Plus court chemin stochastique et application à la maintenance

F. Corset[1] S. Chrétien[2]

[1]LabSAD
Université Pierre Mendes France

[2]Département de Mathématiques
Université de Franche-Comté

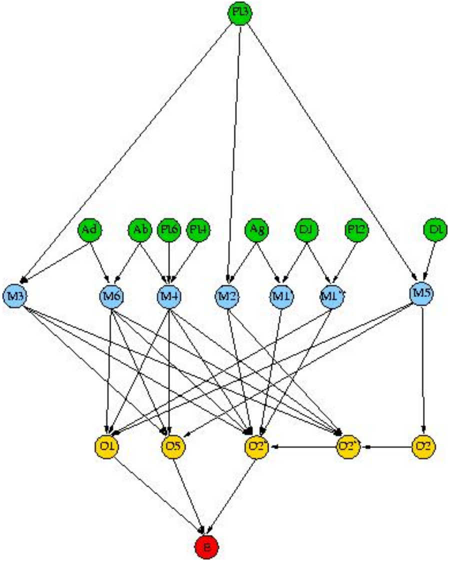
FIMA, 7 juillet 2005

Plan de l'exposé

Origine du problème

- Modélisation de dégradation d'un système par un réseau bayésien (RB)
- RB (graphe orienté sans cycle) où les noeuds sont les états de dégradations du système
- Arcs représentent la probable causalité entre deux dégradations
- Probabilités conditionnelles d'un état sachant les états parents
- Simulateur de scénarii, outil de diagnostic, outil d'aide à la décision

Exemple de RB pour EDF



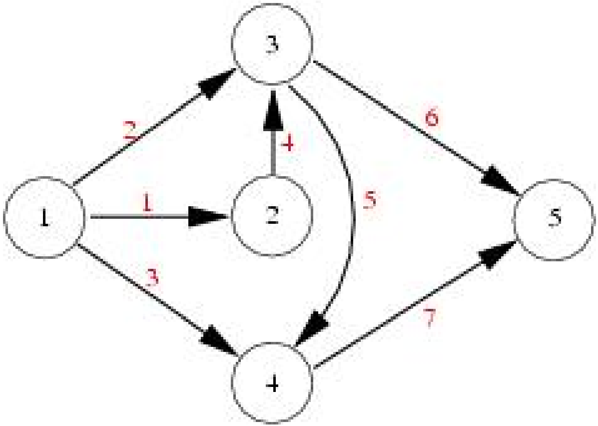
Présentation du problème

- Reprise du même graphe orienté sans cycle
- Sommet 1 représente l'état neuf du système
- Sommet *m* représente un état de dégradation inacceptable
- Le système se dégrade d'un état vers un autre, i.e. d'un sommet vers un sommet voisin
- Le temps de transition entre deux états (pondération de l'arc *i*) est une variable aléatoire $\sim \mathcal{W}(\eta_i, \beta_i)$
- Les v.a. sont indépendantes
- Politique de maintenance : le système doit être inspecté avant d'atteindre le sommet *m* :
  Plus court chemin dans un graphe

Présentation du problème

- Reprise du même graphe orienté sans cycle
- Sommet 1 représente l'état neuf du système
- Sommet *m* représente un état de dégradation inacceptable
- Le système se dégrade d'un état vers un autre, i.e. d'un sommet vers un sommet voisin
- Le temps de transition entre deux états (pondération de l'arc *i*) est une variable aléatoire $\sim \mathcal{W}(\eta_i, \beta_i)$
- Les v.a. sont indépendantes
- Politique de maintenance : le système doit être inspecté avant d'atteindre le sommet *m* :
  Plus court chemin dans un graphe

Un petit exemple



Matrice d'incidence du graphe

Soit $G = (V, E)$ le graphe orienté sans cycle avec m sommets (ici 5) et n arcs (ici 7), et A sa matrice d'incidence

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 \end{pmatrix}$$

Theorem

Soit k le nombre de composantes connexes. Alors $rg(A) = m - k$.

Matrice d'incidence du graphe

Soit $G = (V, E)$ le graphe orienté sans cycle avec m sommets (ici 5) et n arcs (ici 7), et A sa matrice d'incidence

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 \end{pmatrix}$$

Theorem

Soit k le nombre de composantes connexes. Alors $rg(A) = m - k$.

Plus court chemin comme programme linéaire

Soit le vecteur $(c_i)_{i=1,\dots,n}$ de coûts sur les arcs et soit $b = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \end{pmatrix}^t$, le problème de plus court chemin s'écrit

$$\begin{aligned} z = \quad \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

x^* et z^* solution du problème donnent respectivement le meilleur chemin et le coût de ce chemin.

Rappel sur la théorie de la programmation linéaire

Soit x une solution faisable (satisfaisant $Ax = b$ et $x \geq 0$) :

$$\begin{cases} x_i = 0 \text{ forall } i \notin B \text{ and} \\ x_B = A_B^{-1} b \end{cases} \tag{1}$$

où x_B est le vecteur composé des coordonnées indicées par B et A_B la sous matrice de A indicée par B.

Lemma

Soit B une partie de $\{1, \ldots, n\}$ tel que $A_B^{-1} b \geq 0$ et $N = \{1, \ldots, n\} \setminus B$. Un vecteur x sera solution de (2) ssi $c_i \geq c_B^t A_B^{-1} a_i$ pour tout $i \notin B$ où a_i est la $i^{\text{ème}}$ colonne de A.

Inégalité de Dyer-Frieze-McDiarmid (1)

Soit le programme linéaire suivant

$$\begin{aligned} z^\star = \quad \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{2}$$

où c_i sont des v.a. indépendantes, et $A \in \mathbb{R}^{m \times n}$ est une matrice de **rang plein**.

Inégalité de Dyer-Frieze-McDiarmid (2)

Theorem

Supposons que les variables de coûts c_i soient indépendantes et satisfassent

$$\mathbb{E}[c_i \mid c_i \geq h] \geq \mathbb{E}[c_i] + \alpha h$$

pour $\alpha \in (0, 1]$. Alors pour chaque matrice $A \in \mathbb{R}^{m \times n}$ et chaque vecteur $b \in \mathbb{R}^m$, la valeur optimale z^* du problème (2) satisfait

$$\mathbb{E}[z^*] \leq \frac{1}{\alpha} \max_{S\,:\,\#S=m} \sum_{i \in S} \mathbb{E}[c_i] x_i \qquad (3)$$

pour chaque solution faisable x.

Inégalité de Dyer-Frieze-McDiarmid (3)

- pour une loi exponentielle $\alpha = 1$
- pour la loi uniforme sur $[0, 1]$, $\alpha = \frac{1}{2}$
- $\Rightarrow$ donne une borne supérieure pour la moyenne du temps d'inspection
- adaptation pour le problème de plus court chemin ($rang(A) = m - 1$)
- adaptation pour la loi de Weibull
- borne inférieure sur l'espérance du meilleur coût

Matrice d'incidence modifiée

Soit $\tilde{A}$ la matrice d'incidence modifiée :

$$\tilde{A} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & -1 & 1 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 & -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 \end{pmatrix}.$$

Alors, le problème du plus court chemin est équivalent au programme linéaire suivant :

$$\begin{aligned} z^\star = \quad \min \quad & \tilde{c}^T x \\ \text{s.t.} \quad & \tilde{A}x = b \\ & x \geq 0 \end{aligned} \tag{4}$$

où $\tilde{c}$ est le vecteur des coûts auquel on a ajouté un élément nul 0.

La loi de Weibull

Soit X une variable aléatoire distribuée selon une loi de Weibull

$$f_X(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta} e^{-(\frac{t}{\eta})^{\beta}}.$$

Alors, le temps moyen résiduel avant une défaillance (MRTF) est donné par

$$G_X(h) \triangleq E[X \mid X \geq h] = \eta e^{(\frac{h}{\eta})^{\beta}} \Gamma(1 + \frac{1}{\beta}, (\frac{h}{\eta})^{\beta}),$$

où $\Gamma(a, h)$ est la fonction gamma incomplète définie par

$$\Gamma(a, h) = \int_h^{+\infty} t^{a-1} e^{-t} dt.$$

La loi de Weibull (2)

Lemma

Si $\beta \geq 1$ alors

$$\lim_{h \to 0} G'_X(h) = 0 \text{ and } \lim_{h \to +\infty} G'_X(h) = 1$$

et de plus si $\beta < 2$ nous avons

$$G''_X(h) > 0, \qquad \lim_{h \to 0} G''_X(h) = +\infty \text{ and } \lim_{h \to +\infty} G'_X(h) = 0.$$

Theorem

Supposons que X soit distribué selon une Weibull avec $\beta \in]1, 2[$. Alors pour tout $h \geq 0$, on a

$$E[X \mid X \geq h] \leq E[X] + h.$$

La borne inférieure

Theorem

$$\sum_r p_r \sum_{i \in B_r} E[c_i] x_i \leq E[z^*],$$

où (B_r) est la famille de toutes les bases et pour tout r, p_r est la probabilité que B_r soit optimale.

- Difficulté d'appliquer cette borne
- nécessaire d'estimer tous les p_r!
- $\Rightarrow$ Utilisation des bases optimales qui apparaissent le plus fréquemment.

La borne inférieure en pratique

Theorem

Soit $\hat{p}_B$ la borne inf de la probabilité que B soit une base optimale avec un niveau de confiance de $1-\alpha$. Alors avec la probabilité $1-\alpha$ nous avons

$$\hat{p}_B E[c]^t x \leq E[z^*].$$

- Pour η et β fixés, on résoud N problèmes de plus court chemin,
- N bases optimales avec une proportion p_B,
- Etant donné un risque α, nous sous-estimons p_B

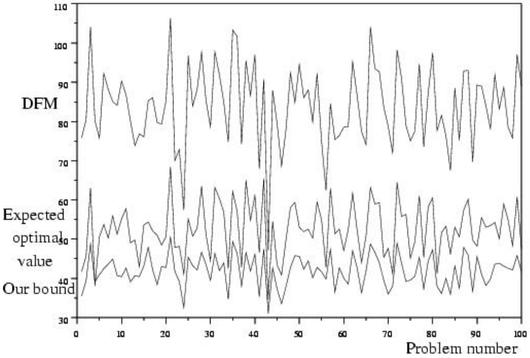
$$\hat{p}_B = p_B - q_{1-\alpha} * \sqrt{p_B(1-p_B)/N},$$

  avec $q_{1-\alpha}$ le quantile de la loi normale.
- Calcul de la borne inférieure correspondante.

Simulations

- Pour un même graphe,
- pour le problème $i = 1, \ldots, 100$, simulation de $\eta(i)$ suivant une loi normale $\mathcal{N}(50, 100)$ et $\beta(i)$ suivant une loi uniforme sur $[1, 2]$.
- Le coût c_j de l'arc j suit une loi de Weibull $W(\eta(i), \beta(i))$.
- Espérance du coût optimal par des simulations de Monte Carlo sur 1000 échantillons.
- Calcul de $\hat{p}_r$ avec un risque $\alpha = 5\%$.
- Calcul de la borne inférieure $\hat{p}_B E[c]^t x$.

Simulations



Théorème de Berry-Essen

Theorem

Soit $X_1, X_2, \ldots, X_n$ n variables aléatoires indépendantes d'espérance nulle et de variance finie. Soit $S_n = \sum_{i=1}^n X_i$ et $B_n^2 = \sum_{i=1}^n \mathrm{E}X_i^2$ et supposons que $E|X_i|^3 < +\infty$. Alors

$$\sup_z \left|P\left(\frac{S_n}{B_n} \leq z\right) - \Phi(z)\right| \leq 0.7975 B_n^{-3} \sum_{i=1}^n \mathrm{E}|X_i|^3, \qquad (5)$$

où Φ est la f.d.r la loi normale centrée réduite.

Une borne inférieure avec Berry-Essen (1)

Lemma

Supposons que les coûts possèdent des moments d'ordre 3 finis. Soit B une base et pour tout $j \in B$ et $i \in B^c$, soit $\alpha_{ji} = \left((A_B)^{-1}A_{B^c}\right)_{ji}$. Alors la probabilité p_B que B soit optimale vérifie

$$
\begin{aligned}
p_B \quad &\geq \quad 1 - \sum_{j \in B^c} \Phi\left(\frac{\mu_i - \sum_{j\in B} \alpha_{ji}\mu_j}{(\sigma_i^2 + \sum_{j\in B} \alpha_{ji}^2 \sigma_j^2)^{\frac{1}{2}}}\right) \\
&+ \quad 0.7975 \frac{\mathrm{E}|c_i - \mu_i|^3 + \sum_{j\in B}^{n} \alpha_{ji}^3 \mathrm{E}|c_j - \mu_j|^3}{(\sigma_i^2 + \sum_{j\in B} \alpha_{ji}^2 \sigma_j^2)^{\frac{3}{2}}}.
\end{aligned}
$$

Une borne inférieure avec Berry-Essen (2)

Theorem

$$
\begin{aligned}
\mathrm{E}[z] \quad & \geq \quad \sum_{B \in \mathcal{B}} \Bigg(1 - \sum_{j \in B^c} \Phi \Bigg(\frac{\mu_i - \sum_{j \in B} \alpha_{ji} \mu_j}{(\sigma_i^2 + \sum_{j \in B} \alpha_{ji}^2 \sigma_j^2)^{\frac{1}{2}}} \Bigg) \\
& + \quad 0.7975 \frac{\mathrm{E}|c_i - \mu_i|^3 + \sum_{j \in B}^n \alpha_{ji}^3 \mathrm{E}|c_j - \mu_j|^3}{(\sigma_i^2 + \sum_{j \in B} \alpha_{ji}^2 \sigma_j^2)^{\frac{3}{2}}} \Bigg) \mathrm{E}[c_B^t] x_B .
\end{aligned}
$$

Conclusions et perspectives

- Notre borne inférieure est "meilleure" que la borne supérieure DFM (en valeur absolue).
- Erreur moyenne relative inférieure à 20%.
- De bons résultats expérimentaux.
- Application sur des systèmes réels.
- Simulations avec Berry-Essen

Références I

Steele, J. M. (1997). *Probability Theory and Combinatorial Optimization*. CBMS-NSF Conference Series in Applied Mathematics 69, SIAM.

Dyer, M.E., Frieze, A.M., and McDiarmid, C.J.H. (1986). *On linear programs with random costs*. Math. Programming **35**, 3 – 16.