Networks

A Bayesian network is a probabilistic graphical model to study and analyse the dependencies of components (random variables) that make up a system.

• The Joint Probability Distribution (JPD) describes entirely network’s dependability,

\[ P(x_i) = \prod_{i=1}^{n} P(x_i | \pi_i) \]

• By introducing evidence, infer updated outcomes.
• Intuitive and relatively easy to implement.
Enhanced Bayesian Networks

Bayesian Networks enhanced* with Structural Reliability Methods (SRM) permit to calculate the conditional probability values of discrete children that come from continuous-parent nodes.

- Calculation of conditional probabilities consist in the approximation of the failure probability.

\[ P(C|B) = \int_{\Omega_{C,b}^c} f(A) dA \]

*f(A): Probability Density Function of continuous node A. \( \Omega_{C,b}^c \) is the domain when C=c in the space of C given B=b.

---

H.D. Estrada-Lugo

Imprecise data sets (discrete): Credal Networks

Generalization of BN to implement imprecise discrete variables in the form of intervals.

• Imprecision is represented through the so called credal sets \( K(x_i) \).

\[
K(x_i) := CH \left\{ P(x_i) \Big| P(x_i) = \prod_{i=1}^{n} P(x_i|x_i) \right\}
\]

• CNs inherent all the probabilistic and graphical characteristics of BNs.

• A CN is a set of BNs, each with different probability values.

Different extreme points combinations make a set of BNs that makes up a CN.
Imprecise datasets (continuous): Probability boxes

A characterization of an uncertain continuous measure in the cumulative distribution space.

• When using SRM failure probability is now represented as:

\[ P_f = \max_\theta \int_{g(x)<0} p(x, \theta)dx \]

• In this way, the continuous probability distributions affected by \textit{aleatoric} and \textit{epistemic uncertainty} are taken into account.
Computational toolbox

OpenCossan

- It takes advantage of Object-Oriented programming in Matlab.
- Parallelization of high demanding tasks.
- Easy connectable with 3rd party toolboxes.
- Excellent platform for EBN.

www.cossan.co.uk
Enhanced BN to Credal nets

Enhanced Bayesian network [*] (Advanced BN)
- Rectangle-Discrete
- Ellipse-Interval
- Circle-Continuous
- Trapezoid- P-box

Credal network [*]

Reduction process

- Rectangle-Interval

Study cases

Oscillating Water Column.
- Probability wave overtopping for different configurations of OWC.
- Experimental data.
- Exact Inference method.

General Railway system.
- Comparison of exact and approximate inference method.
- Probability of having an accident (derailment) due to different rail tracks and train conditions.
- Synthetic data.
Bayesian updating (Inference)

Computation of posterior distribution of a query node given (or not) evidence.

Exact inference methods:
  • Variable elimination (Marginalization).
  • Junction tree algorithm (Clique tree).
  • Recursive conditioning.
  • And/Or search.

Approximate inference.
  • Inner and outer approximation.
  • Importance sampling.
  • Stochastic MCMC simulation.
  • Mini-bucket elimination.
  • Generalized belief propagation.
  • Variational methods.
Exact inference

Computation of posterior distribution of a query node given (or not) evidence.

Exact inference methods:

- Variable elimination (Marginalization).
- Junction tree algorithm (Clique tree).
- Recursive conditioning.
- And/Or search.

Approximate inference:

- Inner and outer approximation.
- Importance sampling.
- Stochastic MCMC simulation.
- Mini-bucket elimination.
- Generalized belief propagation.
- Variational methods.
Inference with intervals

Computation of posterior distribution of a query node given (or not) evidence.

Exact inference methods:
- Variable elimination (Marginalization).
- Junction tree algorithm (Clique tree).
- Recursive conditioning.
- And/Or search.

Approximate inference:
- Inner and outer approximation.
- Importance sampling.
- Stochastic MCMC simulation.
- Mini-bucket elimination.
- Generalized belief propagation.
- Variational methods.
Inference with intervals

It is based on the joint credal set definition to calculate the bounds of the marginal probability as:

\[
P(x_0) := \min_{P(X)\in K(X)} P(x_0) = \min_{P(X_i|\pi_i)\in K(X_i|\pi_i)} \sum_{\pi_i \in \Omega_{\Pi_i}, i=0,\ldots,n} \prod_{i=0}^{n} P(x_i|\pi_i)
\]

\[
\overline{P}(x_0) := \max_{P(X)\in K(X)} P(x_0) = \max_{P(X_i|\pi_i)\in K(X_i|\pi_i)} \sum_{\pi_i \in \Omega_{\Pi_i}, i=0,\ldots,n} \prod_{i=0}^{n} P(x_i|\pi_i)
\]

This represents a non-linear optimization problem with a multilinear objective function. (The head ache of CN inference).
Method 1: Exact inference*

• Take the joint probability distribution function of upper bounds of all the variables in the net. Artificial JPDs are created (each containing a case of the query node).

\[
P(F, S, A) = \begin{bmatrix}
p(F_1, S_1, A_1) & p(F_1, S_2, A_1) \\
p(F_2, S_1, A_1) & p(F_2, S_2, A_1) \\
p(F_1, S_1, A_2) & p(F_1, S_2, A_2) \\
p(F_2, S_1, A_2) & p(F_2, S_2, A_2)
\end{bmatrix}
\]

• Outer approximation is obtained by computing inference in the artificial JPD containing all-upper and all-lower bounds.

Method 1: Exact inference

• Take the joint probability distribution function of upper bounds of all the variables in the net. Artificial JPDs are created (each containing a case of the query node).

\[
P(F, S, A) = \begin{bmatrix}
p(F_1, S_1, A_1) & p(F_1, S_2, A_1) \\
p(F_2, S_1, A_1) & p(F_2, S_2, A_1) \\
p(F_1, S_1, A_2) & p(F_1, S_2, A_2) \\
p(F_2, S_1, A_2) & p(F_2, S_2, A_2)
\end{bmatrix}
\]

• Inner approximation is obtained by finding the artificial JPD that maximizes and minimizes the posterior probability of queried variable.

\[
\begin{align*}
\max \left[ \frac{P(F_1, S, A)}{P(F_2, S, A)} \right] &= \max \left[ \frac{p(F_1, S, A_1)}{p(F_2, S, A_1)} \right. \\
& \left. \frac{p(F_1, S, A_2)}{p(F_2, S, A_2)} \right] \\
\min \left[ \frac{P(F_1, S, A)}{P(F_2, S, A)} \right] &= \min \left[ \frac{p(F_1, S, A_1)}{p(F_2, S, A_1)} \right. \\
& \left. \frac{p(F_1, S, A_2)}{p(F_2, S, A_2)} \right]
\end{align*}
\]
Case study 1: Oscillating water column

- Power generator that converts the energy provided by rise and fall of water inside the column due to sea waves close the shoreline.

- Experimental scaled (1:20) model of an OWC was built in the laboratory.

- To study the hydrodynamic efficiency by the addition of harbour walls to the OWC.

- In a real model, harbour walls would increase probabilities of flooding and station damage.


H.D. Estrada-Lugo

Conventional OWC structure[*].
Case study 1: Oscillating water column

- Power generator that converts the energy provided by rise and fall of water inside the column due to sea waves close the shoreline.

- Experimental scaled (1:20) model of an OWC was built in the laboratory.

- To study the hydrodynamic efficiency by the addition of harbour walls to the OWC.

- In a real model, harbour walls would increase probabilities of flooding and station damage.

- Experimental layout of OWC with harbour walls[*].


H.D. Estrada-Lugo
Case study: Addition of harbour walls to an OWC*


H.D. Estrada-Lugo
Case of study 1: OWC

Results

Harbour wall inclination top view.

Wave overtopping

Harbour wall top view.

Wave overtopping
Method 1: Exact inference

✓ This method is computationally cheap.
✓ Reliable when **extreme scenarios** are of the interest.
✓ Uncertainty attached to the bounds provided.

- Boolean variables.
- Overestimation of upper bounds.
- Underestimation of lower bounds.
- Not suitable for large networks, number of inference computations increase as $2^n$. 
Method 2: Approximate inference

• Approximate inference with Linear programming. Optimization task.
• Reduce credal sets to singletons called Extreme Points $\hat{P}(X_i|\pi_i) \in \text{ext}[K(X_i|\pi_i)]$
different from the Free variable $X_j$.

So the constrained queried-variable ($x_0$) lower bound is:

$$P'(x_0) := P(X_j|\pi_j) \min_{\pi_j \in \Omega_{\Pi_j}} \sum_{x_j, \pi_j} \left[ \tilde{P}(x_0|x_j, \pi_j) \cdot \tilde{P}(\pi_j) \right] \cdot P(x_j|\pi_j)$$

Linear combination of $X_j$ local probabilities.

Method 2: Approximate inference

\[ P'(x_0) := P(X_j | \pi_j) \min_{\pi_j \in \Omega_{\pi_j}} \sum_{x_j, \pi_j} \left[ \bar{P}(x_0 | x_j, \pi_j) \cdot \bar{P}(\pi_j) \right] \cdot P(x_j | \pi_j) \]

- Iterations over Xj are done to perform a local search.
- Once an approximation (extreme point) to the optimal solution is calculated. The Xj variable released and a new Xj is used as the free variable.
- The programme stops iterating when no further improved approximation is found.

Method 2: Approximate inference

- $P'(x_0)$ upper approximation of lower probability bound $P'(x_0)$ of the CN.
- $\overline{P}'(x_0)$ is lower approximation of the upper bound $\overline{P}'(x_0)$ of the CN.
Method 2: Approximate inference

- $P'(x_0)$ upper approximation of lower probability bound $P'(x_0)$ of the CN.
- $P'(x_0)$ is lower approximation of the upper bound $P'(x_0)$ of the CN.
Case of study 2: Railway system

Derailment probability, taking into account:

- Obstructions in the railway due to:
  - Earthworks
  - Terrain
- Train speed.
- Damage in the tracks.
Results

- Embankment slope over which the rail tracks are placed.
- Terrain quality depending on:
  - Earthworks
  - Cut slopes
  - Embankment slope steepness
- Derailment, due to factors:
  - Final train speed
  - Track obstructions
  - Track defects
Results

- Embankment slope over which the rail tracks are placed.
- Terrain quality depending on:
  - Earthworks
  - Cut slopes
  - Embankment slope steepness
- Derailment, due to factors:
  - Final train speed
  - Track obstructions
  - Track defects

### Embankment slope vs. Terrain quality

- Approx
- Exact

<table>
<thead>
<tr>
<th>Embankment slope</th>
<th>Terrain quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steep</td>
<td>Good, Bad</td>
</tr>
<tr>
<td>Gradual</td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Bad</td>
<td></td>
</tr>
</tbody>
</table>

H.D. Estrada-Lugo
Results

- Embankment slope over which the rail tracks are placed.
- Terrain quality depending on:
  - Earthworks
  - Cut slopes
  - Embankment slope steepness
- Derailment, due to factors:
  - Final train speed
  - Track obstructions
  - Track defects

H.D. Estrada-Lugo
Results

- The number of nodes in the network increases exponentially, the computational time needed to compute inference with exact inference.

Computation done with machine with an Intel i5 @3.2GHz, 8Gb RAM.
Method 2: Approximate inference

✓ Does not suffer from large credal sets.
✓ Follows the same topology of BN.
✓ Does not require to indicate the extreme points.
✓ It can be used with variables with many states and/or parents.
✓ Provides inner approximate solutions.
✓ Fast and accurate.

- Local credal sets specified by lean constraints.
- Not for local credal sets given by explicit enumeration of the extreme points.
- Outer approximations are currently excluded.
- A combination of inner with outer approximations can bring reliable inferences.
Conclusions

- Two different inference computation methods were tested to compare their performance.
- Using optimization methods reduces substantially notably inference computation on CN.
- Approximate inference methods prevents from combinatorial explosion.
- The use of interval probabilities allows to consider a broader range data types (imprecise data sets).
- Imprecise probabilities allows to take into account epistemic uncertainty due to the vagueness or lack of data.
- This model can be applicable to different complex technological facilities.
- Work is carried out to provide a reliable method to provide an outer approximation of the probability bounds and study convergence.
Inference method for Bayesian networks with imprecise datasets

Hector Diego Estrada Lugo
h.d.estrada-lugo@liverpool.ac.uk

https://www.liverpool.ac.uk/risk-and-uncertainty/