Modélisation du conflit dans les bases de données évidentielles

Mouna Chebbah*, Arnaud Martin**
Boutheina Ben Yaghlane***

*LARODEC, ISG Tunis, 41 Rue de la Liberté, Cité Bouchoucha 2000 Le Bardo, Tunisie
Mouna.Chebbah@gnet.tn
**E^3I^2, EA3876, ENSIETA, 2 rue François Verny, 29806 Brest Cedex 9
Arnaud.Martin@ensieta.fr
***LARODEC, IHEC Carthage, Carthage Présidence 2016, Tunisie
boutheina.yaghlane@ihec.rnu.tn

Résumé. La combinaison de différentes sources imparfaites fait inévitablement apparaître du conflit. Dans le cadre de la théorie des fonctions de croyance, la résolution du conflit se fait avant ou pendant l'étape de combinaison. Lors de la combinaison, plusieurs règles permettent d'éliminer le conflit en le redistribuant sur les informations disponibles de différentes manières. Par contre la gestion du conflit avant la combinaison revient à affaiblir les informations avant de les combiner ce qui nécessite une connaissance préalable des degrés de fiabilité des sources. Dans cet article, nous proposons une nouvelle méthode d'estimation de la fiabilité d'une source à partir de toutes les informations disponibles dans une base de données évidentielle. Les expérimentions sur des données radar réelles ont montré une amélioration remarquable des fiabilités des sources après affaiblissement.

1 Introduction

Les bases de données permettent de stocker une grande quantité d'information qui sont, la plupart du temps, incertaines ou imprécises. Pour aborder ce problème, des bases de données évidentielles ont été proposées par Hewawasam et al. (2005) et Bach Tobji et al. (2008). La fusion d'informations permet, d'une part, de réduire la quantité d'informations disponibles dans les bases de données évidentielles, et d'autre part, d'aider les utilisateurs qu'ils soient humains ou logiciels à la prise de décision en résumant les degrés de confiances en un seul facilement interprétable. Un degré de confiance associé à chaque information incertaine doit alors être défini en vu d'être stocké dans une base de données évidentielle.

Fusionner revient à combiner différentes informations, pouvant être imparfaites, en provenance de diverses sources. Pour ce faire la *théorie des fonctions de croyance*, introduite par Dempster (1967) et Shafer (1976), offre plusieurs avantages. En effet, cette dernière est utilisée pour sa robustesse en terme de représentation de données incertaines et de combinaison. La prise en considération de plusieurs sources hétérogènes lors de la combinaison peut induire l'apparition d'un conflit dû à une contradiction au niveau des informations fournies.

L'existence du conflit a favorisé l'apparition de plusieurs méthodes visant à le résoudre dans le cadre de la théorie des fonctions de croyance. Certaines méthodes proposent la résolution du conflit lors de la combinaison en proposant des règles adéquates. Ces règles permettent à la fois de combiner et de redistribuer le conflit de manières différentes sur les informations disponibles.

Une autre façon de gérer le conflit est de le réduire avant de combiner en affaiblissant les informations fournies par une source avec son degré de fiabilité. Cette méthode permet de différencier les informations fournies par une source fiable de celles fournies par une source de fiabilité moindre. Cette méthode nécessite une connaissance préalable des degrés de fiabilité des sources. Il est donc nécessaire de pouvoir estimer la fiabilité d'une source. Martin et al. (2008a) proposent une approche sans *a priori* fondée sur les informations fournies par les sources.

Dans cet article, nous proposons une méthode d'estimation de la fiabilité d'une source à partir de toutes les fonctions de masse qu'elle fournit et non pas à partir d'une seule fonction. Notre méthode est particulièrement applicable sur les bases de données évidentielles du moment où ces dernières permettent de stocker toutes les fonctions de masse fournies par une source. Nous proposons également d'améliorer le niveau de fiabilité des différentes sources.

Le reste de cet article est organisé comme suit : dans la deuxième section nous présentons brièvement les notions de base de la théorie des fonctions de croyance ensuite nous définissons les bases de données évidentielles dans la troisième section. La quatrième section présente notre méthode d'estimation du conflit et de la fiabilité d'une source et enfin nous présentons les résultats expérimentaux sur des données radar dans la cinquième et dernière section.

2 La théorie des fonctions de croyance

2.1 Formalisme

La théorie des fonctions de croyance, appelée aussi théorie de l'évidence initiée par Dempster (1967) et Shafer (1976), est un outil robuste pour la représentation des données imparfaites (imprécises et/ou incertaines). Nous présentons ici quelques concepts de base de cette théorie.

Soit $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ un ensemble fini de toutes les hypothèses possibles ω_i pour un problème donné. L'ensemble Ω représente *l'ensemble de discernement* ou *l'univers de discours* du problème en question.

Une fonction de masse est définie sur l'ensemble de tous les sous-ensembles possibles de Ω, noté 2^Ω et affecte à chaque sous-ensemble une valeur entre 0 et 1 représentant sa masse de croyance élémentaire. Une fonction de masse m est donc telle que :

$$m : 2^\Omega \mapsto [0, 1] \tag{1}$$

Un sous-ensemble ayant une masse de croyance non-nulle est *un élément focal*. Une fonction de masse doit satisfaire les conditions suivantes :

$$\sum_{X \subseteq \Omega} m(X) = 1 \tag{2}$$

On impose aussi en général $m(\emptyset) = 0$ qui permet de rester en monde fermé.

Une fonction de masse permet de représenter les connaissances incertaines et imprécises fournies par un expert (une source, un classifieur, ...). La masse affectée à un élément focal X représente le degré de croyance élémentaire d'une source à ce que la valeur réelle de l'attribut en question soit incluse ou égale à X, c'est donc son degré de croyance élémentaire en X.

La fonction de croyance (ou de crédibilité) *bel* représente le degré de croyance minimal affecté à un sous-ensemble de 2^Ω justifié par les informations disponibles. *bel(A)* mesure le degré auquel les informations données par une source ($B \subseteq A$) soutiennent A.

$$bel : 2^\Omega \to [0,1] \quad (3)$$

$$A \mapsto \sum_{B \subseteq A, B \neq \emptyset} m(B) \quad (4)$$

La fonction de plausibilité *pl* représente le degré de croyance alloué aux propositions non contradictoires avec A. C'est le degré de croyance maximal en A. La fonction de plausibilité est la fonction duale de la fonction de crédibilité.

$$\begin{array}{ll} \text{pl :} & 2^\Omega \to [0,1] \\ \text{A} \mapsto & pl(A) = \sum_{A \cap B \neq \emptyset} m(B) \end{array} \quad (5)$$

2.2 La règle conjonctive normalisée

La théorie des fonctions de croyance représente un outil robuste de combinaison de différentes fonctions de masse définies sur le même ensemble de discernement et fournies par différentes sources. La combinaison permet de confronter différents avis d'experts afin d'en obtenir un seul en vue de la décision.

La règle conjonctive de combinaison, présentée dans le modèle de croyances transférables par Smets et Kennes (1994), permet de combiner deux fonctions de masse distinctes m_1^Ω et m_2^Ω. Elle est définie comme suit :

$$m_{1 \bigcirc\!\!\!\!\cap\, 2}^\Omega(A) = (m_1^\Omega \bigcirc\!\!\!\!\cap\, m_2^\Omega)(A) = \sum_{B \cap C = A} m_1^\Omega(B) \times m_2^\Omega(C) \quad (6)$$

Cette règle est non normalisée du moment qu'elle autorise l'affectation d'une masse nulle à l'ensemble vide après la combinaison. Elle est applicable sous l'hypothèse du monde ouvert dans lequel la masse attribuée à l'ensemble vide représente le degré de croyance en une hypothèse inconnue et non énumérée dans Ω. Une normalisation de cette règle est présentée par Dempster (1967).

La fonction de masse $m_{1\oplus 2}^\Omega$ résultat de la combinaison de m_1^Ω et m_2^Ω par la règle de Dempster (1967) est obtenue comme suit :

$$m_{1\oplus 2}^\Omega(A) = (m_1^\Omega \oplus m_2^\Omega)(A) = \begin{cases} \frac{m_{1 \bigcirc\!\!\!\!\cap\, 2}^\Omega(A)}{1 - m_{1 \bigcirc\!\!\!\!\cap\, 2}^\Omega(\emptyset)} & , \forall A \subseteq \Omega \text{ si } A \neq \emptyset \\ 0 & , \text{si } A = \emptyset \end{cases} \quad (7)$$

Cette règle est normalisée et elle est applicable avec l'hypothèse du monde fermé dans lequel on suppose que toutes les valeurs possibles que peuvent prendre un attribut sont énumérées dans l'ensemble de discernement.

Il existe toute une panoplie de règles de combinaison par exemple, la règle de combinaison de Yager (1987) et la règle de combinaison de Dubois et Prade (1988).

2.3 L'affaiblissement

La majorité des règles de combinaison ne font pas la distinction entre le conflit existant entre les sources et l'auto-conflit dû à la non-idempotence de la règle de combinaison utilisée. Une des origines du conflit est la non fiabilité d'au moins une des sources. La non fiabilité d'une source peut être réglée par l'affaiblissement des fonctions de masse avant la combinaison. Du moment qu'on arrive à quantifier les fiabilités α de chaque source, on peut affaiblir les fonctions de masse associées comme suit :

$$\begin{cases} m^{\alpha}(A) = \alpha\, m^{\Omega}(A), \text{ pour } A \subset \Omega \\ m^{\alpha}(\Omega) = (1-\alpha) + \alpha\, m^{\Omega}(\Omega) \end{cases} \tag{8}$$

avec α le degré de fiabilité de la source.

3 Les bases de données évidentielles

Une base de données évidentielle est une base de données qui contient des données parfaites, imparfaites ou même des données manquantes. L'imperfection (incertitude et/ou imprécision) dans les bases de données évidentielles est représentée grâce à la théorie des fonctions de croyance précédemment décrite.

Formellement, une base de données évidentielle est une base de données ayant X attributs (colonnes) et Y enregistrements (lignes), chaque attribut j ($1 \leq j \leq X$) possède un domaine D_j représentant toutes les valeurs que peut prendre cet attribut : *C'est son ensemble de discernement* tel que défini par Bach Tobji et al. (2008). Une base de données évidentielle doit contenir au moins un attribut évidentiel qui prendra des valeurs évidentielles décrite par une fonction de masse au lieu d'une valeur certaine et précise. Une valeur évidentielle V_{ij} de l'enregistrement i ($1 \leq i \leq Y$) pour l'attribut j ($1 \leq j \leq X$) est une fonction de masse m_{ij} telle que :

$$\begin{array}{l} m_{ij} : 2^{D_j} \rightarrow [0,1] \text{ avec :} \\ m_{ij}(\emptyset)\text{=0 et } \sum_{x \subseteq D_j} m_{ij}\text{(x)=1} \end{array} \tag{9}$$

Les bases de données évidentielles sont utilisées dans différents domaines notamment pour le stockage des fonctions de masse de différents classifieurs présenté par Hewawasam et al. (2005).

4 Estimation du conflit et de la fiabilité d'une source

Une base de données évidentielle stocke différentes fonctions de masse fournies par une source. Avec la présence de plusieurs sources, experts ou classifieurs il y aura autant de bases de données évidentielles que de sources donc une quantité énorme de données à traiter. L'intégration des différentes bases de données en une seule permet de réduire la quantité d'information disponible afin de faciliter les requêtes sur la base et la prise de décision éventuelle. Lors de l'intégration de différentes bases de données évidentielles, le principal problème rencontré réside dans l'intégration des valeurs évidentielles surtout quand elles sont conflictuelles.

Dans cet article, nous nous concentrons sur la fusion de plusieurs fonctions de masse stockées dans différentes bases de données. Nous proposons aussi d'enrichir ces bases de données par des informations sur le niveau de fiabilité des sources et de fiabilité des combinaisons. Ces fiabilités seront des indicateurs importants sur le degré de confiance des résultats des requêtes effectuées sur la base. L'utilisation d'une règle de combinaison paraît une solution simple permettant à la fois de combiner plusieurs fonctions de masse et de résoudre le conflit. Cependant, la résolution du conflit se fait alors par la redistribution de la masse affectée à l'ensemble vide sur les éléments focaux de différentes manières selon la règle utilisée sans prendre en considération les fiabilités des sources et le degré de véracité des fonctions de masse fournies.

Le degré de fiabilité de la source doit être pris en considération avant la combinaison pour prévenir le conflit au maximum. La difficulté réside dans la quantification de la fiabilité d'une source afin d'en tenir compte avant la combinaison en affaiblissant ses fonctions de masse. Dans une base de données évidentielle, plusieurs fonctions de masse relatives à une source y sont stockées, d'où la nécessité d'attribuer un degré de fiabilité global à la source qui prend en considération toutes les fonctions de masse qu'elle fournies.

4.1 La mesure du conflit

Martin et al. (2008b) considèrent que plus deux fonctions de masse sont éloignées plus les sources sont en conflit. Ainsi, une mesure de distance entre différentes fonctions de masse permet de quantifier le conflit entre leurs sources.

La distance de Jousselme et al. (2001) est utilisée dans cet article parce qu'elle permet de tenir compte des spécificités des fonctions de croyance. Elle est donnée par :

$$d(m_1, m_2) = \sqrt{\frac{1}{2}(m_1 - m_2)^t \underline{\underline{D}}(m_1 - m_2)} \tag{10}$$

avec :

$$D(A, B) = \begin{cases} 1, \text{ si A=B=}\emptyset \\ \frac{|A \cap B|}{|A \cup B|}, \forall A, B \in 2^\Omega \end{cases} \tag{11}$$

Le conflit entre deux sources S_1 et S_2 n'est autre que la distance séparant leurs fonctions de masse m_1 et m_2 :

$$Conf(S_1, S_2) = d(m_1, m_2) \tag{12}$$

Avec la présence de plusieurs sources, la mesure de conflit correspond à la distance entre la distribution de masse fournie par une source donnée et les autres distributions. Cette mesure de conflit peut être calculée de deux manières différentes. La première méthode consiste à calculer la moyenne des distances entre la fonction de masse de la source S_j en question avec les $s-1$ autres fonctions de masse ϵ :

$$Conf(j, \epsilon) = \frac{1}{s-1} \sum_{j=1, j' \neq j}^{s} Conf(j, j') \tag{13}$$

La seconde correspond au calcul de la distance entre la fonction de masse fournie par la source S_j et la fonction de masse m_{s-1} résultat de la combinaison des fonctions de masse des

$s-1$ différentes sources autre que la source en question :

$$Conf(j,s) = d(m_j, m_{s-1}) \tag{14}$$

4.2 Estimation de la fiabilité relative d'une source

L'affaiblissement peut être utilisé afin de prendre en considération la non-fiabilité relative d'une source avant la combinaison pour éliminer ou réduire le conflit qui pourra apparaître après. Nous faisons ici l'hypothèse que le conflit est issu de la non-fiabilité des sources.

Pour pouvoir affaiblir une fonction de masse, l'estimation de la fiabilité est nécessaire. La méthode d'estimation de fiabilité proposée par Martin et al. (2008b) est fondée sur la mesure du conflit. La *fiabilité relative* α_j d'une source S_j est une fonction décroissante de son conflit avec les autres sources telle que :

$$\alpha_j = (1 - Conf(j,s)^{\lambda})^{\frac{1}{\lambda}} \tag{15}$$

avec λ un réel quelconque.

4.3 Estimation de la fiabilité absolue d'une source

Pour s sources, l'estimation de la fiabilité relative revient à calculer pour chacune le conflit de sa fonction de masse par rapport aux autres qui servira à estimer sa fiabilité, cette fiabilité est utilisée pour affaiblir la fonction de masse correspondante avant la combinaison.

Dans une base de données évidentielle, différentes fonctions de masse fournies par la même source sont stockées. Il faut alors tenir compte de toutes les fonctions de masse pour estimer la fiabilité absolue de la source que nous introduisons ici. La fiabilité relative ne prend en compte qu'une seule fonction de masse alors que la *fiabilité absolue* reflète le niveau général de fiabilité de toutes les fonctions de masse fournies par une source.

Ces informations concernant les fiabilités relatives et la fiabilité absolue d'une source pourront servir à l'enrichissement de la base de données évidentielle en indiquant le niveau de fiabilité de chaque source.

À partir de s bases de données évidentielles concernant s sources, chacune ayant X attributs évidentiels et Y enregistrements, chaque base de données évidentielle stocke $X \times Y$ différentes fonctions de masse par source. À partir de chaque fonction de masse, l'estimation des fiabilités relatives pourra être faite ; il y aura donc $X \times Y$ degrés de fiabilités par source.

Dans cet article nous proposons de prendre la moyenne des différentes fiabilités relatives comme fiabilité absolue de la source. Le choix de la moyenne se justifie par le fait que la fiabilité d'une source est fixe, bien qu'elle puisse se tromper parfois donc sa fiabilité peut augmenter ou diminuer légèrement mais en moyenne elle garde le même niveau.

En effet, affaiblir avec une fiabilité relative minimale réduit la fonction de masse au maximum ce qui peut mener à l'ignorance totale. De plus, l'affaiblissement avec une fiabilité relative maximale ne permettra pas dans certains cas de réduire le conflit du moment où une source peut être au moins une fois complètement fiable alors qu'elle ne l'est pas en général. Affaiblir avec la fiabilité moyenne permet de réduire le conflit tout en gardant l'intégrité de la fonction de masse. Le choix de la fiabilité moyenne permettra ainsi d'éviter les valeurs aberrantes.

La *fiabilité absolue* α_j^a d'une source S_j est donc la moyenne de ses Y fiabilités relatives α_{yj} :

$$\alpha_j^a = \frac{1}{Y}\sum_{y=1}^{Y}(\alpha_{yj}) \tag{16}$$

4.4 Estimation de la fiabilité de la combinaison

Un degré de fiabilité relative α_j est attribué à chaque fonction de masse, qui pourra servir à estimer la fiabilité de la combinaison.

Si on a s sources fournissant chacune une distribution de masse différente, l'estimation de la fiabilité de chaque source se fera en calculant la distance entre sa distribution de masse et le reste des distributions ce qui représente son conflit relatif. La fiabilité relative est calculée à partir de ce degré de conflit relatif. La fiabilité de la combinaison est la moyenne des s fiabilités relatives propres aux différentes sources concernant les fonctions de masse à combiner pour une observation donnée. La fiabilité d'une combinaison représente le degré de confiance moyen attribué à la fonction de masse résultante.

$$\alpha_c = \frac{1}{s} \times \sum_{j=1}^{s}(\alpha_j) \tag{17}$$

Ces fiabilités associées aux combinaisons pourront enrichir la base de données évidentielle en indiquant à l'utilisateur à quel point il pourra faire confiance à la fonction de masse fournie. Ces fiabilités ne sont utiles que pour l'utilisateur lors de la prise de décision.

Les équations (16) et (17) sont différentes : la première équation permet d'obtenir la fiabilité moyenne d'une source à partir de toutes ses fiabilités relatives correspondant à ses fonctions de masse, tandis que la seconde équation permet d'obtenir la fiabilité de la combinaison à partir des fiabilités relatives de toutes les fonctions de masse des différentes sources combinées.

5 Expérimentation

Afin de pouvoir tester la méthode précédemment décrite, nous avons considéré une base de données radar. Ces données ont été recueillies dans la chambre anéchoïque de l'ENSIETA en plaçant une cible (maquette d'avion) et un capteur radar pouvant détecter la cible sous différents points angulaires. Le système d'acquisition est présenté par Martin et Radoi (2004). Une base de données a été proposée pour l'acquisition et le stockage des signaux par Toumi (2007). Nous considérons ainsi cinq cibles radar différentes (Mirage, F14, Rafale, Tornado, Harrier). Chaque table contient 250 représentations fréquentielles obtenues dans un domaine angulaire d'environ 60° et utilisant une bande de fréquence d'environ 6 GHZ. Pour caractériser les cibles, et donc renseigner la bases de données, nous avons utilisé trois classifieurs différents : le k-plus proche voisin flou, le k-plus proche voisin crédibiliste et les réseaux de neurones. Ces trois classifieurs sont considérés comme des sources, sur lesquelles on déduit des fonctions de masse tel que présenté par Martin et Radoi (2004). Ils ont donc fourni 250 fonctions de masse stockées dans trois tables différentes et permettant de classifier les cinq cibles radar différentes.

Notre but est d'intégrer ces trois tables en combinant les 250 fonctions de masse fournies par chaque source (classifieur) pour obtenir une seule table facilitant les requêtes sur la base et ainsi aider à la prise de décision.

La première étape consiste à estimer les conflits relatifs à chaque source pour chaque fonction de masse, donc chaque source aura 250 conflits relatifs. Le conflit absolu d'une source n'est autre que la moyenne de ses 250 conflits relatifs. Afin de calculer les conflits relatifs, nous avons utilisé deux types de méthode de calcul de distance :

– **Distance type1 :** correspond au calcul du conflit donné par l'équation (13), *i.e.* à la moyenne des distances séparant une fonction de masse fournie par une source et les autres fonctions de masse sans utiliser une règle de combinaison.
– **Distance type2 :** correspond au calcul du conflit donné par l'équation (14), *i.e.* à la distance séparant la fonction de masse fournie par une source et la fonction de masse combinée relative aux autres sources. Il existe plusieurs règles de combinaison pouvant être utilisées pour la combinaison des fonctions de masse telles que rappelées par Smets (2007) et Martin et Osswald (2007), mais dans cet article nous avons utilisé uniquement la moyenne des fonctions de masse et la règle de combinaison de Dempster donnée par l'équation (7). Cette dernière a un comportement conjonctif telle que la règle de Yager (1987) et la règle de Dubois et Prade (1988) mais distribue le conflit de façon différente.

Le conflit absolu initial propre à chaque source et la variance des conflits relatifs sont donnés dans le tableau 1.

Experts	Type de distance	Règle de combinaison	Conflit initial	Variance
K-ppv flou	Type1	-	0.19582	0.01370547
K-ppv flou	Type2	Règle de Dempster	0.1250316	0.02006961
K-ppv flou	Type2	Moyenne	0.147562	0.01127087
K-ppv crédibiliste	Type1	-	0.2232652	0.01772268
K-ppv crédibiliste	Type2	Règle de Dempster	0.0856512	0.02090321
K-ppv crédibiliste	Type2	Moyenne	0.2137072	0.01685384
Réseau de neurones	Type1	-	0.301354	0.03387129
Réseau de neurones	Type2	Règle de Dempster	0.3341664	0.04060376
Réseau de neurones	Type2	Moyenne	0.2922108	0.03484972

TAB. 1 – *Conflits absolus initiaux des sources et les variances associées*

Nous avons utilisé les conflits absolus initiaux pour affaiblir les fonctions de masse des différentes sources avec différentes valeurs de λ (le paramètre de calcul de la fiabilité à partir du conflit de l'équation (11(r)-280.38)9(9(9(9(9(924.240155 [illegible]

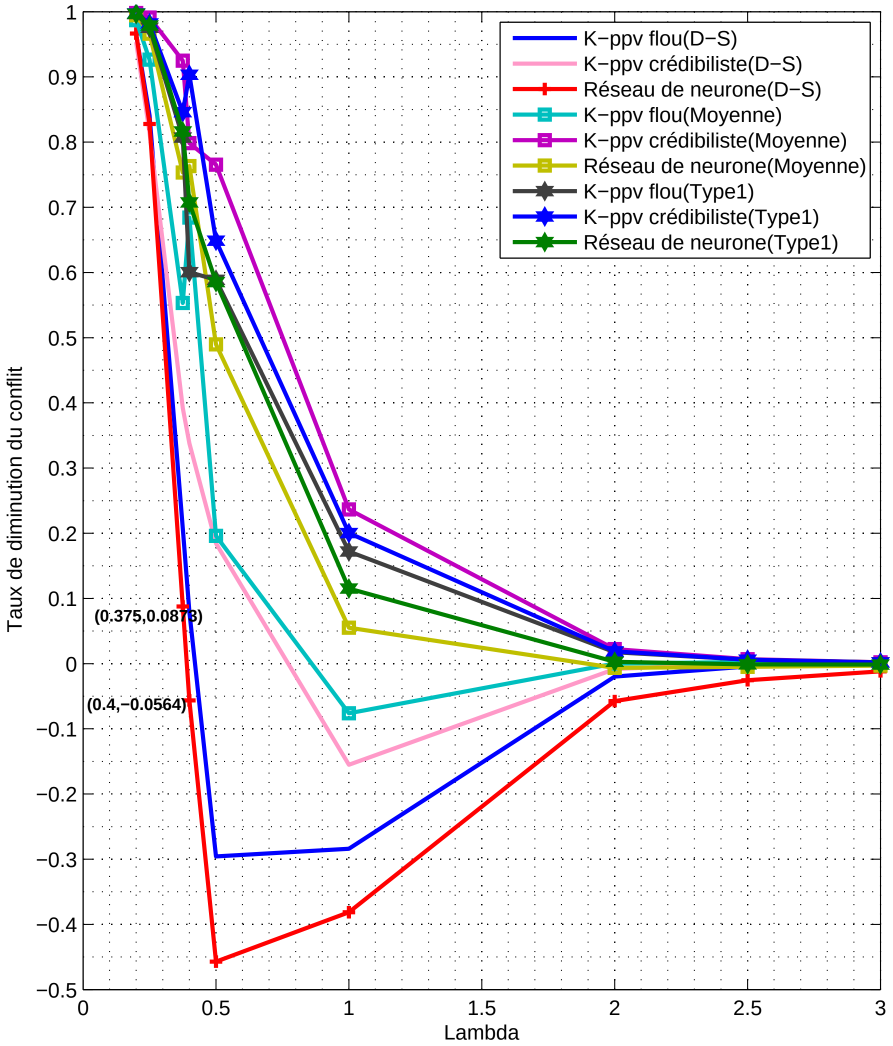

FIG. 1 – *Distributions du conflit avant et après affaiblissement en fonction de* λ

Le choix de λ est conditionné par un pourcentage de diminution positif. D'après la figure 1, pour $\lambda \geq 0.4$ le réseau de neurones a un taux de diminution négatif et pour $\lambda \leq 0.375$ son taux de diminution est positif, c'est pour cela que nous choisissons $\lambda \leq 0.375$ qui garantit un taux de diminution positif pour les deux types de distance.

Pour une valeur $0.375 \leq \lambda_0 \leq 0.4$, le pourcentage de diminution devient nul pour le réseau de neurones avec la distance type 2 en utilisant la règle de combinaison de Dempster et positif pour les autres classifieurs indépendamment de la mesure de distance. Il est difficile de déterminer λ_0 dû à la taille réduite de l'intervalle auquel elle appartient ainsi que l'impossibilité d'estimer la fonction liant le pourcentage de diminution à λ sur ces données réelles.

La fiabilité est une fonction croissante de λ (voir les distributions des fiabilités initiales en fonction de λ dans la figure 2), plus λ est petite plus l'affaiblissement est important d'où le choix de $\lambda = 0.375$ et non pas $\lambda < 0.375$. $\lambda = 0.375$ garantie un pourcentage de diminution positif pour les trois sources indépendamment du type de la distance et de la règle de combinaison tout en affaiblissant le moins possible.

Les conflits absolus propres à chaque source après affaiblissement avec un λ=0.375 sont donnés dans le tableau 2.

Experts	**Type de distance**	**Règle de combinaison**	**Conflit absolu**	**Variance**
K-ppv flou	Type1	-	0.0378216	0.000039045
K-ppv flou	Type2	Règle de Dempster	0.0990748	0.00032168
K-ppv flou	Type2	Moyenne	0.0659344	0.000052011
K-ppv crédibiliste	Type1	-	0.034408	0.000074342
K-ppv crédibiliste	Type2	Règle de Dempster	0.0520692	0.00072428
K-ppv crédibiliste	Type2	Moyenne	0.0160376	0.00020475
Réseau de neurones	Type1	-	0.0553784	0.00010062
Réseau de neurones	Type2	Règle de Dempster	0.304864	0.00044402
Réseau de neurones	Type2	Moyenne	0.072086	0.00012238

TAB. 2 – *Conflits absolus et variances après affaiblissment*

Nous remarquons bien l'amélioration au niveau des conflits après affaiblissement ainsi qu'au niveau des variances des conflits relatifs. Prenons l'exemple de l'utilisation de la distance type 2 avec la moyenne comme règle de combinaison, le conflit initial du réseau de neurones était de 0.29 avec une variance des conflits relatifs de 0.034. Après affaiblissement, le conflit absolu est passé à 0.072 et la variance des conflits relatifs est réduite à un taux de 0.00012.

D'après la figure 2, il est clair que pour $\lambda = 0.375$, les fiabilités après affaiblissement sont supérieures aux fiabilités initiales ce qui signifie une amélioration considérable des fiabilités. Nous remarquons également que les distributions de fiabilités après affaiblissement sont des fonctions croissantes qui tendent vers les distributions des fiabilités initiales. En augmentant la valeur de λ, les fiabilités initiales et les fiabilités après affaiblissement se rejoignent pour atteindre la valeur 1.

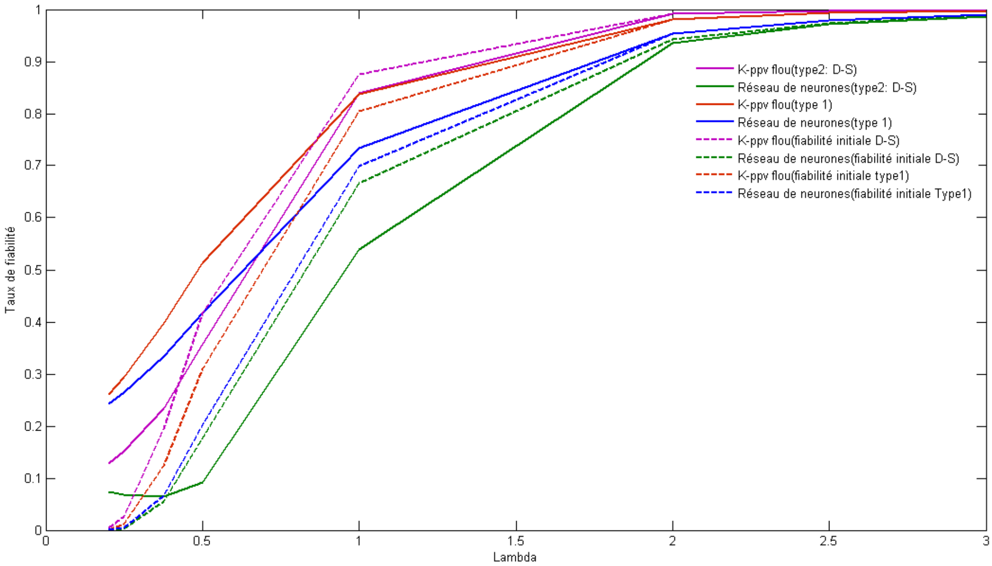

FIG. 2 – *Distributions des fiabilités avant et après affaiblissement en fonction de* λ

6 Conclusion

Dans cet article, nous avons proposé une méthode permettant d'estimer la fiabilité d'une source à partir de toutes les fonctions de masse qu'elle fournit. Cette méthode a pour objectif d'utiliser les fiabilités estimées pour affaiblir les fonctions de masse stockées dans une base de données évidentielle afin de fusionner ces fonctions avec d'autres stockées dans différentes bases de données évidentielles. Cette fusion permettra d'aider l'utilisateur à la prise de décision en réduisant la quantité d'information à traiter et en lui indiquant les degrés de fiabilité des sources et des informations combinées.

Comme perspective à ce travail, la proposition d'une méthode de transfert de masse autre que l'affaiblissement des masses permettrait de modifier l'ensemble des éléments focaux ce qui pourrait améliorer la qualité de classification des fonctions de masse, résultats de la combinaison.

Références

Bach Tobji, M.-A., B. Ben Yaghlane, et K. Mellouli (2008). A new algorithm for mining frequent itemsets from evidential databases. In *Information Processing and Management of Uncertainty*, Malaga, Spain, pp. 1535–1542.

Dempster, A. P. (1967). Upper and Lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics 38*, 325–339.

Dubois, D. et H. Prade (1988). Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence 4*, 244–264.

Hewawasam, K., K. Premaratne, S. Subasingha, et M.-L. Shyu (2005). Rule mining and classification in imperfect databases. In *International Conference on Information Fusion*, Philadelphia, USA, pp. 661–668.

Jousselme, A.-L., D. Grenier, et E. Bossé (2001). A new distance between two bodies of evidence. *Information Fusion 2*, 91–101.

Martin, A., A.-L. Jousselme, et C. Osswald (2008a). Conflict measure for the discounting operation on belief functions. In *International Conference on Information Fusion*, Cologne, Germany.

Martin, A. et C. Osswald (2007). Toward a combination rule to deal with partial conflict and specificity in belief functions theory. In *International Conference on Information Fusion*, Québec, Canada.

Martin, A., C. Osswald, J. Dezert, et F. Smarandache (2008b). General combination rules for qualitative and quantitative beliefs. *Journal of Advances in Information Fusion 3*(2), 67–82.

Martin, A. et E. Radoi (2004). Effective ATR Algorithms Using Information Fusion Models. In *International Conference on Information Fusion*, Stockholm, Sweden.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton University Press.

Smets, P. (2007). Analyzing the combination of conflicting belief functions. *Information Fusion 8*, 387–412.

Smets, P. et R. Kennes (1994). The Transferable Belief Model. *Artificial Intelligent 66*, 191–234.

Toumi, A. (2007). *Intégration des bases de connaissances dans les systèmes d'aide à la décision : Application à l'aide à la reconnaissance de cibles radar non-coopératives*. Ph. D. thesis, Université de Bretagne Occidentale, ENSIETA, Brest.

Yager, R. R. (1987). On the Dempster-Shafer Framework and New Combination Rules. *Informations Sciences 41*, 93–137.

Summary

The conflict appearing while combining several uncertain informations reflects the degree of conflict between their sources. This conflict can be managed before the combination step by discounting belief functions using sources' reliability. In this paper, we propose a generalization method for sources' reliability estimation taking into account all its belief functions stored in an evidential database. This method is evaluated on real radar data and supplied good results in terms of sources' reliability improvement.