

Classification multisource via la théorie des croyances

Anis BEN AÏSSA*, Nour-Eddin EL FAOUZI*, Eric LEFEVRE**

**Laboratoire d'Ingénierie Circulation Transports
INRETS - ENTPE / LICIT*

*25, Avenue François Mitterrand Case 24
69675 Bron Cedex*

Anis.Ben-aissa@inrets.fr nour-eddin.elfaouzi@inrets.fr

***Laboratoire LGI2A*

Université d'Artois

Faculté des Sciences Appliquées

62400 Béthune

eric.lefevre@iut-geii.univ-artois.fr

Résumé. La classification en présence de plusieurs sources de données hétérogènes est abordée comme un problème de fusion de classifieurs. Dans ce travail, la méthodologie de la fusion, opérée par la théorie des croyances, est présentée et plusieurs solutions techniques sont abordées. Les résultats de ces différentes propositions sont illustrés dans le cadre de l'estimation des temps de parcours sur un réseau routier. Ces schémas montrent la propension de la fusion de classifieurs à améliorer la qualité, en terme de taux de bien classés.

1. Introduction

Les méthodes de classification et plus particulièrement les méthodes de discrimination tentent de construire des règles de décision permettant d'affecter un élément au groupe dont il est proche, au sens d'une certaine distance, connaissant uniquement le vecteur de ses caractéristiques. Ces règles de décision sont élaborées de telle sorte qu'elles produisent un minimum d'erreurs lorsqu'elles seront appliquées dans une approche prédictive. Plusieurs approches ont été proposées pour répondre à ces impératifs. A titre d'exemple, on peut citer les méthodes discriminantes (Mardia et al., 1979 ; McLachlan, 1992), méthodes des plus proches voisins (Fix et Hodges, 1951), méthodes neuronales (Ripley, 1996), arbres de classification (Breiman et al., 1984) et plus récemment les méthodes basées sur la théorie de l'apprentissage statistique (Vapnik 2000). Le lecteur intéressé pourra trouver une revue et comparaison de ces méthodes dans Boutou et al. 1994.

Ces différentes propositions reposent sur l'exploitation de l'information sur les éléments à classer sous forme d'un tableau de données. Avec l'amélioration des systèmes de recueil et de collecte d'informations et l'apparition de nouvelles sources de données, les informations sont le plus souvent multiformes et multisources interdisant l'application directe de telles propositions. On est alors face à un problème de classification dans une configuration multisource. L'exploitation des spécificités de chacune des sources et la prise en compte de leurs imperfections sont à l'origine des difficultés opérationnelles et nécessitent le développement d'un cadre méthodologique spécifique.

Classification de données multisources

Un exemple type de ce problème de classification est celui de l'estimation des temps de parcours sur un axe urbain. Traditionnellement, la notion du temps de parcours est utilisée pour répondre aux préoccupations primaires de la gestion du trafic, qui consistent à apporter la réponse la plus satisfaisante aux besoins de déplacements. Le temps de parcours est alors utilisé comme indicateur permettant de qualifier la *qualité de service* du réseau de transport concerné. Le temps de parcours utilisé dans ce contexte est principalement de type instantané.

La réponse à ce besoin primaire de gestion de flux de trafic fait apparaître, sous la demande forte des usagers, un besoin secondaire qui concerne l'information routière. Dans ce cas, le temps de parcours constitue une information, associée à la notion d'impédance (ou coûts) liée aux itinéraires du réseau ou/et à un indicateur de congestion.

Face à ces diverses utilisations, la question d'une estimation avec une précision acceptable du temps de parcours se pose. Ce problème est particulièrement ardu en milieu urbain, où l'on doit faire face à un certain nombre de difficultés théoriques, techniques et méthodologiques. Ainsi, pour connaître l'état du trafic sur un axe urbain, les capteurs classiques utilisés pour mesurer les conditions de circulation s'avèrent inefficaces dans certaines circonstances. Avec l'apparition de nouveaux moyens de mesure (caméras, moyens de localisation de type GPS ou téléphones cellulaires ...), on recourt donc de plus en plus à d'autres sources de données visant à compléter l'information fournie par les moyens de mesure classiques et, en conséquence, à améliorer la qualité de l'estimation du temps de parcours. Le problème de l'estimation du temps de parcours devient alors un problème typique de fusion de données.

La collecte des données par les véhicules traceurs en complément des données de capteurs au sol, permettrait de fournir une meilleure image (globale et complète) de l'état du trafic et, par la même, de faire face aux imperfections des mesures (données manquantes, aberrantes...). Cette collecte, de plus en plus accessible, notamment avec le développement récent de l'usage du téléphone cellulaire et plus généralement celui des réseaux UMTS (*Universal Mobile Telecommunications System*), laisse entrevoir la mise en place de systèmes de gestion de trafic économiquement bon marché, utilisant conjointement des données collectées par des véhicules tests et des données issues de capteurs au sol traditionnels.

Les propriétés de complémentarité et de redondance de ces deux sources de données peuvent donc être mises à profit afin d'élaborer une solution multisources pour le problème d'estimation du temps de parcours en milieu urbain. De cette approche multisources sera exigée d'exploiter au mieux les avantages de chacune des sources d'information, tout en essayant de pallier leurs limitations individuelles respectives.

Le présent article portera sur le problème de classification multisource pour l'estimation du temps de parcours. Il s'agit essentiellement de la fusion des classificateurs basés sur chacune des sources en utilisant la théorie des croyances.

2. Fusion de classifieurs

Dans cette section, après une présentation succincte du problème générique que l'on cherche à résoudre et de la théorie des croyances, nous présenterons trois schémas de fusion de classifieurs basés sur cette théorie et l'apprentissage statistique.

2.1. Problème générique

Le problème générique de la classification est la détermination de la classe d'appartenance d'un élément sur la base de la connaissance seule de ses caractéristiques donnée par un vecteur $x \in \mathbb{R}^n$, $x = \{x_1, \dots, x_n\}$. Ce vecteur est assimilé à un échantillon issu d'une population \mathcal{D} constituée de r classes c_1, \dots, c_r . On appelle donc classifieur tout outil de reconnaissance qui reçoit un vecteur de mesures x en entrée et donne des attributs de la classe de cette mesure x . Plus précisément, soient un vecteur $x \in \mathbb{R}^n$ et un ensemble de classes $\Lambda = \{c_1, \dots, c_r\}$ un classifieur \mathcal{D} est fonction définie par :

$$\begin{aligned} \delta : \mathbb{R}^n &\rightarrow [0,1]^r \\ x &\mapsto \delta(x) = (\delta_1(x), \dots, \delta_r(x)) \end{aligned}$$

Les composantes δ_h ; $h = 1, \dots, r$ peuvent être vu comme des estimations des probabilités *a posteriori* des classes, pour x donné. Autrement dit, $\delta_h(x) = P(y = h | x)$ avec y désigne le label des classes et δ_h à valeurs dans Λ .

La décision de δ peut être durcie de sorte qu'un indice net de classes dans Λ soit assignée à x . Ceci est typiquement fait par la règle du maximum d'adhésion (règle bayésienne) :

$$\begin{cases} \delta_k(x) = 1 & \text{si } k = \arg \min_{i=1, \dots, r} \{\delta_i(x)\} \\ \delta_h(x) = 0 & \text{si } h \neq k \end{cases}$$

On notera par la suite $\delta(x) = k$ lorsque $\delta_k(x) = 1$.

Dans le cas où plusieurs classifieurs sont disponibles, l'agrégation de ces classifieurs permet le plus souvent d'améliorer la qualité de la classification en terme de taux de bien classés. Cette propension à l'amélioration est d'autant plus importante que les classifieurs soient complémentaires. Le plus souvent les différents classifieurs reposent soit sur des différentes méthodologies (Kuncheva et al., 2001), soit sont issus de sources différentes (El Faouzi, 2000a).

Classification de données multisources

Les schémas d'agrégation, quant à eux, vont de la simple moyenne arithmétique (Alexandre et al., 2000) jusqu'aux approches basées sur les théories de l'incertain (approche bayésienne et réseaux du même nom) et de l'imprécis (approches crédibiliste et possibiliste), en passant par des schémas fondés sur les réseaux de neurones, l'approche des k plus proches voisins (Xu et al., 1992) et des techniques de rééchantillonnages (Breiman, 1998 ; El Faouzi, 1997).

Le travail rapporté dans le présent article s'inscrit dans le cadre de la fusion de classifieurs issus de sources différentes et l'approche d'agrégation utilisée repose sur la théorie des crédibilités.

2.2. Théorie des croyances

Cette théorie, basée sur les travaux de Dempster (1967, 1968), a été formalisée par Shafer (1976). Cette théorie, dont le principe est de pouvoir associer à un événement une probabilité inférieure et une probabilité supérieure, constitue une généralisation de l'approche bayésienne.

L'idée de base de cette théorie est qu'à partir de l'ensemble des classes Λ , appelé cadre de discernement, on définit sur l'ensemble des parties de Λ une fonction de masse m ayant ses valeurs dans l'intervalle $[0,1]$ vérifiant $m(\emptyset) = 0$ et $\sum_{A \subseteq \Lambda} m(A) = 1$.

La quantité $m(A)$ s'interprète comme la part de croyance placée strictement sur A . Cette quantité se différencie d'une probabilité par le fait que la totalité de la croyance est répartie non seulement sur les classes singletons mais aussi sur les classes composites. La plupart du temps la construction de cette fonction dépend de l'application envisagée (Dromigny-Badin, 1998). Toutefois, quelques approches ont vu le jour afin de généraliser l'affectation des croyances (Denoeux, 2000 ; Appriou 1991).

Une fois le jeu de masse défini pour chacune des sources en présence, la décision finale est élaborée d'une part en combinant les masses par une règle propre à cette théorie, appelée règle de combinaison orthogonale de Dempster et d'autre part en adoptant une des multiples règles de décision qu'offre la théorie des croyances.

Le principe de la règle de combinaison est relativement simple. Pour le besoin de l'exposé, considérons deux sources S_1 et S_2 auxquelles sont associées deux fonctions de masses m_1 et m_2 . Si u et v désignent les sorties des deux classifieurs δ^1 et δ^2 issus des sources S_1 et S_2 respectivement. La masse associée à leur jonction $w = u \cap v$ est proportionnelle à $m_1(u) \times m_2(v)$. Comme plusieurs paires d'ensembles u et v peuvent avoir w pour intersection, la masse associée à w doit alors intégrer les produits des masses de ces ensembles. Cependant, si u et v sont disjoints, la masse associée à leur intersection

est portée sur l'ensemble vide. Ceci violerait l'hypothèse imposée d'une masse nulle de l'ensemble vide. Cette difficulté peut être contournée par une simple normalisation.

$$m_1 \oplus m_2(w) = \begin{cases} 0 & \text{si } w = \emptyset \\ \frac{1}{1 - \lambda_c} \sum_{u \cap v = w} m_1(u) \times m_2(v) & \text{si } w \neq \emptyset \end{cases}$$

où $\lambda_c = \sum_{u \cap v = \emptyset} m_1(u) \times m_2(v)$ représente le conflit entre les sources, lequel est pris en compte dans la combinaison sous forme d'un facteur de normalisation. Cette règle de combinaison a été critiquée dans plusieurs travaux dont (Yager, 1987). Ainsi, d'autres opérateurs alternatifs à cette somme orthogonale ont été étudiés dans la littérature : opérateurs non normalisés (Smets, 1990), opérateurs disjonctifs (Dubois et al. 1988), généralisation d'opérateurs (Lefevre, 2001),...

La théorie des croyances propose plusieurs critères de décisions. Nous reprenons certaines de ces règles ici. Pour plus d'informations le lecteur pourra se reporter à (Denoeux, 1997).

1. **Règle 1** : La décision combinée est la classe ayant la valeur de crédibilité maximale :

$$\theta_k = \arg \min_{\theta_i, i=1, \dots, r} \left[Cr(\theta_i) = \sum_{c_j \subseteq \{\theta_i\}} m(c_j) \right]$$

La crédibilité Cr constitue le degré de croyance minimal que l'on attribue à une classe.

2. **Règle 2** : La décision combinée est la classe ayant la valeur de plausibilité maximale :

$$\begin{aligned} \theta_k &= \arg \min_{\theta_i, i=1, \dots, r} \left[Pl(\theta_i) = \sum_{c_j \cap \{\theta_i\} \neq \emptyset} m(c_j) \right] \\ &= \arg \min_{\theta_i, i=1, \dots, r} \sum_{\{\theta_i\} \subseteq c_j} m(c_j) \end{aligned}$$

La plausibilité Pl constitue le degré de croyance maximal que l'on attribue à une classe.

Classification de données multisources

3. **Règle 3** : La décision combinée est la classe ayant la valeur de probabilité pignistique maximale :

$$\theta_k = \arg \min_{\theta_i, i=1, \dots, r} \left[P_g(\theta_i) = \sum_{c_j \in \{\theta_i\}} \frac{m(c_j)}{|\mathbf{c}_j|} \right]$$

Ce critère consiste à choisir la classe simple la plus probable en équirépartissant la masse placée sur chaque classe composée sur les classes simples qui la composent.

C'est cette dernière règle qui sera utilisée par la suite car elle garantit que la classe retenue après fusion soit une classe singletons, i.e. élément de Λ .

2.3. Schémas de fusion de classifieurs

Les schémas de fusion de classifieurs que l'on va introduire reposent essentiellement sur la prise en compte des erreurs des classifieurs individuels. Les erreurs de chaque classifieur δ^h sont usuellement consignés dans la matrice de confusion donnée par :

$$M_h = \begin{pmatrix} n_{11}^{(h)} & \dots & n_{1n}^{(h)} \\ \vdots & \ddots & \vdots \\ n_{r1}^{(h)} & \dots & n_{rr}^{(h)} \end{pmatrix} \quad h = 1, \dots, \ell$$

La ligne i correspond à la classe c_i de l'apprentissage et la colonne j correspond à la classe reconstituée par le classifieur δ^h , i.e. $\delta(x) = j$. Cette matrice est obtenue par apprentissage sur un échantillon test et peut être considérée comme la connaissance *a priori* sur les performances du classifieur δ^h . Les éléments diagonaux sont les pourcentages des concordances entre les classes reconstituées par le classifieur et les classes de référence. En d'autres termes, ceci représente le nombre de fois où la classe de référence et la classe reconstituée par le classifieur coïncident. Les éléments hors diagonale donnent quant à eux les pourcentages de discordances (i.e. de confusion).

A partir de cette matrice, on définit le taux de reconnaissance (τ_{rec}^h) et le taux de confusion (τ_{conf}^h) par :

$$\tau_{rec}^h = \frac{\sum_{i=1, \dots, r} n_{ii}^{(h)}}{\sum_{i, j=1, \dots, r} n_{ij}^{(h)}} \quad \text{et} \quad \tau_{conf}^h = 1 - \tau_{rec}^h$$

Nous supposons que l'on dispose d'une matrice de confusion par classifieur et nous présentons dans les sections suivantes trois approches utilisées pour l'élaboration des masses à partir de ces matrices de confusion.

La première approche, proposée initialement par Xu et al. (Xu et al., 1992) constitue à la fois une méthode de référence en fusion de classifieurs et le point de départ de ce travail. Nous avons cherché à en améliorer les performances par une exploitation optimale de la matrice de confusion.

2.3.1 Premier schéma :

Dans ce premier schéma, le taux de reconnaissance et de confusion sont utilisés pour définir les masses de croyances. On définit ce jeu de masse de la façon suivante :

$$(SF.1) \quad \begin{cases} m_k(c_i) = \tau_{rec}^h \\ m_k(\bar{c}_i) = \tau_{conf}^h \end{cases} \quad \forall i = 1, \dots, r$$

Pour cette première approche, le support des masses de croyance est formé des classes et de leurs complémentaires. Par ailleurs, ce schéma impose qu'un même jeu de masses soit associé à une classe abstraction faite de la classe retenue par le classifieur, i.e. $\delta^h(x) = s_h$, $s_h \in \Lambda$.

2.3.2 Deuxième schéma

Cette approche constitue une première amélioration de l'approche précédente et consiste à mieux exploiter les informations contenues dans la matrice de confusion en conditionnant la définition des masses par la sortie de chaque classifieur. En d'autres termes, si

Classification de données multisources

$\mathcal{D}^h(x) = j$, on définit un taux de reconnaissance sachant que le classifieur \mathcal{D}^h a retenu la classe j par :

$$\tau_{rec|j}^h = \frac{n_{jj}^{(h)}}{\sum_{i,j=1,\dots,r} n_{ij}^{(h)}} \quad \text{et} \quad \tau_{conf|j}^h = 1 - \tau_{rec|j}^h$$

et le jeu de masse est obtenu en substituant au taux de reconnaissance et au taux de confusion leurs versions conditionnelles :

$$(SF.2) \quad \begin{cases} m_k(c_i) = \tau_{rec|i}^h \\ m_k(\bar{c}_i) = \tau_{conf|i}^h \end{cases} \quad \forall i = 1, \dots, r$$

2.3.3 Troisième schéma :

Cette troisième approche consiste à un conditionnement dans lequel les éléments de la colonne $\mathcal{D}^h(x) = j$ vont être utilisés pour élaborer un jeu de masse associé aux r classes contenues dans Λ .

$$(SF.3) \quad \begin{cases} m_k(c_j) = \frac{n_{ij}^{(k)}}{\sum_{i=1,\dots,r} n_{ij}^{(k)}} \\ m_k(\Lambda) = \sum_{i=1}^r n_{ij}^{(k)} - n_{jj}^{(k)} \end{cases} \quad \forall i, j = 1, \dots, r$$

Cette fonction de croyance n'étant pas normalisée, un coefficient de normalisation est appliqué afin de respecter les propriétés des jeux de masses. Nous obtenons alors la fonction de croyance suivante :

$$m_k(A) = \frac{m_k(A)}{\sum_{B \subseteq \Lambda} m_k(B)} \quad \forall A \subseteq \Lambda$$

Les deux derniers schémas constituent deux propositions originales développés dans le cadre de ce travail que nous avons cherché à comparer avec la méthode de Xu et al., 1992.

Une fois les jeux de masses associés à chaque classifieurs élaborés, la fusion est opérée en faisant appel à la combinaison orthogonale de Dempster. La décision finale est prise en utilisant la règle du maximum de probabilité pignistique.

3. Mise en oeuvre

3.1. Données

Les données utilisées dans cette étude ont été recueillies à l'issue de la campagne de mesures réalisée sur un axe urbain de Toulouse. La campagne de mesure s'est déroulée pendant huit jours, du 16 au 25 novembre 1993, à raison de trois heures par jour, réparties en deux périodes : 14h30 – 16h et 16h30 – 18h. Le choix de ces périodes est dicté par le besoin de rendre compte de conditions de trafic variées. L'axe urbain sur lequel a été réalisée cette expérimentation mesure deux kilomètres environ et comporte 4 tronçons par sens de circulation. Seul l'itinéraire correspondant au sens sud-nord a été retenu dans le cadre de cette étude.

L'enquête de mesure a consisté en un recueil de données issues de plusieurs sources d'information. Nous en avons considéré trois qui ont été disponibles simultanément :

- **Source « capteurs »** : cette source est composée d'un réseau de capteurs à boucles électromagnétiques (une douzaine), qui délivrent des données de trafic, principalement le débit et le taux d'occupation, toute les 12,5ms , soit environ de 115200 observations .
- **Source « traceurs »** : Quatre véhicules disposant de capteurs embarqués (dits traceurs), qui fournissent des données cinématiques et donc leurs temps de parcours sur chaque tronçon. Ces véhicules avaient comme consigne de parcourir l'axe étudié pendant les deux périodes temporelles suscitées (74 800 réalisations).
- Un recueil exhaustif des temps de parcours sur cet axe vient compléter les sources d'information dont on dispose. Ce recueil est réalisé par la reconnaissance des plaques minéralogiques des véhicules (110 400 réalisations)

Les deux premières sources servent à bâtir deux classifieurs que nous souhaitons fusionner, alors que la troisième fournit le temps de parcours de référence que l'on cherche à reconstituer.

Une particularité du problème d'estimation de temps de parcours est son double dimensionnement à la fois spatial et temporel. Une première étape consiste à recalculer temporellement et spatialement les différentes données disponibles à savoir les temps de parcours.

Dans le cadre de cette étude, on dispose d'un recueil d'informations sur le temps de parcours sur un axe routier soit par mesure directe (cas des véhicules traceurs) soit par estimation (cas des capteurs à boucles). Une première tâche a consisté à une correspondance temporelle des différentes données en les agrégeant sur une période temporelle de six minutes.

Classification de données multisources

A l'issue de cette étape, on obtient 240 observations pour les capteurs, 156 observations pour les traceurs et 230 pour les temps de parcours de référence.

La conversion des données trafic en temps de parcours est rendue nécessaire par l'un des principes de fusion qui veut que, pour pouvoir être combinées et comparées, les différentes informations doivent être manipulées dans un espace de représentation commun. Ainsi, la conversion des données issues de capteurs à boucles électromagnétiques est réalisée par la méthode proposée par Bonvalet et Robin-Prévallée, dite BRP (Bonvalet et Robin-Prévallée, 1987).

Cette méthode se fonde sur l'existence d'une relation linéaire entre le temps de parcours et le taux d'occupation sur chaque tronçon qui compose l'axe à étudier. Le temps de parcours associé à l'itinéraire est alors défini par la somme des temps de parcours sur les tronçons.

Nous disposons ainsi des temps de parcours élaborés soit à partir de données trafic soit à partir de données traceurs. A l'issue de cette phase de recalage, les temps de parcours résultants sont de dimensions spatiales et temporelles équivalentes.

Afin de transformer le problème de l'estimation en un problème de classification, un découpage préalable des temps de parcours en classes a été effectué. Dans cette application, nous avons choisi six classes de temps de parcours sur l'axe routier considéré. Elles sont définies à partir des temps de parcours de référence (enquête minéralogique).

Notons la complémentarité entre la source « capteurs » et la source « traceurs ». En effet, les données fournies par des capteurs au sol sont des mesures quasi exhaustives, c'est-à-dire couvrant l'ensemble des véhicules ayant empruntés le tronçon, avec un échantillonnage et une résolution temporelle excellente. Cependant, ces mesures sont très imprécises (principalement le taux d'occupation qui est très bruité...) avec un échantillonnage spatial qui dépend de la densité des capteurs. En effet, ces mesures ne représentent l'état du trafic qu'à l'endroit où le capteur est placé et non sur l'ensemble du tronçon.

À l'inverse, les données fournies par des véhicules munis de capteurs embarqués (véhicules tests), sont quant à elles, des mesures très précises et de qualité inégale avec une excellente couverture spatiale. Elles expriment l'état du trafic sur l'ensemble du tronçon. Cependant, elles sont non exhaustives, car ne couvrant qu'une partie des véhicules ayant emprunté le réseau pendant une période temporelle donnée.

3.2. Evaluation des résultats

Pour quantifier l'apport de la fusion, nous avons calculé trois taux de bien classés (exprimés en pourcentage) correspondant d'une part aux deux sources d'informations prises individuellement : informations issues du réseau de capteurs et par les véhicules traceurs et d'autre part au taux bien classés obtenu après fusion.

Ces différents taux ont été calculés à partir d'un échantillon de validation (échantillon n'ayant pas fait partie des données d'apprentissage) obtenu par tirage aléatoire.

Classifieurs individuels et fusionné	Schémas de fusion via la théorie des croyances Règle de max de probabilité pignistique		
	SF.1	SF.2	SF.3
Capteurs	33,91 %		
Traceurs	30,07 %		
Fusion	33,91 %	33,91 %	42,17 %

TABLE I. Performances des différents classifieurs et de leurs fusions.

On voit alors que l'application des trois schémas de fusion précédents au problème de l'estimation de la classe des temps de parcours montre des résultats variables d'une stratégie à l'autre.

On remarque que la fusion, basée sur les deux premiers schémas, ne fait guère mieux que le meilleur classifieur. La raison essentielle est que les performances de chacun des classifieurs ne sont pas correctement modélisées. Bien que le conditionnement introduit dans le second schéma tend à mieux appréhender les performances de ces classifieurs, il n'est cependant pas suffisant pour apporter une amélioration notable dans le processus de fusion. Le premier schéma qui est considéré comme une méthode de référence pour la fusion de classifieurs s'est donc révélée inefficace dans le cadre de notre application.

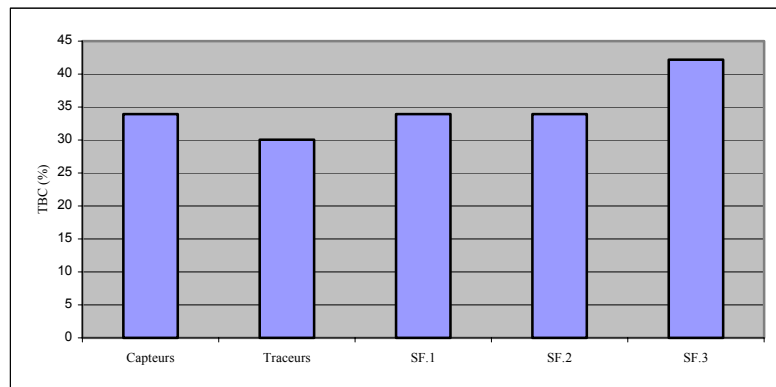


FIGURE I. Performances de la fusion en terme de taux de bien classés.

Le troisième schéma qui intègre toute la connaissance contenue dans la matrice de confusion tend à détecter correctement les vraies classes de temps de parcours. La fusion dans ce cas surclasse le meilleur classifieur avec une amélioration de l'ordre de 24,4 %.

4. Conclusion

L'objectif du travail effectué était de proposer un cadre méthodologique ainsi que des solutions au problème d'estimation de temps de parcours en présence de données issues de sources hétérogènes. Deux sources ont été considérées ici : des capteurs classiques de trafic constitués d'une boucle électromagnétique, qui permettent de mesurer le débit et le taux d'occupation et d'en déduire une estimation de temps de parcours moyen, et un échantillon réduit de véhicules traceurs, qui recueillent les temps de parcours qu'ils ont réalisé.

Abordant le problème de l'estimation du temps de parcours comme un problème typique de fusion de classifieurs et utilisant la théorie des croyances pour sa finesse de modélisation de la connaissance, nous avons utilisé trois méthodes différentes de fusion de classifieurs : la méthode de Xu (qui a constituée le point de départ de notre travail) et deux autres approches permettant de tirer meilleur profit de la matrice de confusion. Parmi ces méthodes, la troisième approche s'est avérée la plus efficace d'entre elles. Il est évident que l'on ne peut pas généraliser ces résultats car ils ont été obtenus dans le cadre d'une application spécifique. Une étude sur d'autres bases de données est envisagée.

Références

- [Alexandre et al., 2000] Alexandre L, Campilho A., Kamel M. – «Combining independent and unbiased classifiers using weighted average». In : Proc. 15th Conf. On Pattern Recognition, vol 2. IEEE Press, 2000.
- [Appriou, 1991] Appriou A. – «Probabilités et incertitude en fusion de données multi-senseurs», Revue Scientifique et Technique de la Défense, 11, pp 27-40, 1991.
- [Ben Aïssa, 2003] Ben Aïssa A – «Fusion de classifieurs pour l'estimation des temps de parcours : Théorie de l'évidence», Rapport de stage LICIT, N°0303, 2003, sous la direction de N.E. El Faouzi.
- [Bonvalet, 1987] Bonvalet F., Robin-Prévallée Y. – Mise au point d'un indicateur permanent des conditions de circulation en Ile-de-France., Revue *T.E.C.*, N° 84-85, Septembre/Décembre, 1987.
- [Breiman et al., 1984] Breiman L., Friedman J.-H., Olshen R., and Stone C. J., – «*Classification and regression trees*», The Wadsworth statistics/probability series, 1984.
- [Breiman, 1998] Breiman L. – « Arcing classifiers», *Annals of Statistics*, 26, pp. 801-824, 1998.
- [Dempster, 1967] Dempster A. P. – Upper and Lower Probabilities Induced by Multivalued Mapping. *Annals of Mathematical Statistics*, Vol. 38, pp. 325-339, 1967.

- [Dempster, 1968] Dempster A.P. – «A Generalization of Bayesian Inference», *Journal of the Royal Statistical Society*, 30B, pp. 205-247, 1968.
- [Denoeux, 1997] Denoeux T. – «Analysis of evidence-theoretic decision rules for pattern classification», *Pattern Recognition*, 30(7), pp. 1095-1107, 1997.
- [Denoeux, 2000] Denoeux T. – «A neural network classifier based on Dempster-Shafer theory», *IEEE Transactions on Systems, Man and Cybernetics, Part A*, 30(2), 2000
- [Dromigny-Badin, 1998] Drogmigny-Badin A. – «*Fusion d'image par la théorie de l'évidence en vue d'applications médicales et industrielles*», Thèse de doctorat : INSA Lyon, 1998.
- [Dubois et Prade *et al.*, 1988] Dubois D. et Prade H. – «Representation and combination of uncertainty with belief functions and possibility measures», *Computer Intelligent*, 4, pp. 224-264, 1988.
- [El Faouzi, 1997] El Faouzi N.E. – «Fusion linéaire d'estimateurs multiples : Méthodes de pondération et de rééchantillonnage. », *Rapport LICIT n°9701*, 1997.
- [El Faouzi, 2000a] El Faouzi N.-E.,– «Fusion de données pour l'estimation des temps de parcours via la théorie de l'évidence», *Recherche Transport Sécurité n° 68*, 2000.
- [El Faouzi, 2000b] El Faouzi N.-E. – «Fusion de données : Concepts et méthodes», *Rapport de Synthèse n°2002*, INRETS-LICIT, 2000.
- [Fix and Hodges, 1951] Fix E., Hodges J. – «Discriminatory analysis, non parametric discrimination : consistency properties», *Technical report, Randolph Field, Texas : USAF School of aviation Medicine*, 1951.
- [Kuncheva *et al.*, 2001] Kunchava L. I., Bezdek C. J., Duin R.P.W. – «Decision templates for multiple classifier fusion : On experimental comparison», *Pattern Recognition*, 34, pp.299-314, 2001.
- [Lefèvre, 2001] Lefèvre E. – «*Fusion adaptée d'informations conflictuelles dans le cadre de la théorie de l'évidence*», Thèse de doctorat, INSA Rouen, 2001.
- [Lu, 1996] Lu Y. – «Knowledge integration in a multiple classifier system», *Appl. Intell.*, 6, pp.75-86, 1996.
- [Mac Lachlan, 1992] Mac Lachlan G. J. – «*Discriminant analysis and statistical pattern recognition*», Wiley, New York , 1992
- [Mardia, 1979] Mardia K. V., J. T. Kent, J. M. Bibby – «*Multivariate Analysis.* », Academic Press Inc., San Diego, 1979.

Classification de données multisources

- [Ripley, 1996] Ripley B. D. – « *Pattern recognition and neural networks*», Cambridge University Press, Cambridge, New York, 1996.
- [Rogova, 1994] Rogova G. – «Combining the results of several neural networks classifiers» , *Neural Networks*, 7(5), 1994.
- [Shafer 1976] Shafer G. – «*A Mathematical Theory of Evidence* », *Princeton University Press*, Princeton, New Jersey, Etats-Unis, 1976.
- [Smets, 1990] Smets P. – «The combination of evidence in the transferable belief model», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5), pp. 447-458, 1990.
- [Xu et al., 1992] Xu L., Krzyzak A., Suen C. Y. – «Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition», *IEEE Transactions on Systems, Man and Cybernetics*, 22(3), 1992.
- [Yager, 1987] Yager R.R. - «On the Dempster-Shafer framework and new combination rules. », *Information Sciences*, 41, pp. 93-138, 1987.
- [Vapnik, 2000] Vapnik V. – « *The nature of statistical learning theory* », Springer, New York, 2nd Edition, 2000.