Dependence modelling and quantification methods

Application to civil infrastructure reliability and maintenance

Alex Kosgodagan
PhD student

Outline

Bayesian networks

- ↳ Non-parametric BNs
- ↳ Limited Memory Influence diagrams

Structured Expert Judgment

- ↳ Cooke's method for combined expert opinion and measures of dependence

Why BNs in this context ?

- Handle randomness → physical quantities impacting degradation can behave randomly
- Handle probabilistic dependencies → account for dependencies/correlations between these quantities
- Adequacy/intuitive visual representation between network and the graph pattern in BNs
- Dynamically propagate evidence → update forecasts

Non-parametric Bayesian networks

- BN for continuous variables (also referred to as pair-copula BNs)
- “Non-parametric” because it relaxes the joint normality assumption
- Dependence induced by copulae
- Copulae parametrized by rank/product moment correlations (Hanea 2006, Morales-Napoles 2014 application bridge reliability)
- Copula assumption (has to be validated) → necessary condition: zero correlation means independence (e.g. Gaussian)

Limited Memory Influence diagrams

- Compact graphical representations of decision-making problems under uncertainty (Howard & Matheson 1984, Lauritzen & Nilsson 2001)
- Solving a LIMID refers to:

“finding an optimal plan of action, that is, a combination of decision rules, or policies, that associate any possible observation to an action; by optimality we mean maximizing/minimizing the total expected utility”

Intractable when the LIMID gets too complex (number of nodes and arcs) → use NPBNs to overcome the dimensionality issue (work in progress)

Limited Memory Influence diagrams

For a discrete LIMID $\mathcal{L}$

- $\mathcal{C}, \mathcal{D}$ and $\mathcal{U}$ sets of chance, decision and utility nodes
- pa_X sets of parents and children for node X, $X \in \Omega_X$
- for each $C \in \mathcal{C}$, associate $\left\{p_C^\pi : \pi \in \Omega_{pa_C}\right\}$ conditional pdf; if $pa_C = \emptyset$, C has a single pdf assigned

Policies and strategies

- For each $D \in \mathcal{D}$, a policy $\delta_D : \Omega_{pa_D} \rightarrow \Omega_D$ and $\Delta_D = \cup_{D\in\mathcal{D}}\, \delta_D$
- Let $\Delta = \times_{D\in\mathcal{D}} \Delta_D$ denote the space of possible combination of policies
- An element $s = (\delta_D)_{D\in\mathcal{D}} \in \Delta$ is said to be a strategy for $\mathcal{L}$.

Limited Memory Influence diagrams

A strategy s induces a joint probability mass function over the variables in $\mathcal{C} \cup \mathcal{D}$ by

$$p_s = \prod_{C \in \mathcal{C}} p_C^{pa_C} \prod_{D \in \mathcal{D}} p_D^{pa_D}$$

with expected utility

$$E_s = \sum_{s \in \mathcal{C} \cup \mathcal{D}} p_s \sum_{v \in \mathcal{U}} u_v$$

Global maximum strategy $\bar{s}$

$$E_{\bar{s}} \geq E_s, \text{ for all strategies } s \in \mathcal{C} \cup \mathcal{D}$$

Limited Memory Influence diagrams

Set

$$s_{-d_0} = \{\delta_d : d \in \mathcal{D} \backslash \{d_0\}\}, \forall\, d_0 \in \mathcal{D}$$

and

$$\delta'_d * s = \{\delta'_d\} \cup s_{-d},$$

Local maximum policy

$$E(\delta_d * s) = \sup_{\delta'_d} E(\delta'_d * s)$$

Local maximum strategy $\bar{s}$ verifies

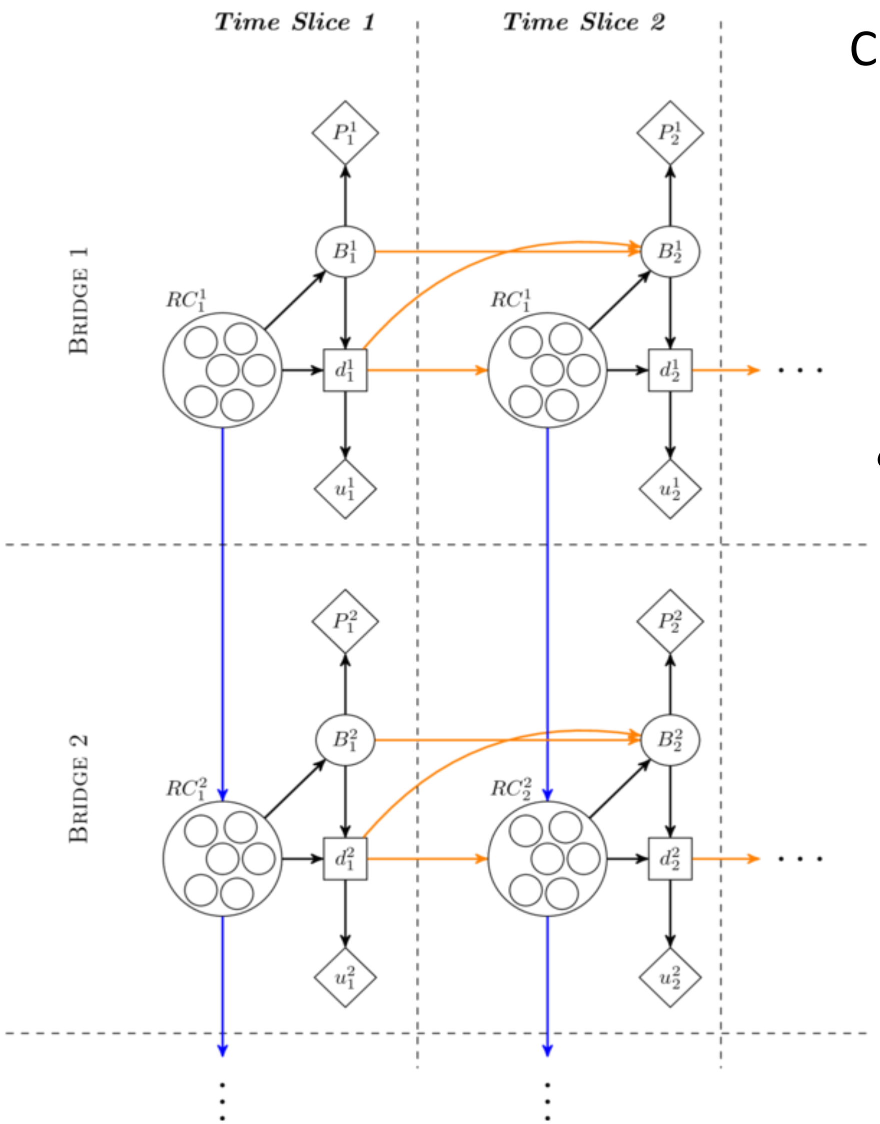
$$E(\bar{s}) \geq E\left(\delta_d \; * \bar{s}\right), \text{for all } d \in \mathcal{D}, \delta_d \in \Delta$$

Limited Memory Influence diagrams

SPU (single policy update) Algorithm (Lauritzen & Nilsson 2001):

1. Transform the network into a junction tree (undirected graph) → allows for message passing algorithm (Pearl 1988) using potential and cliques (all nodes are mutually connected)
   - Standard algorithms for exact inference in discrete BNs
2. Initialize a strategy $s^0 \in \Delta$ (e.g. uniform distribution)
3. At each iteration i, j take $s^i = \delta_{d_j} * s^{i-1}$ and compute E_{s^i} → if $E_{s^{i+1}} > E_{s^i}$, keep $E_{s^{i+1}}$ otherwise keep E_{s^i}
4. Output a local maximum strategy $\bar{s}$
5. Keep iterating until finding a global maximum strategy

Limited Memory Influence diagrams



CPT for node B_t^k

$$P\left(B_{t+1}^k = j\right) = \begin{cases} P\left(B_t^k = j\right) + \phi_{1,j}^k(t+1) - \phi_{1,j+1}^k(t+1)\mathbf{1}_{j<n} \\ P\left(B_0^k = j\right) = 0 \end{cases}$$

where

$$\phi_{i,j}^k(t) = P\left(B_t^k = j, B_{t-1}^k \neq j, \dots, B_1^k \neq j \middle| B_0^k = i\right)$$

Conditional probability table for node B_t^k $\forall k, t > 1$

$$P\left(B_{t+1}^k = i | B_t^k = j, d_t^k = l\right)$$

$$= \begin{cases} 1 & if\ i = 1; j = 1,2; l = 3,4 \\ P(B_t^k = i) & if\ i = j; l = 1 \\ 1 - P(B_t^k = i) & if\ j = i + 1; l = 1 \\ 1 & if\ i = j = 4; l = 1 \\ \dots & \dots \end{cases}$$

TNO innovation for life

MINES Nantes

Limited Memory Influence diagrams

Unconditional policy for B1

$3 \leq t \leq 17$	$Load_t^1 = 1$				$Load_t^1 = 2$			
B_t^1	1	2	3	4	1	2	3	4
$d_t^1 = 1$	1	0	0	0	1	0	0	0
$d_t^1 = 2$	0	1	1	0	0	1	1	0
$d_t^1 = 3$	0	0	0	1	0	0	0	1
$d_t^1 = 4$	0	0	0	0	0	0	0	0

Policy for every bridge conditional on *Standstill* Traffic density in B1

$11 \leq t \leq 17$	$Load_t^1 = 1$				$Load_t^1 = 2$			
B_t^1	1	2	3	4	1	2	3	4
$d_t^1 = 1$	1	0	0	0	1	0	0	0
$d_t^1 = 2$	0	1	0	0	0	1	1	0
$d_t^1 = 3$	0	0	1	1	0	0	0	1
$d_t^1 = 4$	0	0	0	0	0	0	0	0

Policy for bridge 2 conditional on *Standstill* Traffic density in B1

$15 \leq t \leq 17$	$Load_t^2 = 1$				$Load_t^2 = 2$			
B_t^2	1	2	3	4	1	2	3	4
$d_t^2 = 1$	1	0	0	0	1	0	0	0
$d_t^2 = 2$	0	1	0	0	0	1	1	0
$d_t^2 = 3$	0	0	1	1	0	0	0	1
$d_t^2 = 4$	0	0	0	0	0	0	0	0

Structured expert judgment

Cooke's classic method for combined expert opinion
seed variables → measures of performance

- Information: $p = (p_1, \dots, p_n)$ be the subjective distribution over alternatives $\{1, \dots, n\}$ and $s = (s_1, \dots, s_n)$ be the « real » distribution

$$I(s,p) = \sum_{i=1}^{n} s_i \ln \frac{s_i}{p_i}$$ (derived from Shannon's entropy)

- Calibration: let s denote a sample distribution generated by N independent samples from the distribution p

$$C(e) = 1 - \chi_d^2[2NI(s,p)]$$

Different possibilities to weigh experts (global, equal,…)
Robustness tests

Structured expert judgment

Example of output (Excalibur)

Nr.	Id	Calibr.	Mean relative total	Mean relative realization	Numb real	UnNormalized weight	Normaliz.weigh without DM	Normaliz.weig with DM
1	Exp1	1,95E-010	1,125	1,256	15	2,449E-010		5,089E-009
2	Exp2	1,254E-007	1,21	1,442	15	1,808E-007		3,756E-006
3	Exp3	0,0002236	0,9629	1,081	15	0,0002416		0,005021
4	Exp4	1,806E-008	0,93	1,013	15	1,829E-008		3,801E-007
5	Exp5	7,576E-008	1,319	1,471	15	1,114E-007		2,315E-006
6	Exp6	5,731E-012	0,8529	1,049	15	6,01E-012		1,249E-010
7	Exp7	0,0008472	0,6003	0,6605	15	0,0005596		0,01163
8	Exp8	4,458E-005	1,017	1,045	15	4,657E-005		0,0009677
9	IT	0,1272	0,3331	0,3718	15	0,04727		0,9824
10	GL	0,1272	0,2854	0,3111	15	0,03955		0,979
11	EQ	0,09153	0,1596	0,1745	15	0,01597		0,9496

Work in progress: elicitation on bridges reliability and maintenance

TNO innovation for life

MINES Nantes

Structured expert judgment

Goal : assess a dependence structure for an NPBN

2 ways
- ask conditional probability of exceedance: $P(X_1 < 50^{th} | X_2 < 50^{th})$
- ask ratios of rank correlations (Morales *et al.* 2014)

$dCal(e)$ → measure the distance between two correlation matrix

$$dCal(e) = 1 - d_H(\Sigma_e, \Sigma_T)$$

where

$$d_H(\Sigma_e, \Sigma_T) = \sqrt{1 - \eta(\Sigma_e, \Sigma_T)}$$

is denoted as the Hellinger distance between correlation matrix of 2 Gaussian copulae and

$$\eta(\Sigma_1, \Sigma_2) = \frac{|\Sigma_1|^{\frac{1}{4}} |\Sigma_2|^{\frac{1}{4}}}{|\frac{1}{2}\Sigma_1 + \frac{1}{2}\Sigma_2|^{\frac{1}{2}}} \exp\left\{-\frac{1}{8}(\mu_1 - \mu_2)^T \left(\frac{1}{2}\Sigma_1 + \frac{1}{2}\Sigma_2\right)^{-1} (\mu_1 - \mu_2)\right\}$$

Structured expert judgment

Show that $dCal(e)$ is a proper scoring rule (« how to keep the expert honest »)

For « true » probabilities $\mathbf{p} = (p_1, \dots, p_n)$ and subjective probabilities $\mathbf{q} = (q_1, \dots, q_n)$, f_k is a proper scoring rule if

$$\sum_{k=1}^{n} f_k(q_k)p_k \leq \sum_{k=1}^{n} f_k(p_k)p_k$$

Shanon's entropy is a proper scoring rule (McCarthy 1954)

Hints for the proof

Convexity arguments (Boutilier 2012)

$$f(\mathbf{p}, x_i) = G(\mathbf{p}) - G^*(\mathbf{p}) \cdot \mathbf{p} + G_i(\mathbf{p})$$

Where G is a convex function and G^*its subgradient

References

Boutilier, *Eliciting Forecasts from Self-interested Experts: Scoring Rules for Decision Makers,* Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012), Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

Hanea, *Algorithms for non-parametric Bayesian Belief Nets, PhD Thesis, 2007.*

Howard & Matheson, *Influence diagrams,* Decision Analysis Vol. 2, No. 3, September 2005, pp. 127–143

Lauritzen & Nilsson, *Representing and solving decision problems with limited information,* Management Science Vol. 47, No. 9, September 2001 pp. 1235–1251.

McCarthy, *Measures of the Value of information,* Vol. 42, pp. 654-655

Morales-Napoles et al., *Large-Scale Hybrid Bayesian Network for Traffic Load Modeling from Weigh-in-Motion System Data, 2014*

Morales-Nápoles, Joaquín Delgado-Hernández, De-León-Escobedo & Arteaga-Arcos, *A continuous Bayesian network for earth dams' risk assessment: methodology and quantification,* Structure and Infrastructure Engineering, Vol. 10, No. 5, 589–603, 2014.

Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann 1988

alex.kosgodaganacharige@tno.nl

Limited Memory Influence diagrams

Assumptions :

- Same as for Model 1 except for the transitions

→ time-dependent (Dynamic BN)

- Maintenance is considered to be imperfect (A: set of actions available):

given a bridge being in a state i, we assume that each action type r has probability p_r to reach state $i - r + 1$, and probability $1 - p_r$ to reach state i–r+2, with $1 < i, r \leq n$ and $r \leq i$. Note that for the case $i > r, p_r = 1$ to reach state 1.

→ when a maintenance action is performed the structure does not jump to state Green with prob. 1

- all actions take negligible time with respect to our time

Limited Memory Influence diagrams

Example of imperfect maintenance effects:

A = $\{1,2,3,4\}$; $\Omega = \{\mathrm{G}, \mathrm{Y}, \mathrm{O}, \mathrm{R}\}$ and a bridge B is in condition Orange at time *t*

→ Action type 1 is similar to “Do nothing”; B remains in O at time t+1
→ Action type 2 has probability p_2 to make B jump in condition Y and probability $1 - p_2$ for B to stay in O at time t+1
→ Action type 3 has probability p_3 to make B jump in condition G and probability $1 - p_3$ make B jump in condition Y at time t+1
→ Action type 4 has probability $p_4 = 1$ to make B jump in condition G at time t+1

Limited Memory Influence diagrams

Example of variable domain and distribution

Variable	States	Conditional probability distribution
B_0^k	1 - Perfect 2 - Fair 3 - Bad 4 - Poor	$P(B_0^k = i) = \begin{cases} 1 & if\ i = 1 \\ 0 & otherwise \end{cases}$
B_t^k	1 - Perfect 2 - Fair 3 - Bad 4 - Poor	$P(B_{t+1}^k = i \mid B_t^k = j, d_t^k = l) = \begin{cases} 1 & if\ i = 1; j = 1{,}2; l = 3{,}4 \\ P(B_t^k = i) & if\ i = j; l = 1 \\ 1 - P(B_t^k = i) & if\ j = i + 1; l = 1 \\ 1 & if\ i = j = 4; l = 1 \\ 1 & if\ i = 1; l = 2 \\ p_2 & if\ i = j + 1; l = 2 \\ 1 - p_2 & if\ i = j, l = 2 \\ p_3 & if\ i = j + 2, l = 3 \\ 1 - p_3 & if\ i = j + 1, l = 3 \\ 0 & otherwise \end{cases}$

TNO innovation for life

MINES Nantes

Limited Memory Influence diagrams

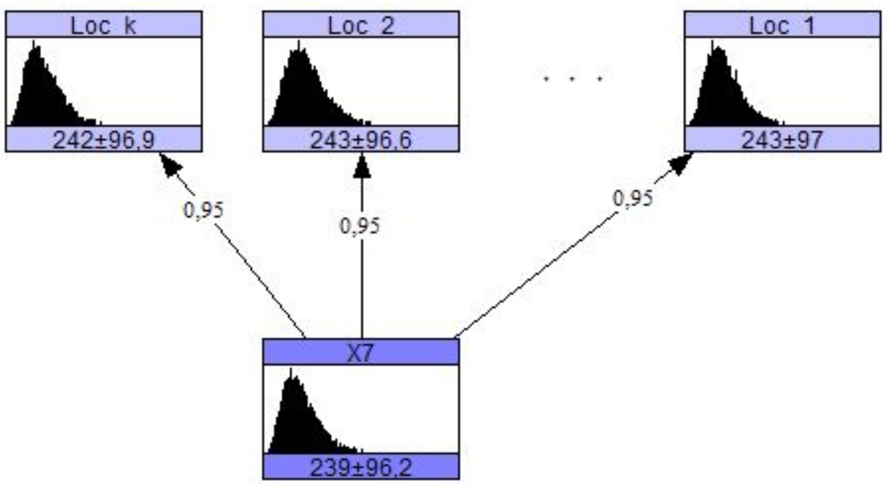
- Exportable to a different set of assets
- Attractiveness of inference in (un)conditional policies
- According to the sensitivity of inference, inserted evidence coming from monitoring could be optimized

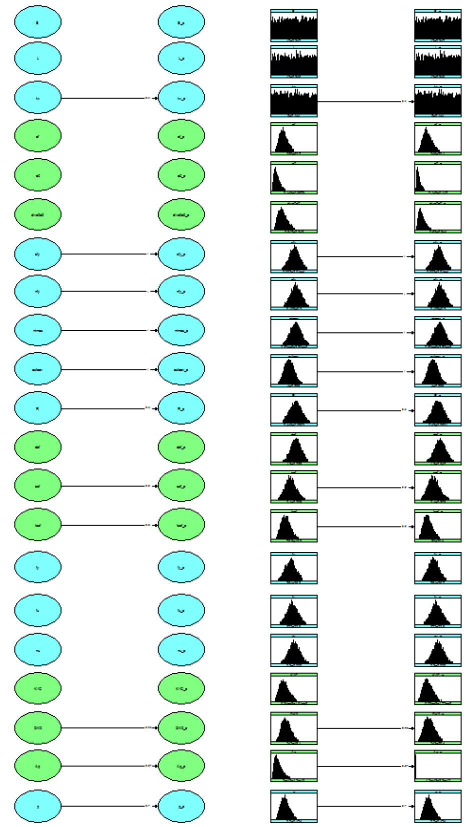
→ Research on coupling another version of BN (NPBNs) with the SPU algorithm to overcome the combinatorial issue

Non-parametric Bayesian networks

21 variables with various distributions

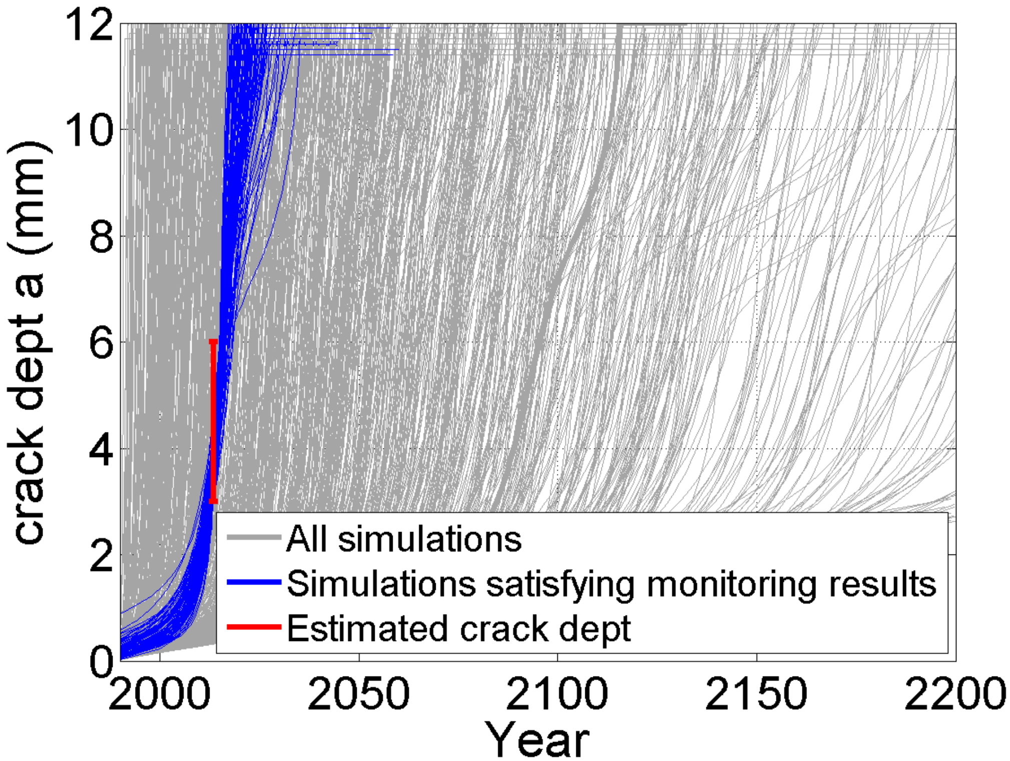
Layout of monitored and non monitored locations





Non-parametric Bayesian networks

Monte Carlo samples to assess crack depth



TNO innovation for life

MINES Nantes

Junction tree algorithm

- Moralisation

→ Add an edge between every pair of parent nodes

- Deletion

→ delete all directions on arcs

- Triangulation

→ Convert a network of nodes into a network of cliques

→ Message passage algorithm can be used

Expert judgment

Global weight
The information score is based on all the assessed seed items.

Item weight
Using information scores for each item rather than the average information score

Cooke
“ The decision maker would perform better than the result of simple averaging, called the equal weight DM, and we should also hope that the proposed DM is not worse than the best expert in the panel.”

General conclusions

- Applicable to a different set of assets

→ Only have to comply with the BN requirements

- Importance of monitoring in propagated evidence

→ Monitoring installation can be expensive and BNs' inference property allows to select elements in order to update forecasts for the whole network

- Important drawback concerns time-consuming computations related to BNs complexity

→ Can be overcome with other versions of BNs (non-parametric for instance)