

Réseau bayésien pour la surveillance à base de mesures et de résidus

ATOUI, Mohamed Amine
Mél : mohamed.atoui@univ-angers.fr

Résumé : Dans cet article, nous présentons une méthode originale de surveillance des procédés permettant sous un même outil de faire cohabiter résidus et mesures prises sur le système à la fois pour la détection et le diagnostic de fautes. L'outil permettant cela est un réseau bayésien. La méthode est évaluée sur une simulation d'un système de chauffe-eau sous différentes hypothèses.

Mots clés : *Surveillance, méthodes statistiques, méthodes basées modèles, réseaux bayésiens*

1 Introduction

La course des entreprises et industries à optimiser le rapport gain/perte ainsi que la croissance de la demande pour des produits de qualité ont énormément contribué à l'évolution et à l'utilisation impérieuse de méthodes de surveillance parmi d'autres moyens de sûreté de fonctionnement. Des méthodes qui s'avèrent de nos jours, nécessaires pour détecter et diagnostiquer l'état (en fonctionnement normal ou anormal (fautes)) dans lequel se trouve une installation à un instant donné.

Au cours de ces dernières années, de nombreuses méthodes de surveillance ont vu le jour [1, 2]. La majorité d'entre elles correspond au développement ou l'amélioration d'une des deux classes de méthodes : celles à base de modèles (model-based) et celles à base de données (data-driven). Les méthodes à base de modèles se basent sur une connaissance à priori du système (ensemble d'équations mathématique) permettant d'expliquer son comportement dynamique. A la différence des méthodes à base de données qui se basent uniquement sur les mesures collectées sur le système à différents instants.

Par ailleurs, pour améliorer la surveillance, nombreux sont les chercheurs [3, 4] suggérant la création d'un cadre commun associant les deux méthodes, permettant ainsi de bénéficier de leurs avantages respectifs. Dans la même idée, nous proposons une nouvelle méthode de surveillance basée sur les réseaux bayésiens. L'intérêt majeur de cette méthode est de permettre l'amélioration de la prise de décision en se servant des complémentarités que peuvent avoir une méthode à base de données et une méthode à base de modèle et cela sous un seul et même outil.

L'article est structuré de la manière suivante : le paragraphe 2 nous décrivons les réseaux bayésiens et les méthodes de surveillances proposées ; dans 3 une évaluation de notre méthode de surveillance sur une simulation d'un système de chauffe-eau est présentée. Les conclusions sont énoncées dans le dernier paragraphe.

2 Préliminaires

2.1 Réseaux Bayésiens

Un réseau bayésien [5] peut se définir formellement par :

- un graphe acyclique orienté G , $G=(V, E)$, où V est l'ensemble des noeuds de G , et E est l'ensemble des arcs de G ,
- un espace de probabilité (Ω, Z, P) , avec Ω un ensemble fini non-vide, Z un ensemble de sous espaces de Ω et P une mesure de probabilité sur Z avec $P(\Omega) = 1$,
- un ensemble de variables aléatoires associées aux noeuds du graphe G est défini sur (Ω, Z, P) , tel que :

$$P(V_1, V_2, \dots, V_n) = \prod_{i=1}^n P(V_i | C(V_i)) \quad (1)$$

où $C(V_i)$ est l'ensemble des noeuds parents de V_i dans le graphe G .

Chaque noeud du réseau représente une variable aléatoire pouvant être discrète (multinomial) ou continue (généralement gaussienne). Chacun de ces noeuds possède une table de probabilités conditionnelles (table de probabilités marginales pour les noeuds racines). Les arcs orientés présentent les dépendances/indépendances conditionnelles existantes entre les différents noeuds du graphe. Chaque arc orienté du réseau permet de lier seulement un noeud père à son noeud fils. Des calculs nommés inférence mettent à jour le réseau étant donné une nouvelle information (évidence).

2.2 Tests statistiques

Pour pouvoir déceler un changement survenu sur le système et le localiser, les méthodes de surveillance à base de modèle calculent la différence existante (résidus) entre les mesures prises sur le système et leurs estimations tirées du modèle. Une fois ces résidus $[r_1, \dots, r_k]^T$ générés, des tests statistiques binaires peuvent être utilisés. Parmi ces tests, on peut citer la statistique T^2 univariée, cas particulier de la statistique multivariée $T^2 = n(x - \mu)^T \Sigma^{-1} (x - \mu)$, où μ est la cible et Σ est la matrice de variance-covariance.

Ainsi pour chaque résidu généré, une analyse du scalaire obtenue en calculant T^2 est effectuée. Pour cela, on vérifie l'appartenance ou non de la mesure obtenue à un intervalle représentatif du fonctionnement normal borné par une limite de contrôle supérieure. Une fois, cette limite franchie, le système est déclaré en état de fonctionnement anormal. Pour pouvoir déceler la faute survenue, le résultat de l'évaluation représenté par un vecteur de signatures binaire (symptômes) $[u_1^t, \dots, u_k^t]^T$, est comparé avec les caractéristiques de chaque faute $F_j \in \{F_1, F_2, \dots, F_n\}$ assemblées généralement dans un tableau (lignes : les symptômes, colonnes : les fautes) appelé matrice d'incidence, traduisant la sensibilité ou la robustesse des résidus envers chacune des fautes.

Dans [6], les auteurs proposent de modéliser sous un réseau bayésien hybride les deux étapes d'évaluation de résidus et d'isolation de fautes, contenant des noeuds discrets (pères : fautes) et gaussiens (fils : résidus), permettant ainsi de décider si une faute apparaît sur le système ou non. Afin de regrouper les probabilités d'appartenance aux différentes fautes n et la probabilité que le système soit sous contrôle SC , nous proposons de rajouter au réseau un noeud père discret S_r à $n+1$ modalités, liant tous les autres noeuds F_n .

2.3 Analyse discriminante

En ce qui concerne les méthodes à base de données, après avoir aperçu un changement dans le fonctionnement normal du système, elles essaient aussitôt de l'expliquer. Dans le cas de non disponibilité d'un historique de fautes, des méthodes d'identification de fautes sont utilisées, se basant uniquement sur les mesures prises sur le système lors de son fonctionnement normal. Dans le cas contraire, lorsqu'un historique de fautes est disponible, des méthodes de classification supervisées, adaptées au diagnostic, peuvent être utilisées. Nous nous intéressons à l'analyse discriminante [7] de type quadratique employant la règle de décision $j^* = \operatorname{argmin}\{(x - \mu_j)^T \Sigma_j^{-1} (x - \mu_j) + \log(|\Sigma_j|) - 2\log(P(C_j))\}$, où $P(C_j)$ est la probabilité a priori d'une classe j , modélisable sous un réseau bayésien [friedman97], lorsque les variables sont sensés suivre une loi normale multivariée. En faisant l'hypothèse que notre connaissance sur les fautes couvre la quasi-totalité de l'espace hors contrôle. Nous proposons de discriminer entre les différentes fautes et l'état de fonctionnement sous contrôle SC . Pour ce faire, nous utilisons un réseau bayésien constitué d'un noeud discret représentant $n + 1$ modalités et un noeud gaussien multivarié regroupant les m variables du système. Ce réseau va permettre de décider à un instant donné dans quel état de fonctionnement évolue le système, parmi ses différents états (SC, F_1, F_2, \dots, F_n).

2.4 Fusion des décisions

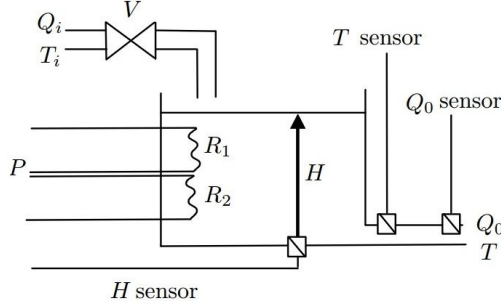
Afin d'améliorer la prise de décision, dans ces travaux nous proposons d'associer les deux méthodes de surveillance vu précédemment en un seul et même outil (un réseau bayésien). Pour cela un nouveau noeud discret $S_{r\&m}$ est rajouté. Ce noeud représente une variable à $n + 1$ modalités. L'une d'elles correspond à l'état de fonctionnement normal (SC) et les n autres modalités correspondent aux différentes fautes connues, pouvant survenir sur le système. Ce nouveau noeud $S_{r\&m}$ permet de connecter les deux noeuds (S_r et S_m) et ainsi combiner les deux méthodes décrites aux paragraphes 2.2 et 2.3.

Une fois le réseau construit (les paramètres établis), on vient rentrer les observations dans le réseau. Ces observations correspondent aux résidus obtenus et aux mesures prises sur le système. Ces évidences sont ensuite propagées vers les autres noeuds non observés du réseau. Leurs probabilités a posteriori sont

alors calculées. Nous utilisons l'algorithme d'inférence dit arbre de jonction. Après cela, le noeud $S_{r\&m}$ indique pour chacune des $n + 1$ modalités sa probabilité d'occurrence. La modalité ayant la plus grande probabilité sachant les nouvelles données, correspond à la décision prise sur l'état de fonctionnement du système.

3 Application

Afin d'illustrer notre approche, nous utilisons une simulation d'un chauffe-eau présenté dans la figure 1. Ce dernier se compose d'un réservoir équipé de deux résistances R_1 et R_2 . Ses entrées sont le débit d'eau Q_i , la température de l'eau T_i et la puissance électrique du chauffage P . Ses sorties sont le débit d'eau Q_0 et la température T réglée autour d'un point de fonctionnement.



	SC	T	H	Q_0
u_1	0	1	0	0
u_2	0	0	1	0
u_3	0	0	1	1

TABLE 1 – Matrice d'incidence

FIGURE 1 – le système étudié : système chauffe-eau

L'objectif du système est de fournir un certain débit d'eau à une température donnée. Dans cette analyse, seule les fautes capteurs sont prises en compte : capteur du niveau du liquide H , capteur de la température sortante T , capteur du débit d'eau sortant Q_0 . Le modèle mathématique détaillé du système de chauffe-eau est présenté dans [8]. Un générateur de résidus classique est utilisé (observateur de Luenberger). Le vecteur de sortie correspond à $[H, T]^T$ et le vecteur d'entrée à $[Q_i, P]$. Des résidus structurés sont générés et évalués afin de détecter les fautes aperçues sur le capteur de niveau du liquide H et le capteur de température T . Selon les équations physiques entre le taux de flux Q_0 et le niveau du liquide H , d'autres résidus peuvent être établis. La matrice d'incidence définie dans la table 1 contribuera au fondement de notre réseau bayésien.

Tout d'abord, les deux réseaux bayésiens représentant les deux méthodes de surveillance (données et modèle) sont construits. La structure des deux réseaux est complètement différente. En effet, celui représentant une méthode à base de modèle (S_r) se base uniquement sur la matrice d'incidence du chauffe-eau, contrairement à l'autre réseau (S_d) relativement simple, liant un noeud discret représentant les différents états de fonctionnement du système à un noeud continu représentant les différentes mesures prises sur le chauffe-eau. Chaque noeud des deux graphes est associé à une table de probabilités conditionnelles (TPC) quantifiant les relations entre variables. En associant ces deux réseaux, nous obtenons notre réseau final (voir figure 2).

Nous avons simulé le système selon les scénarios décrits en table 2. Pour illustrer l'intérêt de notre méthode. Ces scénarios serviront à tester la méthode proposée sous différentes hypothèses. En effet, l'intérêt de combiner les deux méthodes est de pouvoir bénéficier de bons résultats même lorsque l'une ou l'autre des méthodes n'est pas très efficiente. Ainsi, nous nous proposons de tester le réseau bayésien en prenant en compte un modèle précis ($R+$) ou un modèle moins précis ($R-$), et un jeu de données complet de taille convenable ($M+$) ou un jeu de données incomplet (données d'une faute manquante, peu de données) ($M-$). Ainsi, le scénario sera testé sous chacune des hypothèses décrites dans la table 3.

Période	1-40	41-80	81-120	121-160
Cas	SC	Faute T	Faute H	Faute Q_0

TABLE 2 – Les différents scénarios

Hypothèse	R(modèle)	M(données)
H_I	R^+	M^+
H_{II}	R^+	M^-
H_{III}	R^-	M^+
H_{IV}	R^-	M^-

TABLE 3 – les différentes hypothèses

Toutes les simulations ont été réalisées sous Matlab/ Simulink et BNT (BayesNet Toolbox). Le nombre d'erreur (mauvaise décision) pour chaque hypothèse est présenté dans la figure 3. On peut remarquer que

pour chacune des hypothèses, la méthode proposée permet généralement d'égaliser les performances et même parfois d'améliorer la surveillance. On constate donc que la proposition effectuée est relativement performante et qu'elle permet de tirer parti des deux méthodes de bases.

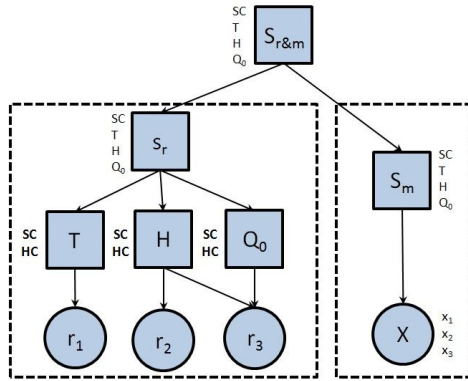


FIGURE 2 – le réseau Bayésien proposé

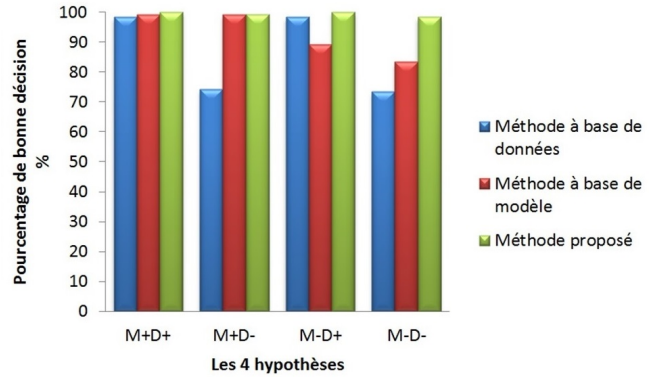


FIGURE 3 – résultats obtenus

4 Conclusions et perspectives

Nous avons présenté une structure particulière de réseau bayésien pour la surveillance des systèmes. Cette structure comprend des noeuds discrets et des noeuds gaussiens modélisant et combinant deux réseaux bayésiens dédiés à la surveillance. Un réseau bayésien classifieur pour la surveillance à base de données et un réseau bayésien classifieur représentant la matrice d'incidence pour la surveillance à base de modèle. Cette structure originale permet de renforcer la prise de décision lors de la surveillance en utilisant simultanément les données et les résidus. Cette méthode a été testée sur un système de chauffage, où une amélioration de la prise de décision a été réalisée et cela dans la plupart des cas (modèle précis, modèle dégradé, plus ou moins de données).

Références

- [1] L.H. Chiang, E.L. Russel, and R.D. Braatz. *Fault Detection and Diagnosis in Industrial Systems*. Springer, 2001.
- [2] Rolf Isermann. *fault-diagnosis system*. Springer, in robotics (vol. xviii) edition, 2006.
- [3] Venkat Venkatasubramanian, Raghunathan Rengaswamy, Surya N. Kavuri, and Kewen Yin. A review of process fault detection and diagnosis : Part iii : Process history based methods. *Computers & Chemical Engineering*, 27(3) :327 – 346, 2003.
- [4] S.X. Ding, P. Zhang, A. Naik, E.L. Ding, and B. Huang. Subspace method aided data-driven design of fault detection and isolation systems. *Journal of Process Control*, 19(9) :1496 – 1510, 2009.
- [5] F.V. Jensen. *An Introduction to Bayesian Networks*. Taylor and Francis, London, United Kingdom, London, United Kingdom, 1996.
- [6] Sylvain Verron, Philippe Weber, Didier Theilliol, Teodor Tiplica, Abdessamad Kobi, and Christophe Aubrun. Decision with bayesian network in the concurrent faults event. In *7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes (SafeProcess'09)*, 2009.
- [7] R O Duda, P E Hart, and D G Stork. *pattern classification 2nd edition*. Wiley, 2001.
- [8] P. Weber, D. Theilliol, and C. Aubrun. Component reliability in fault-diagnosis decision making based on dynamic bayesian networks. *Journal of Risk and Reliability in the Proceedings of the Institution of Mechanical Engineers, Part O*, 222(2) :161, 2008.