

Gestion de l'incertain en diagnostic: l'apport de la théorie des fonctions de croyance

Thierry Denœux¹

¹Université de Technologie de Compiègne
HEUDIASYC (UMR CNRS 6599)

Ecole des JDMACS
19 mars 2009



Diagnostic et Incertitude

- Diagnostic un système : déterminer son **mode de fonctionnement** à partir de mesures réalisées sur ce système et d'informations a priori (modèles, données historiques).
- Les mesures ne permettent généralement pas de déterminer avec certitude le mode de fonctionnement → elles apportent une **information partielle** sur l'état du système.
- Il est souvent nécessaire de combiner différentes informations (mesures, connaissances a priori, prédictions fournies par différents modèles) → **fusion d'informations**.



Fusion d'informations

- **Fusion d'informations** = combinaison d'informations partielles (imprécises et/ou incertaines) issues de plusieurs sources en vue de déterminer la valeurs de certaines variables.
- Nécessité de définir des formalismes suffisamment **généraux** pour représenter et combiner tous types d'informations susceptibles d'être rencontrés.
- Les problématiques du diagnostic des systèmes, de la fusion de l'information et de la **gestion de l'incertain** (au sens large) sont fortement liées.

Typologie des imperfections de l'information

- Soit X une variable à valeurs dans Ω (domaine, cadre de discernement).
- Un élément d'information sur X peut être représenté sous forme d'un couple (**valeur, confiance**) :
 - l'élément "valeur" correspond à un sous-ensemble de Ω ;
 - l'élément "confiance" exprime un jugement sur la validité de l'information.
- L'élément d'information est
 - **imprécis** si "valeur" n'est pas un singleton de Ω ;
 - **incertain** si l'élément "confiance" exprime un doute sur la véracité de l'information.

Exemple

- Soit X = la température de la pièce
- "Il fait entre 15 et 25 degrés" = $([15, 25], \text{certain}) \rightarrow$ élément d'information **certain et imprécis**.
- "Il est plausible qu'il fasse 20 degrés" = $(20, \text{plausible}) \rightarrow$ élément d'information **incertain et précis**.
- "Il fait probablement entre 15 et 25 degrés" = $([15, 25], \text{probable}) \rightarrow$ élément d'information **incertain et imprécis**.

Formalismes classiques

1 Théorie de ensembles

- Calcul par intervalle, estimation à erreur bornée
- Modélise bien l'imprécision mais n'exprime aucune notion d'incertitude
- Peu robuste et excessivement conservative

2 Théorie des probabilités

- Modélise l'incertitude aléatoire (variabilité dans une population ou au cours des répétitions d'une expérience aléatoire).
- N'exprime pas de notion d'imprécision.

Théorie des fonctions de croyance

- Introduite par Dempster (1968) et Shafer (1976), puis développée par Smets (**Modèle des Croyances Transférables**) à partir de 1978.
- Généralise à la fois la théorie des ensembles et la théorie des probabilités :
 - Une fonction de croyance peut être vue comme un **ensemble généralisé** ou comme une **mesure non additive**
 - Extension de notions probabilistes (conditionnement, marginalisation) et ensemblistes (intersection, union, inclusion, etc.)
- La combinaison des informations (éléments d'évidence) occupe une place centrale dans la théorie : nombreuses applications en fusion d'informations.



Plan de la présentation

- 1 Théorie des fonctions de croyance
 - Concepts fondamentaux
 - Fusion d'informations
 - Décomposition canonique et règle prudente
 - Décision
- 2 Application au diagnostic par RdF
 - Problème et motivations
 - Construction d'un classifieur crédibiliste
 - Décision
 - Exemple de fusion de classifieurs
- 3 Etudes de cas
 - Surveillance acoustique de structure
 - Diagnostic de machine asynchrone
 - Diagnostic de circuits de voie ferroviaires

Fonction de masse de croyance

- Soit Ω un ensemble fini, domaine d'un variable X .
- Fonction de masse de croyance : $m : 2^\Omega \rightarrow [0, 1]$ telle que

$$\sum_{A \subseteq \Omega} m(A) = 1.$$

- Tout sous-ensemble A de Ω tel que $m(A) > 0$ est appelé **ensemble focal**.
- Interprétation : m représente
 - Un élément d'évidence relatif à la valeur de X .
 - Un état de connaissance (croyance) induit par cet élément d'évidence.



Cas particuliers

- m peut être vue comme :
 - Une famille d'ensembles pondérés $\{(A_i, m(A_i)), i = 1, \dots, r\}$.
 - Une distribution de probabilité généralisée (masses distribuées sur 2^Ω et non sur Ω).
- Cas particuliers :
 - $r = 1$: fonction de masse **catégorique** (\sim ensemble). On note m_A la fonction de masse catégorique d'ensemble focal A .
 - $|A_i| = 1, i = 1, \dots, r$: fonction de masse **bayésienne** (\sim distribution de probabilité).

Fonctions de croyance et de plausibilité

- Fonction de croyance :

$$bel(A) = \sum_{\substack{B \subseteq A \\ B \not\subseteq \bar{A}}} m(B) = \sum_{\emptyset \neq B \subseteq A} m(B), \quad \forall A \subseteq \Omega$$

(degré de croyance (support) en l'hypothèse " $X \in A$ ")

- Fonction de plausibilité :

$$pl(A) = \sum_{B \cap A \neq \emptyset} m(B), \quad \forall A \subseteq \Omega$$

(degré maximal de croyance susceptible d'être alloué à A après intégration de nouvelles informations)

- $bel \leq pl$.



Relations entre m , bel et pl

- Relations :

$$bel(A) = pl(\Omega) - pl(\bar{A}), \quad \forall A \subseteq \Omega$$

$$m(A) = \begin{cases} \sum_{\emptyset \neq B \subseteq A} (-1)^{|A|-|B|} bel(B), & A \neq \emptyset \\ 1 - bel(\Omega) & A = \emptyset \end{cases}$$

- m , bel et pl sont donc **trois représentations équivalentes** d'un même élément d'information.

Relations d'inclusion

- Soient m_1 et m_2 deux fonctions de masse sur Ω .
- En quel sens peut-on dire que m_1 est **plus informative** que m_2 ?
- Cas particulier :
 - Soient m_A et m_B deux fonctions de masse catégoriques.
 - m_A est plus informative (au sens large) que m_B ssi $A \subseteq B$.
- Généralisation à des fonctions de masse quelconques ?

Inclusion faible

- m_1 plus informative (riche) que m_2 **au sens des plausibilités** (noté $m_1 \sqsubseteq_{pl} m_2$) si

$$pl_1(A) \leq pl_2(A), \quad \forall A \subseteq \Omega.$$

- Propriétés :
 - Généralise l'inclusion : $m_A \sqsubseteq_{pl} m_B \Leftrightarrow A \subseteq B$.
 - Plus grand élément : m_Ω t.q. $m_\Omega(\Omega) = 1$ (fonction de masse vide).

Inclusion forte

- m_1 est une **spécialisation** de m_2 (noté $m_1 \sqsubseteq_s m_2$) si m_1 s'obtient à partir de m_2 en répartissant chaque masse $m_2(B)$ entre des sous-ensembles de B :

$$m_1(A) = \sum_{B \subseteq \Omega} S(A, B) m_2(B), \quad \forall A \subseteq \Omega,$$

avec $S(A, B) =$ proportion de $m_2(B)$ transférée à $A \subseteq B$.

- S : matrice de spécialisation.
- Propriétés :
 - Généralise l'inclusion
 - Plus grand élément : m_Ω .
 - $m_1 \sqsubseteq_s m_2 \Rightarrow m_1 \sqsubseteq_{pl} m_2$.

Principe d'engagement minimal

Définition (Principe d'engagement minimal)

*Lorsque plusieurs fonctions de croyance sont compatibles avec un ensemble de contraintes, la **moins informative** (au sens d'un certain ordre) doit être choisie.*

- Exemple : On sait seulement que $p_l(A) = 0$.
- La fonction de masse la moins informative (au sens de \sqsubseteq_{pl} et \sqsubseteq_s) vérifiant cette contrainte et $m_0(\bar{A}) = 1$. Elle vérifie $p_l(A) = 0$ et $p_l(B) = 1$ pour tout $B \not\subseteq A$.

Plan de la présentation

- 1 **Théorie des fonctions de croyance**
 - Concepts fondamentaux
 - **Fusion d'informations**
 - Décomposition canonique et règle prudente
 - Décision
- 2 **Application au diagnostic par RdF**
 - Problème et motivations
 - Construction d'un classifieur crédibiliste
 - Décision
 - Exemple de fusion de classifieurs
- 3 **Etudes de cas**
 - Surveillance acoustique de structure
 - Diagnostic de machine asynchrone
 - Diagnostic de circuits de voie ferroviaires

Opérateurs conjonctifs et disjonctifs

- Soient m_1 et m_2 deux fonctions de masse sur Ω issues de 2 sources.
- Un opérateur de combinaison fusionne m_1 et m_2 en une nouvelle fonction de masse $m_{1*2} = m_1 * m_2$ qui traduit notre nouvel état de connaissance après avoir pris en considération les deux 2 sources.
- L'opérateur $*$ est
 - **conjonctif** si m_{1*2} est **plus informative** que m_1 et m_2 .
 - **disjonctif** si m_{1*2} est **moins informative** que m_1 et m_2 .
- Le choix d'un opérateur conjonctif suppose que les sources soient toutes deux fiables.
- Combinaison disjonctive : stratégie prudente (l'une au moins des deux sources est fiable).

Conditionnement

- Soient deux fonctions de masse m et m_B avec $m_B(B) = 1$ pour $B \subset \Omega$.
- Etat de connaissance après fusion conjonctive de m et m_B ? Soit $m[B]$ le résultat de la combinaison.
- Contraintes :
 - $pl[B](\bar{B}) = 0$.
 - $m[B] \sqsubseteq_s m$.
- Principe d'engagement minimal : solution la moins informative au sens de \sqsubseteq_{pl} :

$$m[B](A) = \sum_{\{C \mid C \cap B = A\}} m(C).$$



Propriétés

- Généralisation de l'**intersection** : $m_A[B] = m_{A \cap B}$.
- Généralisation du **conditionnement probabiliste** :
 - Si $m(\emptyset) > 0$, soit m^* la fonction de masse normalisée

$$m^*(A) = \frac{m(A)}{1 - m(\emptyset)}.$$

- Expression du conditionnement normalisé :

$$pl[B]^*(A) = \frac{pl(A \cap B)}{pl(B)}$$

- Si m est bayésienne, $pl = P$: on retrouve le conditionnement probabiliste.



Plausibilité, communalité

- Interprétation de $pl(A)$:
 - $pl(A) = belA = \max_B bel[B](A)$
 - **degré maximal de croyance** susceptible d'être alloué à A après conditionnement.
- Fonction de **communalité** : soit $q : 2^\Omega \rightarrow [0, 1]$ définie par $q(A) = mA$:
 - Masse attachée au plus grand ensemble possible (**degré d'ignorance**) après conditionnement par A .
 - Autre expression

$$q(A) = \sum_{B \supseteq A} m(B).$$

- Nouvelle relation d'inclusion : $m_1 \sqsubseteq_q m_2$ si $q_1 \leq q_2$.

Règle de Dempster : hypothèses

- Soit $*$ un opérateur de fusion conjonctif, tel que

$$m_1 * m_2 = S_1 \cdot m_2, \quad m_1 * m_2 = S_2 \cdot m_1,$$

S_1 et S_2 étant des opérateurs de spécialisation.

- Hypothèses :
 - **Indépendance** : S_1 ne dépend pas de m_2 , S_2 ne dépend pas de m_1 .
 - Généralisation du conditionnement : $m_1 * m_B = m[B]$.
 - Commutativité : $m_1 * m_2 = m_2 * m_1$.
- Solution : **règle de Dempster**.

Règle de Dempster

Définition (Règle de combinaison de Dempster)

$$\begin{aligned}\forall A \subseteq \Omega, \quad (m_1 \oplus m_2)(A) &= \sum_{B \subseteq \Omega} m_1[B](A) m_2(B), \\ &= \sum_{B \cap C = A} m_1(B) m_2(C).\end{aligned}$$

Propriétés :

- commutativité, associativité.
- élément neutre : m_Ω
- $(q_1 \oplus q_2) = q_1 \cdot q_2$
- $(m_1 \oplus m_2)(\emptyset) \geq 0$: degré de conflit.

Règle disjonctive

Définition (Règle disjonctive)

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

Propriétés :

- commutativité, associativité.
- élément neutre : m_\emptyset
- Soit $b = bel + m(\emptyset)$ (fonction d'implicabilité). On a :

$$(b_1 \oplus b_2) = b_1 \cdot b_2$$

Choix d'une règle

- Les règles \cap et \cup supposent que les sources d'information sont **distinctes** (la même information n'est pas comptée deux fois). Si cette hypothèse n'est pas vérifiée : **règle prudente**.
- La règle conjonctive suppose que les deux sources sont fiables : produit une information de plus en plus précise, mais génère du conflit.
- La règle disjonctive suppose que l'une au moins des sources est fiable. Ne génère pas de conflit mais produit une information de moins en moins précise.



Gestion du conflit

- Un conflit important révèle souvent (toujours ?) un **problème de modélisation** :
 - Crédit excessif accordé à certaines sources ;
 - Hypothèse d'indépendance des sources non vérifiée ;
 - Cadre de discernement non exhaustif, etc...
- Approches :
 - Revenir sur la modélisation de l'information ;
 - Utiliser des règles **robustes** (résultat cohérent en cas de conflit).

Règle de Dubois et Prade

- Définition :

$$(m_1 *_{DP} m_2)(A) = \sum_{B \cap C = A} m_1(B)m_2(C) + \sum_{\{B \cap C = \emptyset, B \cup C = A\}} m_1(B)m_2(C), \quad \forall A \neq \emptyset$$

$$(m_1 *_{DP} m_2)(\emptyset) = 0.$$

- Propriétés :

- $m_1 *_{DP} m_2 = m_1 \odot m_2$ si $(m_1 \odot m_2)(\emptyset) = 0$,
- $m_1 *_{DP} m_2 = m_1 \oplus m_2$ si $(m_1 \odot m_2)(\emptyset) = 1$.
- Commutative mais **non associative**.

Prise en compte de la fiabilité des sources

Exposé du problème

- Une source d'information fournit :
 - une valeur
 - un ensemble de valeurs
 - une distribution de probabilité, etc..
- L'information fournie est :
 - non totalement fiable ou
 - non totalement pertinente.
- Exemples :
 - Défaillance possible d'un capteur ;
 - Contexte de la mesure plus ou moins favorable ;
 - Information se rapportant à une situation ou un objet ayant seulement un degré de similarité avec la situation ou l'objet considéré (raisonnement à base de cas).



Prise en compte de la fiabilité des sources

Modélisation

- Deux cadres de discernement :
 - fiabilité ou pertinence de la source : $\mathcal{F} = \{F, NF\}$.
 - Domaine Ω de la variable X d'intérêt.
- Une source S fournit une fonction de masse m_S^Ω .
- Éléments d'information disponibles :
 - $pl^{\mathcal{F}}(\{NF\}) = \alpha, pl^{\mathcal{F}}(\{F\}) = 1$.
 - $m^\Omega[F] = m_S^\Omega$.
 - $m^\Omega[NF] = m_\Omega^\Omega$.
- Information induite sur Ω :

$$\alpha m^\Omega = \left(m^{\mathcal{F} \uparrow \Omega \times \mathcal{F}} \oplus m^\Omega[F] \uparrow \Omega \times \mathcal{F} \right) \downarrow \Omega$$

Prise en compte de la fiabilité des sources

Affaiblissement

- Opération d'affaiblissement :

$${}^{\alpha}m^{\Omega}(A) = (1 - \alpha)m_S^{\Omega}(A), \quad \forall A \subset \Omega,$$

$${}^{\alpha}m^{\Omega}(\Omega) = (1 - \alpha)m_S^{\Omega}(\Omega) + \alpha.$$

- α est appelé **coefficient d'affaiblissement** : ${}^0m^{\Omega} = m_S^{\Omega}$ et ${}^1m^{\Omega} = m_{\Omega}^{\Omega}$.
- Expression équivalente :

$${}^{\alpha}m^{\Omega} = m_S^{\Omega} \circledast \emptyset^{\alpha}.$$

- ${}^{\alpha}m^{\Omega} \supseteq_s m_S^{\Omega}$.

Plan de la présentation

- 1 Théorie des fonctions de croyance
 - Concepts fondamentaux
 - Fusion d'informations
 - **Décomposition canonique et règle prudente**
 - Décision
- 2 Application au diagnostic par RdF
 - Problème et motivations
 - Construction d'un classifieur crédibiliste
 - Décision
 - Exemple de fusion de classifieurs
- 3 Etudes de cas
 - Surveillance acoustique de structure
 - Diagnostic de machine asynchrone
 - Diagnostic de circuits de voie ferroviaires

Fonction de masse simple

- Définition : m est une **fonction de masse simple** si elle est de la forme

$$\begin{aligned}m(A) &= 1 - w_A \\ m(\Omega) &= w_A,\end{aligned}$$

avec $A \subset \Omega$ et $w_A \in [0, 1]$.

- Notation : A^{w_A} .
- Propriété : $A^{w_1} \circledast A^{w_2} = A^{w_1 w_2}$.
- Une fonction de masse est dite **séparable** si elle peut s'obtenir comme combinaison de fonctions de masse simples :

$$m = \circledast_{A \subseteq \Omega} A^{w(A)}$$



Retrait d'information

- Soit $m_{12} = m_1 \otimes m_2$. On a $q_{12} = q_1 \cdot q_2$.
- Supposons qu'on s'aperçoive que m_2 a été ajoutée par erreur.
- Si m_2 est **non dogmatique** ($q(A) > 0, \forall A$), m_1 peut être retrouvée à partir de m_{12} par

$$q_1 = q_{12}/q_2.$$

- On note $m_1 = m_{12} \oslash m_2$.
- Remarque : $m_1 \oslash m_2$ n'est pas toujours une fonction de croyance !

Décomposition canonique

Théorème (Smets, 1995)

Toute fonction de masse non dogmatique ($m(\Omega) > 0$) peut être décomposée de manière unique sous la forme :

$$m = \left(\bigoplus_{A \subset \Omega} A^{w_C(A)} \right) \otimes \left(\bigoplus_{A \subset \Omega} A^{w_D(A)} \right)$$

avec $w_C(A) \in]0, 1]$, $w_D(A) \in]0, 1]$ et $\max(w_C(A), w_D(A)) = 1$ pour tout $A \subset \Omega$.

- On peut poser $w = w_C/w_D$.
- La fonction $w : 2^\Omega \setminus \Omega \rightarrow \mathbb{R}_+^*$ est une nouvelle **représentation équivalente** d'une fonction de masse non dogmatique (avec bel, pl, q, b).

Propriétés de w

- La fonction w est disponible directement lorsque m est construite par **accumulation de fonctions de masses simples** (cas fréquent).
- Calcul de w à partir de q :

$$\ln w(A) = - \sum_{B \supseteq A} (-1)^{|B|-|A|} \ln q(B), \quad \forall A \subset \Omega.$$

- Règle de Dempster :

$$w_1 \circledast w_2 = w_1 \cdot w_2.$$

Relation d'inclusion \sqsubseteq_w

- Soient m_1 et m_2 deux fonctions de masse non dogmatiques. On pose

$$m_1 \sqsubseteq_w m_2 \text{ ssi } w_1 \leq w_2.$$

- Interprétation $m_1 = m_2 \odot m$ avec m séparable.
- Propriétés :

- $m_1 \sqsubseteq_w m_2 \Rightarrow m_1 \sqsubseteq_s m_2 \Rightarrow \begin{cases} m_1 \sqsubseteq_{pl} m_2 \\ m_1 \sqsubseteq_q m_2, \end{cases}$
- m_Ω est le seul élément maximal de \sqsubseteq_w :

$$m_\Omega \sqsubseteq_w m \Rightarrow m = m_\Omega.$$

Règle prudente

Principe

- Un agent reçoit deux fonctions de masse m_1 et m_2 issues de sources fiables.
- Soit m_{12} la fct de masse représentant son état de connaissance après réception de m_1 et m_2 .
- m_{12} doit être plus informative que m_1 et m_2 .
Formellement : $m_{12} \in \mathcal{S}_x(m_1) \cap \mathcal{S}_x(m_2)$, pour un $x \in \{w, s, pl, q\}$, avec $\mathcal{S}_x(m) =$ ens. des fcts de masse plus informatives que m au sens de \sqsubseteq_x .
- Principe d'engagement minimal : on choisit dans $\mathcal{S}_x(m_1) \cap \mathcal{S}_x(m_2)$ la fonction de masse la moins informative au sens de \sqsubseteq_x (si elle existe).

Règle prudente

Définition

Théorème

Soient m_1 and m_2 deux fcts de masse non dogmatiques. L'élément le moins informatif de $\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2)$ au sens de \sqsubseteq_w existe et est unique. Il est défini par

$$w_1 \textcircled{\wedge} w_2(A) = w_1(A) \wedge w_2(A), \quad \forall A \subset \Omega.$$

Définition (Règle prudente)

$$m_1 \textcircled{\wedge} m_2 = \textcircled{\bigcap}_{A \subset \Omega} A^{w_1(A) \wedge w_2(A)}.$$

Règle prudente

Définition

Théorème

Soient m_1 and m_2 deux fcts de masse non dogmatiques. L'élément le moins informatif de $\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2)$ au sens de \sqsubseteq_w existe et est unique. Il est défini par

$$w_1 \textcircled{\wedge} w_2(A) = w_1(A) \wedge w_2(A), \quad \forall A \subset \Omega.$$

Définition (Règle prudente)

$$m_1 \textcircled{\wedge} m_2 = \textcircled{\bigcap}_{A \subset \Omega} A^{w_1(A) \wedge w_2(A)}.$$

Règle prudente

Propriétés

- Commutative, associative
- **Idempotente** : $\forall m, m \otimes m = m$
- Distributivité de \oplus par rapport à \otimes :

$$(m_1 \oplus m_2) \otimes (m_1 \oplus m_3) = m_1 \oplus (m_2 \otimes m_3), \forall m_1, m_2, m_3.$$

L'élément d'évidence m_1 n'est compté qu'une fois !

- Absence d'élément neutre, mais $m_\Omega \otimes m = m$ ssi m est séparable.



Résumé

- Il existe une règle disjonctive duale de la règle prudente (règle hardie \vee).
- Quatre règles fondamentales :

Résumé

sources	distinctes	non distinctes
toutes fiables	\cap	\wedge
au moins une fiable	\cup	\vee

Plan de la présentation

- 1 Théorie des fonctions de croyance
 - Concepts fondamentaux
 - Fusion d'informations
 - Décomposition canonique et règle prudente
 - **Décision**
- 2 Application au diagnostic par RdF
 - Problème et motivations
 - Construction d'un classifieur crédibiliste
 - Décision
 - Exemple de fusion de classifieurs
- 3 Etudes de cas
 - Surveillance acoustique de structure
 - Diagnostic de machine asynchrone
 - Diagnostic de circuits de voie ferroviaires

Formalisation et cas probabiliste

- Etat de connaissance sur X décrit par une **distribution de probabilité** p .
- Ensemble d'**actions** $\mathcal{A} = \{a_1, \dots, a_s\}$.
- Le choix de l'action a_i si $X = \omega$ induit un **coût** $\lambda(a_i, \omega)$.
- **Risque** de a_i :

$$R_p(a_i) = \mathbb{E}_p[\lambda(a_i, \cdot)] = \sum_{\omega \in \Omega} \lambda(a_i, \omega) p(\omega).$$

- Choix de l'action a_i de **risque minimal**.

Approches crédibilistes

I - Risques inférieur et supérieur

- Soit m une fonction de masse normalisée, et \mathcal{P} l'ensemble des mesures de probabilité **compatibles**, c'est-à-dire vérifiant $bel(A) \leq P(A) \leq pl(A)$ pour tout $A \subseteq \Omega$.
- On peut définir pour chaque action :
 - un **risque inférieur** :

$$R_*(a_i) = \min_{P \in \mathcal{P}} R_p(a_i) = \sum_{A \subseteq \Omega} m(A) \min_{\omega \in A} \lambda(a_i, \omega)$$

- un **risque supérieur** :

$$R^*(a_i) = \max_{P \in \mathcal{P}} R_p(a_i) = \sum_{A \subseteq \Omega} m(A) \max_{\omega \in A} \lambda(a_i, \omega)$$

- Deux stratégies : minimisation du risque inférieur (optimiste) ou supérieur (pessimiste).



Approches crédibilistes

II - Risque pignistique

- Principe : **transformer m en une distribution de probabilité** au moment de prendre une décision.
- Modèle à deux niveaux : **crédal/pignistique**.
- Distribution de probabilité pignistique :

$$Betp(\omega) = \sum_{\{A \subseteq \Omega | \omega \in A\}} \frac{m(A)}{|A|}.$$

- Risque pignistique

$$R_{pig}(a_i) = R_{Betp}(a_i) = \sum_{\emptyset \neq A \subseteq \Omega} m(A) \left(\frac{1}{|A|} \sum_{\omega \in A} \lambda(a_i, \omega) \right)$$

- Propriété : $R_*(a_i) \leq R_{pig}(a_i) \leq R^*(a_i)$.

Plan de la présentation

- 1 Théorie des fonctions de croyance
 - Concepts fondamentaux
 - Fusion d'informations
 - Décomposition canonique et règle prudente
 - Décision
- 2 Application au diagnostic par RdF
 - **Problème et motivations**
 - Construction d'un classifieur crédibiliste
 - Décision
 - Exemple de fusion de classifieurs
- 3 Etudes de cas
 - Surveillance acoustique de structure
 - Diagnostic de machine asynchrone
 - Diagnostic de circuits de voie ferroviaires

Problème

- Soit un système S caractérisé à chaque instant par une variable y décrivant son **mode de fonctionnement**, et par un vecteur de mesures $\mathbf{x} \in \mathbb{R}^d$ ("**vecteur forme**").
- Connaissance disponible : ensemble d'observations (éventuellement imprécises ou incertaines) des variables \mathbf{x} et y , réalisées sur le système à différents instants.
- Problème : construire un **classifieur** (règle prédisant la valeur de la variable y pour toute nouvelle valeur de la variable \mathbf{x}).
- Ce problème peut être formalisé comme un problème de **décision dans l'incertain** (choix d'une action parmi un ensemble \mathcal{A} , sur la base de connaissances partielles).

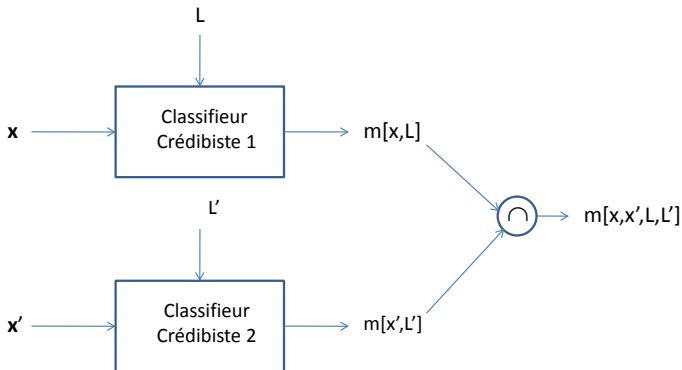


Intérêt des fonctions de croyance

- "Points forts" de la théorie des fonctions de croyance :
 - Modèle général de représentation d'**informations imprécises et incertaines** ;
 - Mécanismes permettant de **combiner des informations** issues de différentes sources, en tenant compte de leur fiabilité et de leur dépendance.
- Applications en diagnostic :
 - Problèmes pour lesquels l'information disponible est relativement "pauvre" : données **partiellement étiquetées** (connaissance imparfaite des modes de fonctionnement du système correspondant aux données d'apprentissage.
 - **Fusion d'informations** issues de différentes sources (capteurs, experts, algorithmes de traitement, etc.).



Approche crédibiliste



Plan de la présentation

- 1 Théorie des fonctions de croyance
 - Concepts fondamentaux
 - Fusion d'informations
 - Décomposition canonique et règle prudente
 - Décision
- 2 Application au diagnostic par RdF
 - Problème et motivations
 - **Construction d'un classifieur crédibiliste**
 - Décision
 - Exemple de fusion de classifieurs
- 3 Etudes de cas
 - Surveillance acoustique de structure
 - Diagnostic de machine asynchrone
 - Diagnostic de circuits de voie ferroviaires

Données

- Soit Ω l'ensemble des modes de fonctionnement du système (classes), et l'ensemble d'apprentissage

$$\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$$

où \mathbf{x}_i est le vecteur forme pour l'exemple e_i , et m_i une fonction de masse relative à la classe de l'exemple e_i .

- Cas particuliers :
 - $m_i(\{\omega_k\}) = 1$: étiquetage classique (**précis**) ;
 - $m_i(A) = 1$ pour $A \subseteq \Omega$: étiquetage **imprécis** ;
 - m_i est une fonction de masse bayésienne : étiquetage **probabiliste** ;
 - m_i est une fonction de masse consonante : étiquetage **possibiliste**, etc...

Problème

- Soit x le vecteur d'attributs pour un nouvel exemple e à classer.
- Problèmes :
 - Construire une **fonction de masse m relative à la classe y de e** .
 - Eventuellement, la combiner avec d'autres informations ;
 - Prendre une décision.

Construction d'une fonction de masse

- Hypothèse : soit α_j la **plausibilité que o et o_j ne soient pas dans la même classe**. On suppose $\alpha_j = \phi(d(\mathbf{x}, \mathbf{x}_j))$, où d est une distance, et ϕ une fonction croissante de \mathbb{R}^+ dans $[0, 1]$.
- Chaque exemple (\mathbf{x}_j, m_j) est une source d'information, qui doit être **affaiblie** avec le coefficient α_j :

$$m[\mathbf{x}, e_j] = \alpha_j m_j.$$

- Les n fonctions de masse sont ensuite **combinées conjonctivement** par \odot (exemples indépendants) où \oslash (exemples non indépendants) :

$$m[x, \mathcal{L}] = m[\mathbf{x}, e_1] * \dots * m[\mathbf{x}, e_n],$$

avec $*$ $\in \{\odot, \oslash\}$.

Cas d'un ensemble d'apprentissage classique

- Pour chaque exemple e_i , on a $m_i(\{\omega_j\}) = 1$ (l'exemple e_i appartient certainement à la classe ω_j).
- La fonction de masse correspondante est alors la **fonction de masse simple**

$$\begin{aligned}m[\mathbf{x}, e_i](\{\omega_j\}) &= \phi(d(\mathbf{x}, \mathbf{x}_i)) \\m[\mathbf{x}, e_i](\Omega) &= 1 - \phi(d(\mathbf{x}, \mathbf{x}_i)).\end{aligned}$$

- La fonction de masse combinée $m[x, \mathcal{L}] = \bigodot_{i=1}^n m[\mathbf{x}, e_i]$ a comme ensembles focaux \emptyset , $\{\omega_j\}$ ($j = 1, \dots, K$) et Ω .

Mise en œuvre

- Choix de la fonction ϕ : par exemple,

$$\phi(d) = \beta \exp(-\gamma d^2).$$

- Détermination des hyperparamètres β et γ de manière heuristique ou par apprentissage.
- Prise en compte uniquement des **k plus proches voisins** de \mathbf{x} dans \mathcal{L} (ou des vecteurs x_i dans un voisinage de \mathbf{x}) ;
- Compilation de \mathcal{L} sous forme de **r prototypes**, déterminés par apprentissage supervisé ou non supervisé.



Méthode des prototypes

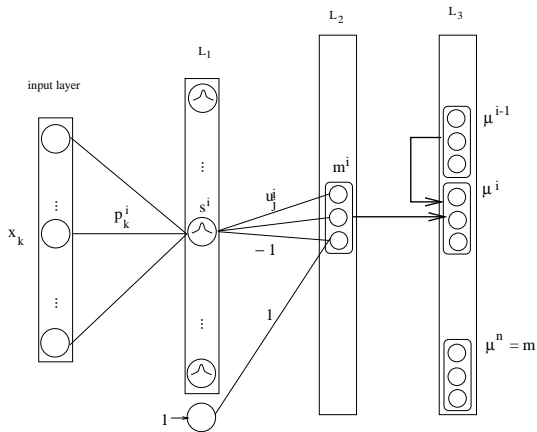
- Idée : pour accélérer les calculs, résumer l'ensemble d'apprentissage sous forme de r prototypes : $\mathbf{p}_1, \dots, \mathbf{p}_r$.
- Chaque prototype i a un degré d'appartenance u_{ik} à chaque classe ω_k avec $\sum_{k=1}^K u_{ik} = 1$
- La distance au prototype i induit une fonction de masse :

$$\begin{aligned}m[\mathbf{x}, \mathbf{p}_i](\{\omega_k\}) &= \alpha_i u_{ik} \exp(-\gamma_i \|\mathbf{x} - \mathbf{p}_i\|^2) \quad \forall k \\m[\mathbf{x}, \mathbf{p}_i](\Omega) &= 1 - \alpha_i \exp(-\gamma_i \|\mathbf{x} - \mathbf{p}_i\|^2)\end{aligned}$$

$$m[\mathbf{x}, \mathcal{L}] = \bigoplus_{i=1}^r m[\mathbf{x}, \mathbf{p}_i]$$

- Initialisation : centres mobiles par exemple.

Modèle connexionniste



Plan de la présentation

- 1 Théorie des fonctions de croyance
 - Concepts fondamentaux
 - Fusion d'informations
 - Décomposition canonique et règle prudente
 - Décision
- 2 Application au diagnostic par RdF
 - Problème et motivations
 - Construction d'un classifieur crédibiliste
 - **Décision**
 - Exemple de fusion de classifieurs
- 3 Etudes de cas
 - Surveillance acoustique de structure
 - Diagnostic de machine asynchrone
 - Diagnostic de circuits de voie ferroviaires

Décision

Ensemble d'apprentissage exhaustif (1)

- On suppose que **toutes les classes (modes de fonctionnement) sont représentées dans l'ensemble d'apprentissage.**
- Soit
 - $\Omega = \{\omega_1, \dots, \omega_K\}$ = ensemble des classes
 - $\mathcal{A} = \{a_0, a_1, \dots, a_K\}$ = ensemble des actions avec a_k = affectation à la classe ω_k et a_0 = rejet.

$$\lambda(a_k, \omega_\ell) = \begin{cases} 0 & k = \ell, \quad k, \ell \in \{1, \dots, K\} \\ 1 & k \neq \ell, \quad k, \ell \in \{1, \dots, K\} \\ c_0 & k = 0, \quad \ell \in \{1, \dots, K\} \end{cases}$$

Décision

Ensemble d'apprentissage exhaustif (2)

- Les **risques pignistiques** sont alors

$$R(a_k) = 1 - \text{Betp}(\omega_k)$$

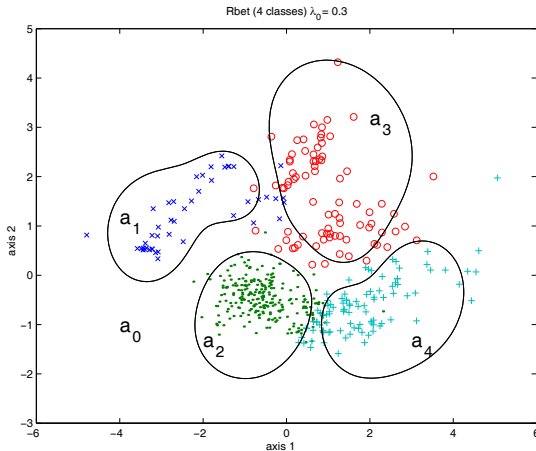
$$R(a_0) = c_0$$

- Règle de décision

$$D(\mathbf{x}) = \begin{cases} a_k & \text{si } \text{Betp}(\omega_k) > \text{Betp}(\omega_\ell) \quad \forall \ell \neq k \\ & \text{et } \text{Betp}(\omega_k) > 1 - \lambda_0 \\ a_0 & \text{sinon.} \end{cases}$$

Décision

Exemple



Décision

Ensemble d'apprentissage non exhaustif (1)

- On suppose l'existence d'une **classe inconnue** ω_U non représentée dans $\mathcal{L} : \Omega = \{\omega_1, \dots, \omega_K, \omega_U\}$ et

$$\mathcal{A} = \{a_0, a_1, \dots, a_K, a_U\}$$

- Il n'y a dans l'ensemble d'apprentissage aucun élément d'évidence en faveur de l'hypothèse ω_U , d'où $bel(\{\omega_U\}) = 0$. Cependant

$$BetP(\omega_U) = \frac{m(\Omega)}{K + 1}$$

- Pour certains choix de la fonction de coût, l'affectation à la classe ω_U (**détection de nouveauté**) est possible.

Décision

Ensemble d'apprentissage non exhaustif (2)

- Soit la matrice de coûts :

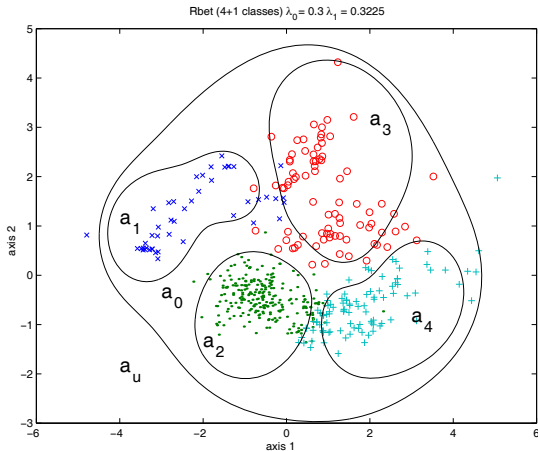
	a_0	a_1	a_2	...	a_{K-1}	a_K	α_u
ω_1	λ_0	0	1	...	1	1	λ_1
ω_2	λ_0	1	0	...	1	1	λ_1
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
ω_{K-1}	λ_0	1	1	...	0	1	λ_1
ω_K	λ_0	1	1	...	1	0	λ_1
ω_u	λ_0	1	1	...	1	1	0

- Le risque de l'action a_u est alors :

$$R(a_u) = \lambda_1 \left(1 - \frac{m(\Omega)}{K+1} \right)$$

Décision

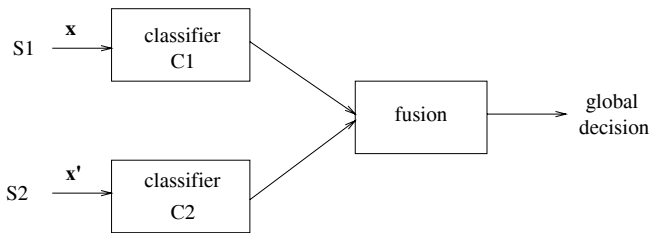
Exemple



Plan de la présentation

- 1 Théorie des fonctions de croyance
 - Concepts fondamentaux
 - Fusion d'informations
 - Décomposition canonique et règle prudente
 - Décision
- 2 Application au diagnostic par RdF
 - Problème et motivations
 - Construction d'un classifieur crédibiliste
 - Décision
 - Exemple de fusion de classifieurs
- 3 Etudes de cas
 - Surveillance acoustique de structure
 - Diagnostic de machine asynchrone
 - Diagnostic de circuits de voie ferroviaires

Problème



- $K = 2$ classes
- $\mathbf{x} \in \mathbb{R}^5, \mathbf{x}' \in \mathbb{R}^3$, distributions normales, indépendantes conditionnellement à la classe
- Ensemble d'apprentissage : $n = 60$, validation croisée : $n_{CV} = 100$
- test : 5000 exemples.

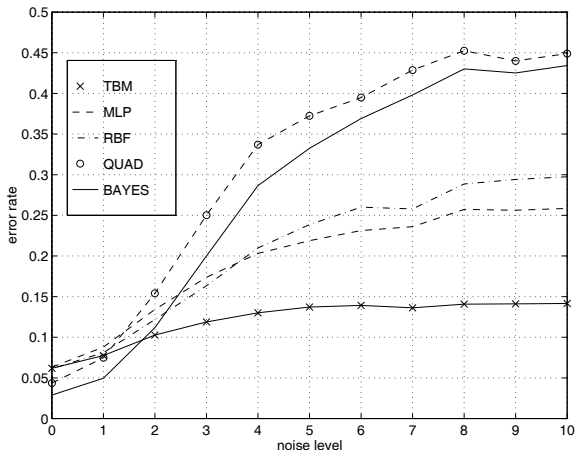
Résultats

Erreurs de tests : données non corrompues

Method	\mathbf{x} seul	\mathbf{x}' seul	\mathbf{x} et \mathbf{x}'
TBM	0.106	0.148	0.061
MLP	0.113	0.142	0.063
RBF	0.133	0.159	0.083
QUAD	0.101	0.141	0.049
BAYES	0.071	0.121	0.028

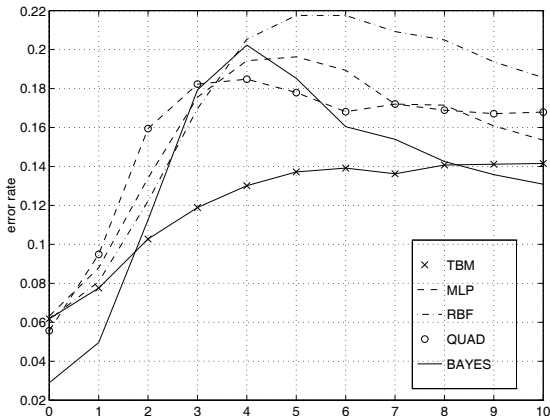
Résultats

Erreurs de test : $\mathbf{x} + \epsilon, \epsilon \sim \mathcal{N}(0, \sigma^2)$



Résultats

Erreurs de test : $\mathbf{x} + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma^2)$ avec rejet

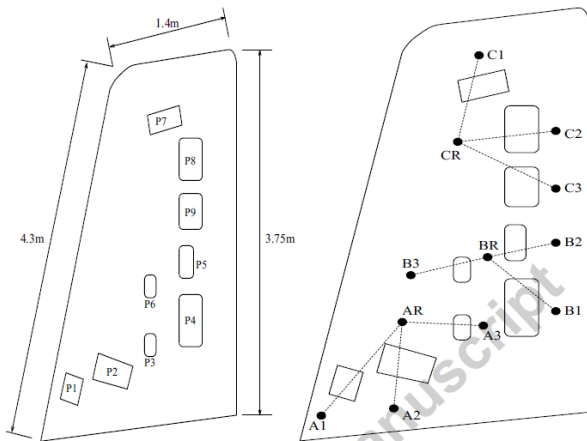


Plan de la présentation

- 1 Théorie des fonctions de croyance
 - Concepts fondamentaux
 - Fusion d'informations
 - Décomposition canonique et règle prudente
 - Décision
- 2 Application au diagnostic par RdF
 - Problème et motivations
 - Construction d'un classifieur crédibiliste
 - Décision
 - Exemple de fusion de classifieurs
- 3 Etudes de cas
 - **Surveillance acoustique de structure**
 - Diagnostic de machine asynchrone
 - Diagnostic de circuits de voie ferroviaires

Problème

K. Worden, G. Manson, T. Denœux. An evidence-based approach to damage location on an aircraft structure. *Mechanical Systems and Signal Processing*, in press, 2009.



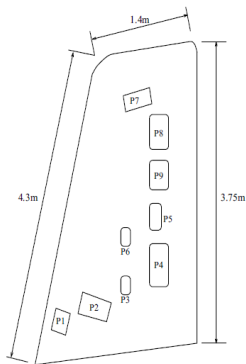
Données

- 9 défauts (retrait des 9 panneaux d'inspection) → $K = 9$ classes.
- Mesures de transmissibilité par 12 capteurs acoustiques.
- 100 mesures pour chaque retrait d'un panneau d'inspection, avec 2 répétitions → 1800 observations.
- Extraction de caractéristiques correspondant à certaines fréquences puis sélection par analyse de donnée et algorithme génétique de 4 variables → $\mathbf{x} \in \mathbb{R}^4$.
- Données partitionnées en ensembles d'apprentissage, de validation et de test de même taille : 594 exemples, 66 exemples par classe.

Résultats

Réseau de neurones évidentiel

29 prototypes, taux de bonnes classifications 89.7 %.

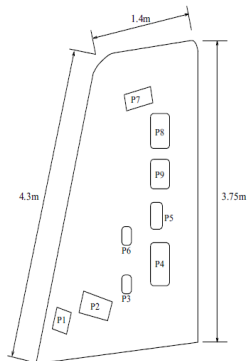


Prediction	1	2	3	4	5	6	7	8	9	Θ
True Class 1	54	5	5	0	0	0	2	0	0	0
True Class 2	0	63	0	0	2	0	0	0	0	1
True Class 3	6	1	56	2	0	0	0	0	0	1
True Class 4	5	0	1	55	0	3	0	2	0	0
True Class 5	0	0	0	0	65	0	0	1	0	0
True Class 6	2	2	2	4	0	54	1	0	0	1
True Class 7	0	1	1	0	0	0	61	2	1	0
True Class 8	0	0	1	0	1	0	0	62	1	1
True Class 9	0	0	0	0	0	0	0	3	63	0

Résultats

Perceptron multi-couches

19 neurones cachés, taux de bonnes classifications 89.2 %.

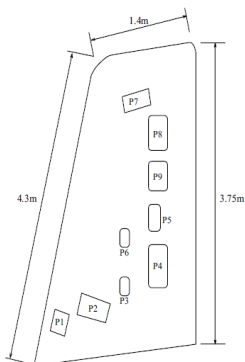


Prediction	1	2	3	4	5	6	7	8	9
True Class 1	61	1	0	0	0	0	1	0	0
True Class 2	0	63	0	0	3	0	0	0	0
True Class 3	1	0	48	8	0	5	7	2	0
True Class 4	0	1	3	56	0	2	0	4	0
True Class 5	0	0	0	0	66	0	0	0	0
True Class 6	4	1	0	9	0	52	0	0	0
True Class 7	1	0	0	0	0	0	59	5	1
True Class 8	1	0	0	0	1	0	1	59	4
True Class 9	0	0	0	0	0	0	0	5	61

Résultats

Réseau de neurones évidentiel + Perceptron multi-couches

Taux de bonnes classifications 92.3 %.



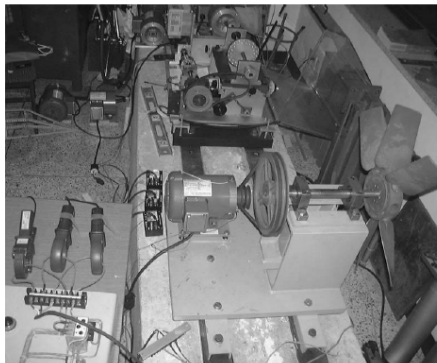
Prediction	1	2	3	4	5	6	7	8	9
True Class 1	63	2	0	0	0	0	1	0	0
True Class 2	0	64	0	0	2	0	0	0	0
True Class 3	0	1	55	6	0	0	2	2	0
True Class 4	0	1	2	59	0	1	0	3	0
True Class 5	0	0	0	0	65	0	0	0	1
True Class 6	1	2	0	6	0	57	0	0	0
True Class 7	0	0	0	0	0	0	60	5	1
True Class 8	1	0	0	0	1	0	0	63	2
True Class 9	0	0	0	0	0	0	0	4	62

Plan de la présentation

- 1 Théorie des fonctions de croyance
 - Concepts fondamentaux
 - Fusion d'informations
 - Décomposition canonique et règle prudente
 - Décision
- 2 Application au diagnostic par RdF
 - Problème et motivations
 - Construction d'un classifieur crédibiliste
 - Décision
 - Exemple de fusion de classifieurs
- 3 Etudes de cas
 - Surveillance acoustique de structure
 - **Diagnostic de machine asynchrone**
 - Diagnostic de circuits de voie ferroviaires

Problème

B.-S. Yang, K.J. Kim. Application of Dempster-Shafer theory in fault diagnosis of induction motors using vibration and current signals. *Mechanical Systems and Signal Processing*, vol. 20, 403-420, 2006.



Dispositif expérimental

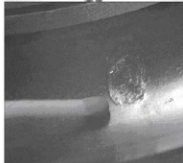
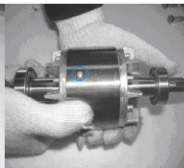
- Six moteurs asynchrones (0.5 kW, 60 Hz, 4 pôles).
- Fonctionnement en pleine charge et en régime permanent.
- Un moteur sans défaut (référence), les 5 autres utilisés pour reproduire des défauts.
- **Signaux vibratoires** acquis par des accéléromètres dans les directions axiale, horizontale et verticale.
- **Signaux électriques** d'entrée mesurés par des sondes de courant.
- 20 mesures pour chacun des **8 modes de fonctionnement** (nominal + 7 défauts) → **160 exemples**.

Défauts générés

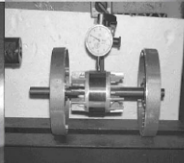
Rotor unbalance



Broken rotor bar



Faulty bearing



Bowed rotor

Fault condition

Normal condition (without fault)

Mechanical imbalance (15.64 g · cm)

Phase imbalance (added 4.3Ω to one phase)

Broken rotor bar (12 EA)

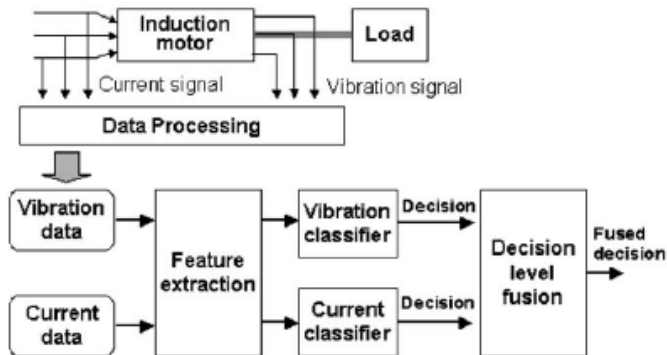
Bowed rotor (0.075 mm)

Fault bearing (a spalling on outer raceway)

Angular misalignment (0.48°)

Parallel misalignment (0.1 mm)

Architecture du système



Extraction de caractéristiques

- 13 caractéristiques extraites dans les domaine temporel et fréquentiel pour chaque signal.
- Sélection de caractéristiques → **3 caractéristiques** retenues pour chaque type de signal :
 - Signal électrique : RMS, kurtosis, RVF (root variance frequency)
 - Signal vibratoire : RMS, RVF, entropie

Distributions des caractéristiques (1)

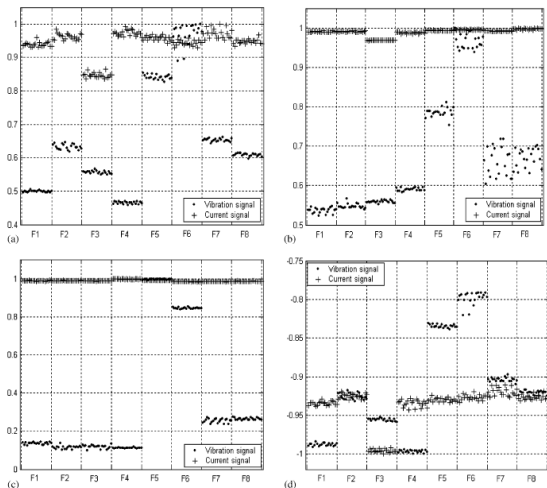


Fig. 4. Typical feature extraction of vibration signals and current signals. (a) RMS. (b) Kurtosis. (c) RVF. (d) Entropy

Distributions des caractéristiques (2)

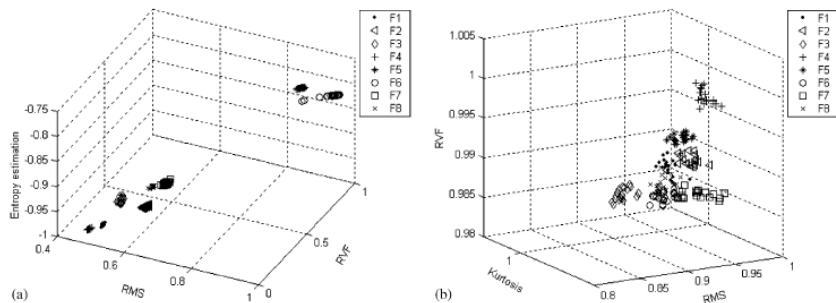


Fig. 5. Feature clusters of vibration and current signals.

Apprentissage

- 160 exemples partitionnés en
 - 120 exemples d'apprentissage (15 par classe) ;
 - 40 exemples de test (5 par classe).
- Un prototype par classe, initialisé par l'algorithme des centres mobiles.



Résultats

Comparison of classification results

Classifier	Classification results								
	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>F4</i>	<i>F5</i>	<i>F6</i>	<i>F7</i>	<i>F8</i>	Average
Vibration alone	0.964	0.970	0.916	0.955	0.976	0.969	0.455	0.484	0.836
Current alone	0.500	0.162	0.955	0.665	0.317	0.334	0.377	0.374	0.461
Vibration and current	0.981	0.938	0.996	0.988	0.976	0.957	0.670	0.675	0.898

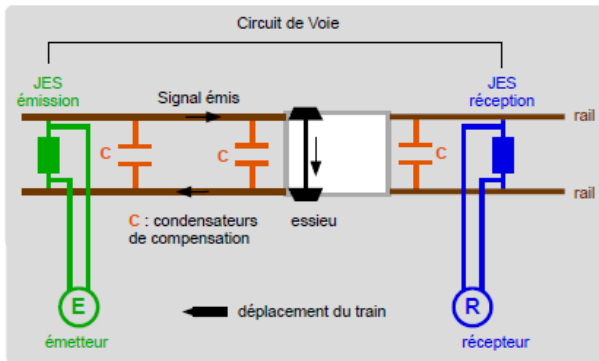
Plan de la présentation

- 1 Théorie des fonctions de croyance
 - Concepts fondamentaux
 - Fusion d'informations
 - Décomposition canonique et règle prudente
 - Décision
- 2 Application au diagnostic par RdF
 - Problème et motivations
 - Construction d'un classifieur crédibiliste
 - Décision
 - Exemple de fusion de classifieurs
- 3 Etudes de cas
 - Surveillance acoustique de structure
 - Diagnostic de machine asynchrone
 - Diagnostic de circuits de voie ferroviaires

Description du système

A. Debiolles. *Diagnostic de systèmes complexes à base de modèle interne, reconnaissance de formes et fusion d'informations. Application au diagnostic des circuits de voie ferroviaires.* Thèse de doctorat, UTC, février 2007.

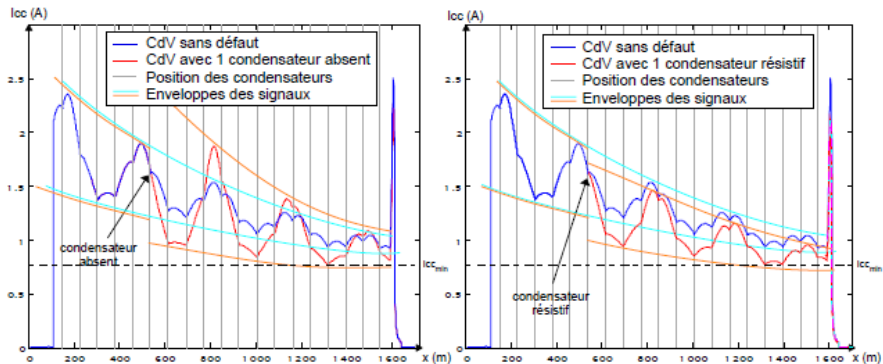
Fonction : détecter la présence d'un train sur la voie.



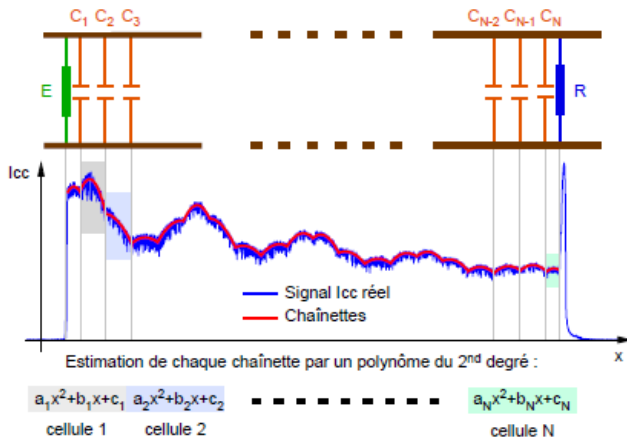
Problème

- **Détecter** la présence d'un défaut sur un condensateur, et **localiser** le défaut.
- Types de défauts :
 - Arrachement (travaux de maintenance) ;
 - Mauvaise connexion entre le condensateur et le rail (corrosion, ...)
 - Perte par conduction ou hystérésis électrique (vieillessement).
- Mesures de courant de court-circuit (I_{cc}) par véhicule d'inspection.

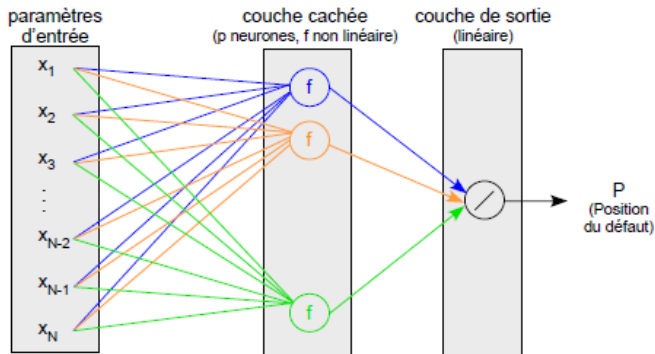
Courant I_{cc}



Extraction de caractéristiques

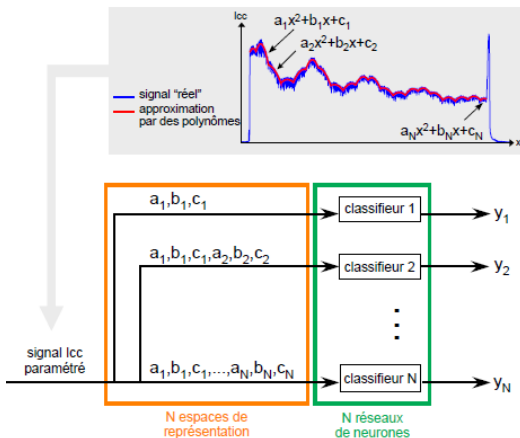


Approche par régression



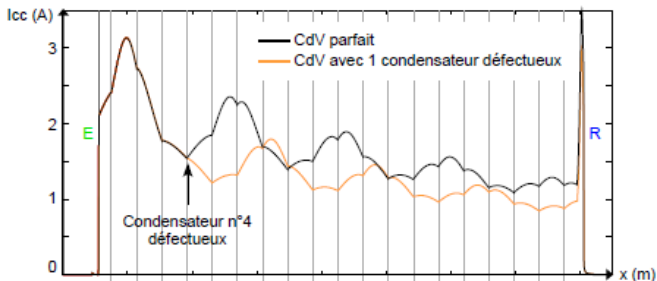
Approche par fusion

Principe



Approche par fusion

Codage des sorties désirées



Numéro de cellule	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
CdV sans défaut	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CdV avec défaut sur le condensateur n°4	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Classifieur	RN_1	-----																RN_N	

Codage thermométrique

Approche par fusion

Modélisation (1)

- Problème : **prédire la position X du défaut**, à valeurs dans $\Omega = \{1, 2, \dots, N + 1\}$ avec $N =$ nombre de condensateurs. Par convention $X = N + 1$ si absence de défaut.
- Soit p_i la sortie du réseau de neurone i . Une valeur proche de 1 signifie qu'il y a un défaut en amont du condensateur i . On pose :

$$\begin{aligned}m_i([1, i]) &= \alpha_i p_i \\m_i([i + 1, N + 1]) &= \alpha_i (1 - p_i) \\m_i(\Omega) &= 1 - \alpha_i,\end{aligned}$$

où $1 - \alpha_i$ est un **coefficient d'affaiblissement** dépendant de la fiabilité du classifieur i .

Approche par fusion

Modélisation (2)

- Détermination du coefficient d'affaiblissement :

$$1 - \alpha_j = \frac{0.5MSE_j}{\max_j(MSE_j)}.$$

- Combinaison des sorties par la règle de Dempster :

$$m = m_1 \circledast \dots \circledast m_N.$$

- Décision : **maximum de probabilité pignistique**

$$\hat{i} = \arg \max_i \text{Betp}(i) = \sum_{A \subseteq \Omega, i \in A} \frac{m(A)}{|A|}.$$



Données et méthodologie

- Simulation de 4256 signaux bruités, $N = 19$ condensateurs.
- 608 signaux sans défaut, 3648 avec un défaut sur un condensateur (résistance entre 1Ω et ∞).
- Partition en 3 ensembles (apprentissage, validation, test).
- Test : 1064 exemples.
- Nombres de neurones cachés déterminés par validation croisée.

Evaluation des résultats

		Truth	
		Fault position k	No fault
Decision	Fault position k	N_{CD}	FA
	Fault position $j \neq k$	N_{FL}	N_{FA}
	No fault	ND	CR
		N_{ND}	N_{CR}

$$t_{CD} = \frac{N_{CD} + N_{CR} + N_{FL}}{N_0 + N_1}, \quad t_{CL} = \frac{N_{CD} + N_{CR}}{N_{CD} + N_{CR} + N_{FL}}$$

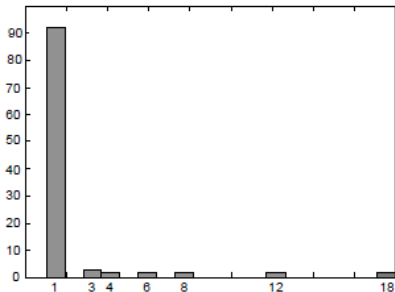
$$t_{FA} = \frac{N_{FA}}{N_0}, \quad t_{ND} = \frac{N_{ND}}{N_1}$$

avec N_0 = nombre de signaux sans défaut, et N_1 = nombre de signaux avec défaut.

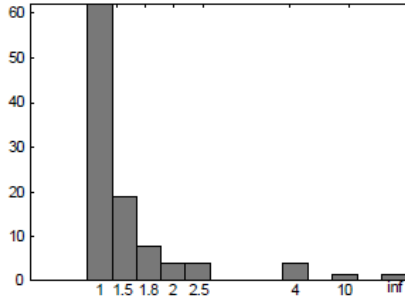
Résultats

	t_{CD}	t_{CL}	t_{FA}	t_{ND}
Régression	84.7	56.1	59.2	1.17
Fusion	95.2	98.3	1.07	6.36

Erreurs de localisation



Résistances en cas de FL



Conclusion

Théorie des fonctions de croyance

- La théorie des fonctions de croyance (TFC) est un **formalisme général** qui étend les notions d'ensemble et de distribution de probabilité : représentation simultanée de l'imprécision et de l'incertitude.
- La TFC n'est pas un ensemble de recettes ad hoc : elle peut être construite rigoureusement à partir d'un petit nombre de principes (engagement minimal, ...).
- La TFC comporte un grand nombre (une infinité) de **règles de fusion** reposant sur différentes hypothèses relatives à la fiabilité et à l'indépendance des sources.

Conclusion

TFC en diagnostic

- Existence d'algorithmes d'apprentissage permettant de construire des **classifieurs crédibilistes**.
- Avantages :
 - Prise ne compte de l'**imprécision et de l'incertitude de l'étiquetage**.
 - **Fusion d'informations**.
- En diagnostic par reconnaissance de formes, la TFC permet de fusionner les informations issues de classifieurs
 - construits par des **algorithmes** différents ;
 - ayant **espaces d'entrée** différents (multi-capteurs) ;
 - ayant des **espaces de sortie** différents.
- La fusion par d'**autres règles que la règle de Dempster** est une voie intéressante encore peu explorée.

Références I

cf. <http://www.hds.utc.fr/~tdenoeux>



T. Denœux.

A neural network classifier based on Dempster-Shafer theory.

IEEE Transactions on Systems, Man and Cybernetics A,
Vol. 30, Issue 2, Pages 131-150, 2000.



T. Denœux and P. Smets.

Classification using Belief Functions : the Relationship between the Case-based and Model-based Approaches.

IEEE Transactions on Systems, Man and Cybernetics B,
Vol. 36, Issue 6, Pages 1395-1406, 2006.



Références II

cf. <http://www.hds.utc.fr/~tdenoeux>



T. Denœux.

Conjunctive and Disjunctive Combination of Belief Functions Induced by Non Distinct Bodies of Evidence. *Artificial Intelligence*, Vol. 172, pages 234-264, 2008.



A. Debiolles, L. Oukhellou, T. Denœux and P. Aknin.

Output coding of spatially dependent subclassifiers in evidential framework. Application to the diagnosis of railway track-vehicle transmission system.

Proceedings of FUSION '06, Florence, Italy, July 2006.

