

# Supervision partielle par des experts d'une analyse en facteurs indépendants. Application au diagnostic d'un système ferroviaire

Zohra CHERFI<sup>1</sup>, Etienne CÔME<sup>1</sup>, Latifa OUKHELLOU<sup>2</sup>, Thierry DENOËUX<sup>3</sup>, Patrice AKNIN<sup>1</sup>

<sup>1</sup>GRETTIA - IFSTTAR - Université Paris-Est  
Le Descartes 2, 2 rue de la Butte Verte F-93166 Noisy-le-Grand Cedex, France

<sup>2</sup>LISSI (E.A. 3956), Université Paris-Est  
61 avenue du Général de Gaulle, 94010 Créteil, France

<sup>3</sup>Heudiasyc  
UTC - UMR CNRS 6599, B.P 20529, 60205 Compiègne - France  
zohra.cherfi@ifsttar.fr, oukhellou@u-pec.fr, etienne.come@ifsttar.fr  
thierry.denoeux@hds.utc.fr, patrice.aknin@ifsttar.fr

**Résumé** – Cet article illustre l'intérêt de l'utilisation d'informations partiellement fiables, obtenues grâce à différents experts, pour améliorer le diagnostic des circuits de voies ferroviaires. L'approche générale adoptée pour le diagnostic est basée sur un modèle génératif, en l'occurrence l'Analyse en Facteurs Indépendants sans bruit formulée dans un cadre d'apprentissage partiellement supervisé.

**Abstract** – This paper investigates the use of partially reliable information elicited from multiple experts to improve the diagnosis of a railway track circuit. The general statistical model used to perform the diagnosis task is based on a noiseless Independent Factor Analysis handled in a soft-supervised learning framework.

## 1 Introduction

Dans cet article le diagnostic d'un système industriel est vu comme un problème de Reconnaissance des formes (RdF) pouvant être résolu à l'aide de techniques issues de l'apprentissage statistique. La tâche de diagnostic consiste à identifier, parmi différents modes de fonctionnement connus, le mode de fonctionnement courant du système à partir de mesures. Pour être efficaces, ces méthodes nécessitent une base de données d'apprentissage exhaustive et suffisamment représentative des différents états du système. Un travail important d'étiquetage de ces données est souvent nécessaire et celui-ci peut utilement mettre à profit l'expertise humaine, même si cette labellisation peut s'avérer coûteuse et/ou fastidieuse. Dans ce cas, l'étiquetage est soumis aux imprécisions et aux incertitudes que peut rencontrer l'expert. La formalisation et la résolution du problème de diagnostic à partir de données étiquetées de manière imparfaite est précisément l'objet de cet article. L'approche proposée pour le diagnostic s'appuie sur l'Analyse en Facteurs Indépendants (IFA) [1] et sur l'utilisation de la théorie des fonctions de croyance pour représenter l'information imparfaite [4, 9, 11]. L'idée étudiée consiste à incorporer des connaissances supplémentaires sur l'appartenance aux classes de certains échantillons pour estimer les paramètres du modèle IFA, en utilisant une extension de l'algorithme EM dans un cadre partiellement supervisé [2, 6]. Les résultats sont présentés

pour le diagnostic d'un organe de l'infrastructure ferroviaire.

Cet article est organisé comme suit. La section 2 introduit quelques concepts de la théorie des fonctions de croyance. La section 3 présente le modèle utilisé pour le diagnostic. Le problème applicatif est présenté en section 4. La section 5 expose les résultats obtenus. La section 6 conclut cet article.

## 2 Théorie des fonctions de croyance

Cette section donne un aperçu des notions fondamentales de la théorie des fonctions de croyance [4, 9, 11].

### 2.1 Représentation de l'information

Soit  $\Omega = \{\omega_1, \dots, \omega_n\}$  un ensemble fini appelé *cadre de discernement*, défini comme l'ensemble des hypothèses exclusives et exhaustives concernant une question  $Q$ . Toute information partielle concernant la réponse à la question  $Q$  peut être représentée par une *fonction de masse* ou *bba* (*basic belief assignment*)  $m : 2^\Omega \rightarrow [0, 1]$ , telle que  $\sum_{A \subseteq \Omega} m(A) = 1$ , où  $m(A)$  représente le degré avec lequel les informations disponibles accréditent l'hypothèse  $A \subseteq \Omega$ , et aucune hypothèse plus spécifique. Tout ensemble  $A \subseteq \Omega$  tel que  $m(A) > 0$  est appelé *élément focal* de  $m$ . Une bba est dite :

- *normale* si  $\emptyset$  n'est pas un élément focal ;
- *dogmatique* si  $\Omega$  n'est pas un élément focal ;

- vide si  $\Omega$  est le seul élément focal (ignorance totale) ;
- simple si elle a au plus un élément focal en plus de  $\Omega$  ;
- catégorique si elle est à la fois simple et dogmatique.

Une bba simple telle que  $m(A) = 1-w, \forall A \neq \Omega$  et  $m(\Omega) = w$  peut être notée  $A^w$ . Ainsi, la bba vide est notée  $A^1$  pour tout  $A \subset \Omega$ , et la bba catégorique est notée  $A^0$  pour tout  $A \neq \Omega$ .

Une bba  $m$  peut être représentée de façon équivalente par une fonction de plausibilité  $pl(A) = \sum_{B \cap A \neq \emptyset} m(B), \forall A \subseteq \Omega$ . Cette fonction désigne la quantité de croyance pouvant être allouée à  $A$  sous réserve d'informations supplémentaires.

## 2.2 Combinaison d'informations

Soient  $m_1$  et  $m_2$  deux bbas définies sur le même cadre de discernement  $\Omega$ . Ces bbas peuvent être combinées à l'aide de règles de combinaison. Les plus courantes sont les règles conjonctive et disjonctive, respectivement définies par :

$$(m_1 \odot m_2)(A) = \sum_{B \cap C = A} m_1(B)m_2(C), \quad \forall A \subseteq \Omega, (1)$$

$$(m_1 \oplus m_2)(A) = \sum_{B \cup C = A} m_1(B)m_2(C), \quad \forall A \subseteq \Omega. (2)$$

La masse affectée à l'ensemble vide mesure le *degré de conflit* entre les deux sources d'information. Une extension de la règle conjonctive proposée par Yager considère qu'en cas de conflit, le résultat n'est pas fiable et que la masse sur  $\emptyset$  doit être reportée sur  $\Omega$ , ce qui permet d'aboutir à une bba normale [12]. Ces règles de combinaison supposent que les sources d'information sont indépendantes. Une autre règle de combinaison dite prudente  $\odot$  et introduite dans [5] est adaptée à la combinaison de bbas issues de sources non indépendantes. Cette règle peut être appliquée à toute bba non dogmatique, mais sera rappelée ici uniquement pour les bbas *séparables*, c'est-à-dire décomposables en combinaisons conjonctives de bbas simples [9, 10]. Soient deux bbas séparables  $m_1 = \odot_{A \subset \Omega} A^{w_1(A)}$  et  $m_2 = \odot_{A \subset \Omega} A^{w_2(A)}$ , où  $A^{w_1(A)}$  et  $A^{w_2(A)}$  sont des bbas simples,  $w_1(A) \in (0, 1]$  et  $w_2(A) \in (0, 1]$ , pour tout  $A \subset \Omega$ . Leur combinaison par la règle prudente est définie par :

$$(m_1 \odot m_2)(A) = \odot_{A \subset \Omega} A^{w_1(A) \wedge w_2(A)}, (3)$$

où  $\wedge$  représente l'opérateur min. Cette règle évite de comptabiliser deux fois l'information issue de sources non distinctes et elle est idempotente ( $m \odot m = m$  pour tout  $m$ ).

## 3 Analyse en Facteurs Indépendants

### 3.1 Formalisme

L'IFA est un modèle génératif qui considère les variables observées comme une combinaison linéaire de variables latentes indépendantes, et qui permet d'estimer à la fois les variables latentes (appelée également sources) et le processus de mélange à partir des seules variables observées [7, 1]. Dans sa version sans bruit, le modèle IFA est formulé par :  $\mathbf{y} = H \mathbf{z}$ , où  $H$  est une matrice carrée non-singulière de taille  $S$  et  $\mathbf{y}$  est le vecteur

des  $S$  variables observées dont les éléments sont des mélanges linéaires des  $S$  variables latentes  $\mathbf{z}$ . La densité marginale de chaque source est modélisée par un modèle de mélange à  $K_j$  composantes gaussiennes et prend la forme :

$$f^{\mathbf{z}}(\mathbf{z}) = \prod_{j=1}^S \sum_{k=1}^{K_j} \pi_k^j \varphi(z_j; \mu_k^j, \nu_k^j). (4)$$

Le vecteur de paramètres  $\Psi$  contient la matrice de mixage  $H$  et les paramètres des densités des sources que sont les proportions des composantes  $\pi^j$  ainsi que les vecteurs  $\mu^j$  et  $\nu^j$  contenant les moyennes et les variances de chaque composante. Dans sa version classique, l'apprentissage du modèle IFA est réalisé dans un contexte non supervisé et repose sur la seule observation des mélanges pour estimer  $\Psi$ . Ceci peut être fait en maximisant la vraisemblance du modèle par une stratégie d'optimisation alternée [1].

### 3.2 IFA partiellement supervisée

L'IFA est présentée ici dans un contexte partiellement supervisé, une connaissance partielle sur l'appartenance aux classes étant supposée disponible pour certains individus sous la forme de bbas. Chaque source est modélisée par un mélange de plusieurs composantes gaussiennes représentant ses différents états discrets. L'idée est d'intégrer au modèle l'information disponible sur la composante d'origine de quelques individus, celle-ci pouvant être imprécise et incertaine. Plus formellement, nous supposons disposer d'un jeu de données de la forme :

$$\mathbf{M} = \{(\mathbf{y}_1, m_1^1, \dots, m_1^S), \dots, (\mathbf{y}_N, m_N^1, \dots, m_N^S)\}, (5)$$

où  $m_i^1, \dots, m_i^S$  est l'ensemble des bbas spécifiant la connaissance disponible sur la composante d'origine de l'individu  $i$  pour les  $S$  sources. Chaque bba  $m_i^j$  est définie sur le cadre de discernement  $\mathcal{U}_j = \{c_1, \dots, c_{K_j}\}$  composé des  $K_j$  composantes possibles pour la source  $j$ . L'ensemble des données complétées est noté  $\mathbf{x}_i = (\mathbf{y}_i, u_i^1, \dots, u_i^S)$  où  $\mathbf{y}_i \in \mathbb{R}^S$  sont les variables observées et où les  $u_i^j \in \mathcal{U}^j, \forall j \in \{1, \dots, S\}$  et  $\forall i \in \{1, \dots, N\}$  sont les variables définissant l'appartenance aux classes et pour lesquelles seule une connaissance imparfaite est disponible. Ce modèle repose sur deux hypothèses d'indépendance. L'hypothèse d'indépendance stochastique induite par le processus aléatoire de génération des données qui s'exprime de la manière suivante :

$$f(\mathbf{X}; \Psi) = \prod_{i=1}^N f(\mathbf{x}_i; \Psi), (6)$$

où  $\Psi$  est le vecteur des paramètres,  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$  le vecteur d'échantillons complet et  $f(\mathbf{x}_i; \Psi)$  la densité d'une observation complète en fonction des paramètres. La connaissance imparfaite sur l'appartenance aux classes induit, quant à elle, l'hypothèse d'indépendance cognitive suivante [9] :

$$pl(\mathbf{X}) = \prod_{i=1}^N pl_i(\mathbf{x}_i) = \prod_{i=1}^N \prod_{j=1}^S pl_i^j(u_i^j), (7)$$

où  $pl(\mathbf{X})$  est la plausibilité que  $\mathbf{X}$  soit le vecteur d'échantillon complet,  $pl_i(\mathbf{x}_i)$  la plausibilité que les données complètes pour l'échantillon  $i$  soient égales à  $\mathbf{x}_i$  et  $pl_i^j(u_i^j)$  la plausibilité que la source  $j$  de l'échantillon  $i$  ait été générée par la composante  $u_i^j$ . Sous les hypothèses (6) et (7) et selon [2] et [6], la log-vraisemblance peut être écrite sous la forme suivante :

$$\mathcal{L}(\Psi; \mathbf{M}) = -N \log(|\det(H)|) + \sum_{i=1}^N \sum_{j=1}^S \log \left( \sum_{k=1}^{K^j} pl_{ik}^j \pi_k^j \varphi((H^{-1} \mathbf{y}_i)^j; \mu_k^j, \nu_k^j) \right) \quad (8)$$

où  $pl_{ik}^j = pl_i^j(c_k)$  est la plausibilité que l'individu  $i$  appartienne à la composante  $k$  de la variable latente  $j$  (obtenue à partir de  $m_i^j$ ). Le critère (8) doit être maximisé par rapport à  $\Psi$  pour estimer les paramètres. Une extension de l'algorithme EM appelée  $E^2M$  pour *Evidential EM* peut être utilisée pour effectuer cette tâche [6].

## 4 Application

### 4.1 Description de l'application

Le modèle présenté précédemment a été mis œuvre pour le diagnostic des circuits de voie (CdV), éléments essentiels de la signalisation ferroviaire [8]. La principale fonction du CdV est de détecter automatiquement la présence d'un train sur une portion de voie donnée. Pour ce faire, la voie est divisée en sections et chaque section est équipée d'un CdV composé de (Figure 1) : un émetteur branché à la sortie de la portion de voie et qui émet un courant alternatif modulé en fréquence, les deux files de rail qui servent de ligne de transmission, un récepteur branché à l'entrée de la portion de voie et des condensateurs de compensation connectés entre les deux files de rail à intervalles réguliers pour compenser le caractère inductif de la voie. Un train est détecté lorsque son premier essieu court-circuite la voie et provoque une perte du signal au niveau du récepteur.

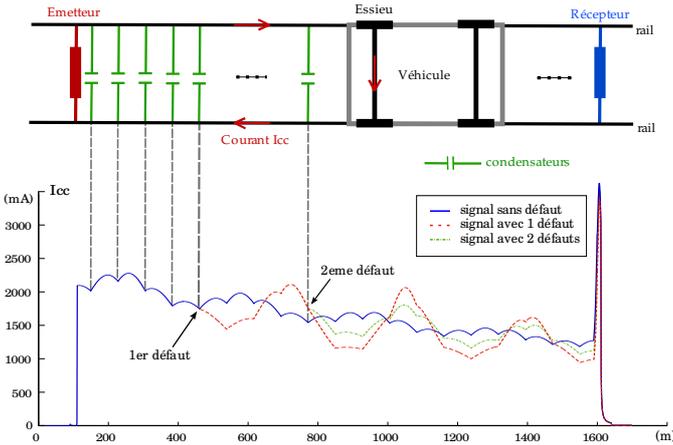


FIGURE 1 – Représentation du CdV et du signal Icc.

L'ensemble de ce système peut être sujet à des dysfonctionnements qu'il convient de diagnostiquer au plus tôt afin de con-

server une disponibilité maximale. Dans des cas extrêmes, une forte atténuation du signal transmis peut provoquer d'importants problèmes de signalisation et entraîner l'arrêt de l'exploitation. Nous nous intéressons dans cet article aux défauts liés aux condensateurs de compensation. Le diagnostic de ces derniers se fait sur la base de signaux d'inspection relevés par un véhicule de mesure (courant de court circuit ou signal Icc) dont l'allure est étroitement liée aux caractéristiques des condensateurs (Figure 1).

### 4.2 Méthodologie du diagnostic

Dans ce qui suit, le circuit de voie est vu comme un système complexe réparti composé de  $S$  sous-systèmes, où chaque sous-système correspond à un condensateur. Un défaut sur un sous-système peut être représenté par une valeur continue liée à la capacité du condensateur ou par une valeur discrète si un nombre fini de modes de dégradation est pris en compte. Des caractéristiques sont alors extraites du signal de mesure, lesquelles servent ensuite à construire le modèle génératif détaillé dans la figure 2. Dans ce modèle, les variables  $y_i^j$  correspondent aux

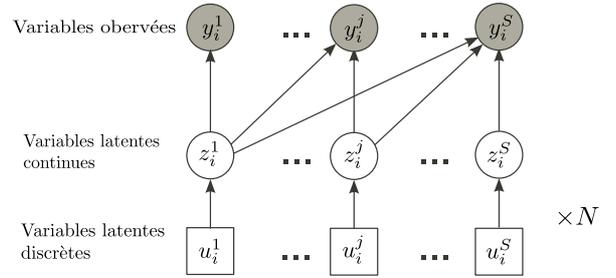


FIGURE 2 – Modèle graphique pour le diagnostic des CdV.

variables observées obtenues en approximant chaque arche du signal Icc par un polynôme quadratique [8]. Les  $z_i^j$  représentent les variables latentes continues liées aux capacités des condensateurs et les  $u_i^j$  représentent les variables latentes discrètes définissant l'état de fonctionnement des condensateurs (sans défaut, défaut intermédiaire ou défaut grave). En supposant une relation linéaire entre variables latentes et observées et en modélisant chaque variable latente par un mélange de trois gaussiennes, le modèle génératif précédent est celui de l'IFA [3].

## 5 Résultats

L'approche proposée a été évaluée sur une base de données réelles de 422 signaux d'inspection fournis par la SNCF. Une campagne de labellisation a été organisée, avec pour objectif, l'étiquetage de ces signaux par quatre experts. Trois classes étaient considérées lors de l'étiquetage, correspondant aux trois modes de fonctionnement des condensateurs. Les experts pouvaient émettre un avis incertain et imprécis lorsqu'une décision claire était difficile. En définissant le cadre de discernement comme l'ensemble des trois modes de fonctionnement possi-

bles  $\Omega = \{\omega_0, \omega_1, \omega_2\}$  (sans défaut, défaut intermédiaire et défaut grave), les avis de chaque expert ont été modélisés par des bbs sur  $\Omega$ . Ainsi, un avis incertain définissait un degré de confiance  $m(\{\omega_k\}) < 1$  où  $k \in \{0, 1, 2\}$  est le mode de fonctionnement (le reste allant sur l'ignorance  $m(\Omega) = 1 - m(\{\omega_k\})$ ) et un avis imprécis définissait un degré de confiance  $m(\{\omega_k, \omega_l\}) \leq 1$  où  $k$  et  $l$  sont les modes de fonctionnement concernés ( $k, l \in \{0, 1, 2\}$  et  $k \neq l$ ). Ces avis ont ensuite été fusionnés à l'aide des opérateurs rappelés dans la section 2.2. Le résultat de chacune des combinaisons a été intégré lors de l'apprentissage du modèle IFA par l'algorithme E<sup>2</sup>M [6]. Une validation croisée (10 blocs) a été utilisée pour évaluer les performances de chaque approche. Celles-ci sont résumées dans le tableau 1. On peut remarquer que de bonnes performances en classification sont obtenues malgré quelques confusions entre classes voisines. De plus, parmi les experts, certains ont plus de facilité à labelliser les défauts graves (experts 1 et 4), tandis que d'autres reconnaissent mieux les défauts intermédiaires (expert 2 et 3). Les apprentissages réalisés avec les labels issus des combinaisons tirent parti de ces tendances et permettent d'obtenir de bonnes performances à la fois sur les défauts intermédiaires et graves. Les meilleurs résultats sont obtenus par la règle prudente, ce qui confirme que les opinions émises par les experts ne peuvent être considérées comme distinctes [5].

TABLE 1 – Matrices de confusion entre les vraies classes  $\omega_0$ ,  $\omega_1$  et  $\omega_2$  et leur prédictions  $d_0$ ,  $d_1$  et  $d_2$  (obtenues pour chaque apprentissage), calculées par validation croisée.

	$\omega_0$	$\omega_1$	$\omega_2$
$d_0$	98.8	33.1	2.1
$d_1$	0.9	51.1	6.9
$d_2$	0.2	15.8	90.9

(Expert 1)

	$\omega_0$	$\omega_1$	$\omega_2$
$d_0$	98.9	34.7	3.0
$d_1$	0.8	58.8	12.2
$d_2$	0.3	6.5	84.7

(Expert 2)

	$\omega_0$	$\omega_1$	$\omega_2$
$d_0$	98.7	22.1	2.1
$d_1$	1.1	63.6	13.8
$d_2$	0.2	14.3	84.1

(Expert 3)

	$\omega_0$	$\omega_1$	$\omega_2$
$d_0$	98.8	34.6	3.3
$d_1$	1.0	49.6	5.8
$d_2$	0.2	15.8	90.9

(Expert 4)

	$\omega_0$	$\omega_1$	$\omega_2$
$d_0$	98.9	23.1	2.9
$d_1$	0.8	62.8	8.0
$d_2$	0.1	14.1	89.2

(Combinaison disjonctive)

	$\omega_0$	$\omega_1$	$\omega_2$
$d_0$	98.9	20.2	2.9
$d_1$	1.0	64.2	6.5
$d_2$	0.1	15.6	90.6

(Combinaison conjonctive)

	$\omega_0$	$\omega_1$	$\omega_2$
$d_0$	98.9	20.4	2.6
$d_1$	1.0	65.3	4.9
$d_2$	0.1	14.2	92.4

(Combinaison prudente)

## 6 Conclusion

Nous nous sommes intéressés dans cet article au diagnostic d'un système complexe dédié à la signalisation ferroviaire. L'apprentissage d'un modèle génératif de type analyse en facteurs indépendants dans un contexte partiellement supervisé est au centre de l'approche adoptée [2][6]. Les expérimentations

ont été réalisées sur des données réelles étiquetées par quatre experts. Ces derniers ont fourni des informations partiellement fiables sur l'état des sous-systèmes, qui ont été modélisées par des fonctions de croyance et fusionnées par la suite par plusieurs règles de combinaison. Nous avons constaté à travers nos expérimentations que la combinaison des avis permettait d'obtenir de meilleures performances pour le diagnostic que la prise en compte d'un avis unique. Plusieurs voies d'approfondissement peuvent être envisagées. L'approche repose sur l'élicitation de connaissances expertes dans le cadre de la théorie des fonctions de croyance, problème important qui a reçu peu d'attention jusqu'à présent. Par ailleurs, des schémas de combinaison plus sophistiqués peuvent être envisagés. Par exemple, des taux d'affaiblissement pourraient être appris à partir des données pour tenir compte des compétences de chaque expert.

Ce travail a été effectué dans le cadre du projet DiagHist financé par l'ANR.

## Références

- [1] H. Attias. Independent factor analysis, Neural Computation. MIT Press, Cambridge, MA. 11(4):803–851, 1999.
- [2] E. Côme, L. Oukhellou, T. Denoeux, P. Aknin. Learning from partially supervised data using mixture models and belief functions. Pattern Recognit. 42(3) :334–348, 2009.
- [3] E. Côme, L. Oukhellou, T. Denoeux and P. Aknin. Fault diagnosis of a railway device using semi-supervised independent factor analysis with mixing constraints. Pattern Analysis and Applications, accepted for publication, 2011. DOI : 10.1007/s10044-011-0212-3.
- [4] A. Dempster. Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Stat. 38:325–339, 1967.
- [5] T. Denoeux. Conjunctive and Disjunctive Combination of Belief Functions Induced by Non Distinct Bodies of Evidence. Artificial Intelligence. 172 :234–264, 2008.
- [6] T. Denoeux. Maximum likelihood from evidential data : an extension of the EM algorithm. C. Borgelt et al. (Eds) Combining soft computing and statistical methods in data analysis, AISC 77. Springer, 2010.
- [7] E. Moulines, J. Cardoso, E. Cassiat. Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models. ICASSP, Munich, 1997.
- [8] L. Oukhellou, A. Debiolles, T. Denoeux, P. Aknin. Fault diagnosis in railway track circuits using Dempster-Shafer classifier fusion. Eng. Appl. of Artif. Intell. 23(1) :117–128, 2010.
- [9] G. Shafer. A mathematical theory of evidence. Princeton Univ. Press, 1976.
- [10] P. Smets. The canonical decomposition of a weighted belief. Int. Joint Conf. on Artif. Intell, San Mateo, Ca, 1995.
- [11] P. Smets, R. Kennes. The Transferable Belief Model. Artif. Intell. 66 :191–234, 1994.
- [12] RR. Yager. On the Dempster-Shafer framework and new combination rules. Inf. Sci. 41(2):93–137, 1987.