

Estimation et fusion des Temps de parcours routiers par la théorie de l'évidence

Eric Lefevre*, Nour-Eddin El Faouzi**

*Laboratoire d'Informatique et d'Automatique de l'Artois (LGI2A), EA 3926
Université d'Artois, Faculté des Sciences Appliquées
Technoparc Futura, 62400 Béthune
eric.lefevre@iut-geii.univ-artois.fr,

**Laboratoire d'Ingénierie Circulation Transports (LICIT)
INRETS-ENTPE
25, Avenue François Mitterrand Case 24
69675 Bron Cedex
nour-eddin.elfaouzi@inrets.fr

Résumé. Nous abordons ici le problème de l'estimation du temps de parcours sur un axe urbain par des méthodes de classification basées sur la théorie de l'évidence. Les informations issues des différentes sources de recueil de temps de parcours (capteurs au sol, boucle magnétique, vidéo, véhicules traceurs,...) sont complémentaires et redondantes. Il est alors nécessaire de mettre en œuvre des stratégies de fusion multi-capteurs. L'approche de fusion que nous avons retenue, est basée sur la théorie des fonctions de croyance. Cette théorie permet de prendre en compte de manière plus naturelle les imprécisions ainsi que les incertitudes liées aux différentes informations traitées. Deux stratégies sont mises en œuvre. Tout d'abord une approche fusion de classifieurs dans laquelle les sources d'information sont considérées comme des classifieurs. La seconde approche est une approche de classification basée sur une approche distance pour la modélisation des fonctions de croyance. Des résultats de ces approches sur des données recueillies à l'issue d'une campagne de mesures réalisée sur un axe urbain de Toulouse montre les avantages de la fusion dans le cadre de cette application.

1 Introduction

Dans cet article nous abordons le problème de l'estimation du temps de parcours sur un axe urbain par des méthodes de classification basées sur la théorie de l'évidence. La notion de temps de parcours est utilisée pour répondre aux préoccupations de la gestion du trafic, qui consistent à apporter la réponse la plus satisfaisante aux besoins de déplacements, mais aussi comme information routière pour l'utilisateur.

Les techniques de recueil de l'information temps de parcours peuvent être classées en deux familles. La première famille repose sur des moyens liés à l'infrastructure routière. Ces moyens sont généralement des détecteurs (capteurs au sol, boucle magnétique, vidéo, ...) et/ou des

observateurs (enquête minéralogique relevée manuellement, ...). La seconde famille concerne les moyens de mesures embarqués à bord de véhicules. On parle alors de *véhicules tests*. Il existe une complémentarité entre les données issues de capteurs au sols et celles mesurées par des capteurs embarqués à bord de véhicules empruntant les différents tronçons d'un réseau.

Les données fournies par les capteurs au sols 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 fortement bruité,...) avec un échantillonnage spatial qui dépend de la densité et de l'emplacement 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.

A l'inverse, les données fournies par des véhicules munis de capteurs embarqués (véhicules tests), sont, quant à elles, des mesures très précises 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.

Les propriétés de complémentarité et de redondance de ces deux sources de données peuvent donc être mises à profit en élaborant une solution fusion de données multi-capteurs pour le problème d'estimation du temps de parcours en milieu urbain. L'objectif de cette approche multi-capteurs sera d'exploiter au mieux les avantages de chacune des sources d'information, tout en essayant de pallier leurs limitations individuelles respectives et de faire face, par la même, aux imperfections des mesures (données manquantes, aberrantes. ...). Ceci afin de fournir une meilleure image (globale et complète) de l'état du trafic.

Parmi les techniques de fusion de données multi-capteurs connues dans la littérature, nous avons retenu les approches basées sur la théorie des fonctions de croyance. En effet, il apparaît que cette théorie permet de prendre en compte de manière plus naturelle les imprécisions ainsi que les incertitudes liées aux informations. Dans cet article, nous présentons deux approches de fusion utilisant cette théorie. Tout d'abord une approche fusion de classifieurs dans laquelle les sources d'information sont considérées comme des classifieurs. La seconde approche est une approche de classification basée sur une approche distance pour la modélisation des fonctions de croyance.

Cet article est décomposé de la manière suivante. Nous présentons dans la section 2, les bases mathématiques de la théorie des fonctions de croyance nécessaires à la compréhension de la suite de l'article. Les deux méthodes de classification employées dans l'estimation du temps de parcours sont présentées dans la section 3. Enfin, dans la section 4, des résultats de ces approches sur des données recueillies à l'issue d'une campagne de mesures réalisée sur un axe urbain de Toulouse montre les avantages de la fusion dans le cadre de cette application.

2 Théorie des Fonctions de Croyances

Dans cette section, nous rappelons brièvement quelques concepts de base de la théorie des fonctions de croyance. Le point de vue du modèle des croyances transférables, proposé par Smets et Kennes (1994), est adopté dans cet article. Celui-ci distingue deux niveaux de traitement de l'information : le niveau crédal où les croyances sont modélisées et révisées et le niveau pignistique dans lequel les fonctions de croyance sont transformées en fonctions de probabilités pour la prise de décision.

2.1 Niveau crédal

Soit $\Omega = \{\omega_1, \dots, \omega_k, \dots, \omega_K\}$ un ensemble fini, généralement appelé cadre de discernement. Une fonction de croyance bel est une mesure floue non additive de 2^Ω dans $[0, 1]$ définie par :

$$bel(A) \triangleq \sum_{\emptyset \neq B \subseteq A} m(B) \quad \forall A \subseteq \Omega \quad (1)$$

où m , appelé généralement jeu de masses, est une fonction de 2^Ω dans $[0, 1]$ qui vérifie :

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

Chaque sous-ensemble $A \subseteq \Omega$ tel que $m(A) > 0$ est appelé élément focal de m . Ainsi, la masse $m(A)$ représente le degré de croyance attribué à la proposition A et qui n'a pas pu, compte tenu de l'état de la connaissance, être affectée à un sous-ensemble plus spécifique que A . Tirés des travaux de Dempster (1968) et de Shafer (1976), les fonctions de croyance sont de nos jours reconnues pour la modélisation des informations incertaines (de l'ignorance totale à la connaissance complète) et imprécises. Nous verrons plus en détail, dans la section 3, les méthodes employées dans notre application afin de construire les fonctions de croyances.

La seconde étape au niveau credal correspond à la révision des croyances. Parmi les outils de la théorie de l'évidence, il en est un qui concerne la combinaison de deux fonctions de croyance. A partir des fonctions de masse m_1 et m_2 , la combinaison conjonctive de ces deux sources d'information ($m_\cap = m_1 \cap m_2$) peut être calculée $\forall A \subseteq \Omega$ par :

$$m_\cap(A) \triangleq \sum_{A=B \cap C} m_1(B)m_2(C). \quad (2)$$

Il est à noter que cette règle, généralement appelée règle de Dempster non normalisée, permet de combiner des informations incertaines extraites sous forme de fonctions de croyance. Si nécessaire, la condition $m(\emptyset) = 0$ peut être retrouvée en divisant chaque masse par un coefficient de normalisation. L'opération résultante, appelée règle de Dempster et notée \oplus est définie $\forall A \subseteq \Omega$ par :

$$(m_1 \oplus m_2)(A) \triangleq \frac{(m_1 \cap m_2)(A)}{1 - m(\emptyset)} \quad (3)$$

où la quantité $m(\emptyset)$ est appelée degré de conflit entre les fonctions m_1 et m_2 et peut être calculé en utilisant l'équation suivante :

$$(m_1 \cap m_2)(\emptyset) = \sum_{B \cap C = \emptyset} m_1(B)m_2(C). \quad (4)$$

L'utilisation de la règle de Dempster est possible si et seulement si m_1 et m_2 ne sont pas en conflit total, c'est-à-dire s'il existe deux éléments focaux B et C de m_1 et m_2 qui satisfassent $B \cap C \neq \emptyset$.

2.2 Niveau pignistique

L'étape d'agrégation précédemment définie permet ainsi d'obtenir un résumé exhaustif de l'information sous forme d'une fonction de croyance unique m qui est utilisée pour la prise

de décision. En basant leur raisonnement sur des arguments de rationalité développés dans le modèle des croyances transférables, Smets et Kennes (1994) proposent de transformer m en une fonction de probabilité *BetP* définie sur Ω (appelée fonction de probabilité *pignistique*) qui se formalise pour tout $\omega_k \in \Omega$ par :

$$BetP(\omega_k) = \frac{1}{1 - m(\emptyset)} \sum_{A \ni \omega_k} \frac{m(A)}{|A|} \quad (5)$$

où $|A|$ représente la cardinalité de $A \subseteq \Omega$. Dans cette transformation, la masse de croyance $m(A)$ est uniformément distribuée parmi les éléments de A .

3 Application à l'estimation de Temps de Parcours

Afin de réaliser l'estimation de temps de parcours à l'aide de la théorie de l'évidence, nous allons mettre en œuvre 2 approches différentes.

Pour la première approche, nous définissons un espace à 2 dimensions où chaque composante correspond à l'estimation d'un capteur (véhicule traceur ou boucle magnétique). Dans cet espace, nous calculons des dissimilarités entre un nouveau couple de mesures et les mesures d'apprentissage. Ces dissimilarités vont nous permettre de construire des fonctions de croyance et ainsi attribuer une classe de temps de parcours au nouveau couple de mesures.

La seconde approche correspond à une approche de fusion de classifieurs. Dans ce cas, les mesures obtenues par les capteurs sont considérées comme des classes. Elle peuvent donc être assimilées à des sorties de classifieurs qu'il convient alors de fusionner afin de prendre une décision globale plus fiable.

Dans les sections suivantes, nous détaillons ces deux approches.

3.1 Approche distance

Dans cette section, nous présentons le travail original de Denœux (Denœux (1995); Zouhal et Denœux (1998)) qui introduit l'approche distance de la modélisation des fonctions de croyance de la manière suivante.

Considérons un nouvel individu de vecteur forme x connu et de vecteur d'appartenance u inconnu. Si un élément de vecteur forme $x^{(i)}$ et d'étiquette $u_n^i = 1$ de l'ensemble d'apprentissage \mathcal{L} est proche de x dans l'espace des caractéristiques, alors une partie de la croyance sera affectée à ω_n et le reste à l'ensemble des hypothèses du cadre de discernement. Ainsi, nous obtenons alors à partir de l'élément i une masse de croyance m_i . Jusqu'à présent, nous n'avons considéré pour l'appartenance de x qu'un seul élément de \mathcal{L} . Si la même opération est répétée pour l'ensemble des I exemples d'apprentissage, on obtient alors I fonctions de croyance qui peuvent être combinées à l'aide de l'opérateur de Dempster. En pratique, les éléments éloignés de x ont peu d'influence et peuvent être négligés. Deux techniques peuvent alors être mises en œuvre. Dans la première approche, on ne prend en compte que les k plus proches voisins de x , méthode que nous appellerons KNN-DS (Zouhal et Denœux (1998)). La seconde approche repose sur la caractérisation des données d'apprentissage à l'aide de prototypes (Denœux (2000)), méthode que nous nommerons ProDS. Chacun des prototypes i , initialisés par

un algorithme de type C-means, permet la construction d'une fonction de croyance ayant l'expression suivante :

$$\begin{cases} m_i(\{\omega_n\}) &= \alpha_i \phi_i(d_i) \\ m_i(\Omega) &= 1 - \alpha_i \phi_i(d_i) \end{cases} \quad (6)$$

où $0 < \alpha_i < 1$ est une constante, $\phi_i(\cdot)$ est une fonction décroissante monotone vérifiant $\phi_i(0) = 1$ et $\lim_{d \rightarrow \infty} \phi_i(d) = 0$, d_i est la distance euclidienne entre le vecteur x et le $i^{\text{ème}}$ prototype. La fonction ϕ_i peut être une fonction exponentielle de la forme :

$$\phi_i(d_i) = \exp^{-\gamma_i(d_i)^2} \quad (7)$$

où γ_i est un paramètre associé au $i^{\text{ème}}$ prototype. Le paramètre α_i empêche l'affectation de toute la masse de croyance à l'hypothèse ω_n lorsque x et le $i^{\text{ème}}$ prototype sont égaux. Il traduit l'incertitude relative à la caractérisation du prototype i . En outre, la contrainte $\alpha_i < 1$ garantit la possibilité de combiner m_i avec n'importe quelle autre fonction de croyance puisque quelque soit d_i , on aura toujours $m_i(\Omega) > 0$ (la certitude de ω_n pourrait entraîner un conflit total avec une autre source de croyance incompatible). Le paramètre γ_i , quant à lui, permet de spécifier la vitesse de décroissance de la masse avec la distance selon le prototype.

Les fonctions de croyance m_i , obtenues pour chacun des prototypes, sont ensuite fusionnées avec la règle de combinaison de Dempster.

Nous pouvons remarquer que cette méthode repose sur l'estimation de plusieurs paramètres : nombre de voisins k (ou nombre de prototypes), position des prototypes, valeurs de γ_i et les valeurs de α_i . Au cours d'une première étude (Zouhal (1997)), il a été montré que l'approche présentée ici est très peu sensible au choix du paramètre k , et donc *a priori* au nombre de prototypes.

Pour les autres paramètres, une méthode d'optimisation, fondée sur l'utilisation de l'information contenue dans l'ensemble d'apprentissage \mathcal{L} , a été introduite par Zouhal et Denoux (1998). Cette optimisation est basée sur la minimisation d'un critère d'erreur quadratique moyenne E_{MS} entre les probabilités pignistiques et les vecteurs d'appartenance aux classes :

$$E_{MS} = \sum_{i=1}^I \sum_{n=1}^N [BetP^{(i)}(\omega_n) - u_n^i]^2 \quad (8)$$

où $BetP^{(i)}$ représente la probabilité pignistique d'un vecteur $x^{(i)}$ de la base d'apprentissage.

3.2 Approche combinaison de classifieurs

Dans la section précédente, nous avons présenté un classifieur basé sur la théorie des fonctions de croyance. 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.

Les schémas d'agrégation de ces classifieurs varient selon le type de leurs sorties. Dans le cas de notre application, les sorties des capteurs (qui peuvent être assimilés aux classifieurs) sont des classes (des estimations de temps de parcours). Certaines techniques d'agrégation de

Temps de parcours et Théorie de l'évidence

classifieurs de type classe sont basées sur les théories de l'incertain (approche bayésienne et réseaux du même nom) et de l'imprécis (approches crédibiliste et possibiliste).

Les schémas de fusion de classifieurs que nous allons introduire reposent essentiellement sur la prise en compte des erreurs des classifieurs individuels. Les erreurs de chaque classifieur sont usuellement consignées dans la matrice de confusion donnée par :

$$\mathcal{M}^j = \begin{pmatrix} n_{11} & n_{12} & \dots & n_{1j} & \dots & n_{1N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ n_{i1} & n_{i2} & \dots & n_{ij} & \dots & n_{iN} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ n_{N1} & n_{N2} & \dots & n_{Nj} & \dots & n_{NN} \end{pmatrix}$$

où N représente le nombre de classes dans le problème. La ligne i correspond à la classe ω_i de l'apprentissage et la colonne j correspond à la classe décidée par le classifieur j , i.e. \mathcal{C}^j . Cette matrice est obtenue sur un échantillon d'apprentissage et peut être considérée comme la connaissance *a priori* sur les performances du classifieur. Les éléments diagonaux sont les pourcentages de 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). On définit alors le taux de reconnaissance τ_{rec}^j et le taux de confusion τ_{con}^j par :

$$\tau_{\text{rec}}^j = \frac{\sum_{i=1}^N n_{ii}}{\sum_{i,j} n_{ij}} \quad \text{et} \quad \tau_{\text{con}}^j = 1 - \tau_{\text{rec}}^j \quad (9)$$

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 de croyance à partir de ces matrices de confusion. Ces différentes approches de fusion de classifieurs sont présentées plus en détail par Aissa et al. (2004).

3.2.1 Méthode de Xu

L'une des premières approches de fusion de classifieurs de type classe basée sur la théorie des fonctions de croyance a été proposée par Xu et al. (1992). Les fonctions de croyances sont alors définies de la façon suivante :

$$\begin{cases} m_j(\{\omega_n\}) & = \tau_{\text{rec}}^j \\ m_j(\overline{\omega_n}) & = \tau_{\text{con}}^j \end{cases} \quad (10)$$

dans le cas où le classifieur \mathcal{C}^j a sélectionné la classe ω_n . Les méthodes que nous avons développées et qui sont décrites par la suite découlent de cette construction de jeu de masses mais nous apparaissent plus précises.

3.2.2 Méthode n°2

Les fonctions de croyance construites dans la méthode précédente sont indépendantes de la classe sélectionnée par le classifieur. Les jeux de masses sont alors identiques quelque soit la classe retenue par le classifieur. Ainsi un classifieur ayant une taux de reconnaissance global important produira une fonction de croyance importante même pour une classe qu'il reconnaîtra très peu. Il apparaît alors plus judicieux de prendre en compte des taux de reconnaissance par classe. C'est cette approche qui est proposée avec cette méthode. Nous obtenons alors les fonctions de croyance de la manière suivante :

$$\begin{cases} m_j(\{\omega_k\}) &= \frac{n_{kk}}{\sum_{i=1}^N n_{ki}} \\ m_j(\overline{\omega_k}) &= 1 - \frac{n_{kk}}{\sum_{i=1}^N n_{ki}} \end{cases} \quad (11)$$

lorsque le classifieur C^j sélectionne la classe ω_k .

3.2.3 Méthode n°3

Comme pour les méthodes précédentes, celle-ci permet de construire des fonctions de croyance à partir de la matrice de confusion. Au contraire de la méthode n°2, elle permet de placer de la masse de croyance non seulement sur la classe retenue par le classifieur, son complémentaire mais aussi sur le cadre de discernement Ω . Les croyances se répartissent alors de la manière suivante :

$$\begin{cases} m'_j(\{\omega_i\}) &= \frac{n_{ki}}{\sum_{j=1}^N n_{jk}} \quad \forall i = 1, \dots, N \\ m'_j(\Omega) &= \frac{\sum_{j=1}^N n_{kj} - n_{kk}}{\sum_{j=1}^N n_{kj}} \end{cases} \quad (12)$$

lorsque le classifieur C^j sélectionne la classe ω_k . Les fonctions de masse m'_j ne sont pas normalisées, il faut alors passer par l'étape de normalisation suivante :

$$m_j(A) = \frac{m'_j(A)}{\sum_{B \subseteq \Omega} m'_j(B)} \quad \forall A \subseteq \Omega \quad (13)$$

4 Résultats

4.1 Les données traitées

Les données utilisées, dans cette étude, ont été recueillies à l'issue de la campagne de mesures réalisée à la ZELT (Zone Expérimentale, Laboratoire de Trafic) de Toulouse. L'enquête de mesure a consisté en un recueil de données issues de plusieurs sources d'information :

- capteurs à boucles magnétiques,
- véhicules traceurs,

Temps de parcours et Théorie de l'évidence

	Véhicule Traceur	Boucle Magnétique
Pourcentage de Classification	26.57	27.27

TAB. 1 – *Pourcentage de bonne classification sans le processus de fusion pour les véhicules traceurs et les boucles magnétiques.*

- une enquête de reconnaissance des véhicules par les plaques minéralogiques.

Les 2 premières sources d'information seront exploitées afin de définir une estimation du temps de parcours alors que la dernière source d'information sera considérée comme le temps de parcours de référence.

Les boucles magnétiques nous donnant des informations trafic (débit, taux d'occupation) alors que les véhicules traceurs nous donne un temps de parcours, il est nécessaire de transformer les informations fournies par les boucles magnétiques en une estimation du temps de parcours. Pour cela, nous utilisons la méthode proposée par Bonvalet et Robin-Prévallée (1987) qui est aussi l'approche retenue dans les premiers travaux sur l'estimation du temps de parcours par les fonctions de croyance (Faouzi (2000)).

Le recueil d'information a permis d'obtenir 230 données en ce qui concerne les boucles magnétiques et l'enquête minéralogique (temps de parcours de référence). Les véhicules traceurs ont fourni eux 156 observations. En réalisant, le croisement de ces deux fichiers, il apparaît que 143 observations sont communes. Pour la suite des essais, nous ne travaillerons que sur ces données.

Par ailleurs, pour les tests nous avons choisi de discrétiser notre temps de parcours en 6 classes. Ces classes sont définies à partir des temps de parcours de référence (enquête minéralogique) de manière à avoir le même nombre d'observations dans chaque classe (méthode des quantiles).

Afin d'évaluer les performances des différentes stratégies de fusion mises en œuvre dans cet article, nous présentons dans la tableau TAB. 1 les résultats de classification sans ces approches c'est-à-dire en prenant de manière individuel les véhicules traceurs et les boucles magnétiques. On peut constater sur ce tableau que l'estimation issue des boucles magnétiques semble donner de meilleurs résultats que ceux issus des véhicules traceurs.

4.2 Tests sur la constitution de la base d'apprentissage

Pour vérifier l'influence de la constitution de la base d'apprentissage nous avons fait varier le nombre d'observations dans la base d'apprentissage de 5 à 95% de la population initiale. Cela consiste alors à prendre une partie des observations initiales qui sont employées dans la base d'apprentissage et le restant des observations constituant alors la base de test. Afin de moyenner les résultats, 300 tirages de bases d'apprentissage ont été réalisés pour une proportion donnée. Ainsi sur les différentes courbes, nous avons représenté la moyenne, la valeur minimale et maximale ainsi que l'écart type du pourcentage de bonne classification.

La figure FIG. 1 montre l'évolution du pourcentage de bonne classification obtenu par la méthode de fusion de classifieurs proposée par Xu et al. (1992), en fonction du pourcentage de points dans la base d'apprentissage.

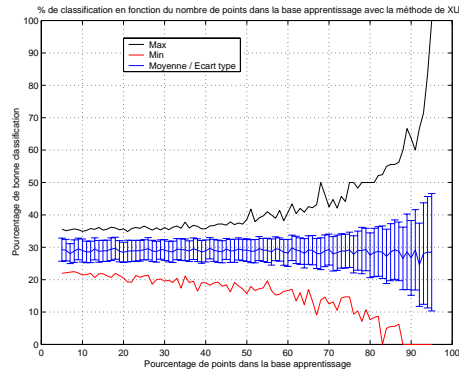


FIG. 1 – Pourcentage de bonne classification en fonction du nombre de points dans la base d'apprentissage pour la méthode de XU.

La figure FIG. 2 montre l'évolution du pourcentage de bonne classification en fonction du pourcentage de points dans la base d'apprentissage en utilisant l'amélioration, proposée dans Aissa et al. (2004), de la méthode de Xu .

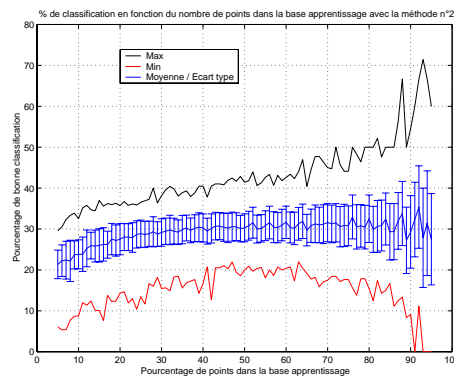


FIG. 2 – Pourcentage de bonne classification en fonction du nombre de points dans la base d'apprentissage pour la méthode $n^{\circ}2$.

La figure FIG. 3 montre l'évolution du pourcentage de bonne classification en fonction du pourcentage de points dans la base d'apprentissage pour la méthode $n^{\circ}3$ Aissa et al. (2004).

La figure FIG. 4 récapitule les pourcentages de bonne classification pour les 3 méthodes de fusion de classifieurs.

On peut constater sur ces figures que les performances obtenues par la méthode de Xu sont moins sensibles aux nombres d'observations constituant la base d'apprentissage. Toutefois, ces performances restent globalement moins bonnes que celles obtenues par les méthodes $n^{\circ}2$ et $n^{\circ}3$ lorsque 20% des observations initiales constituent la base d'apprentissage. Par ailleurs,

Temps de parcours et Théorie de l'évidence

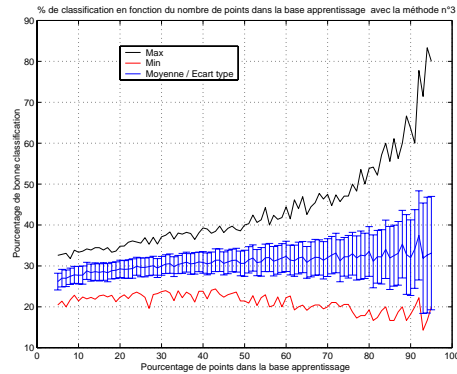


FIG. 3 – Pourcentage de bonne classification en fonction du nombre de points dans la base d'apprentissage pour la méthode n°3.

la méthode n°3 donne les meilleurs résultats, en terme de pourcentage de bonne classification, une fois que la base d'apprentissage est constituée de plus de 20% des observations initiales.

La figure FIG. 5 représente respectivement l'évolution du pourcentage de classification en fonction de la constitution de la base d'apprentissage en travaillant avec des méthodes de classification. La méthode KNN-DS donne des résultats de bonne classification inférieurs à 20% quelque soit la constitution de la base d'apprentissage. La méthode ProDS donne elle des résultats similaires à ceux obtenus par la méthode n°3 de fusion de classifieurs.

5 Conclusion

L'objectif du travail effectué dans cet article est 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 boucles électromagnétiques, qui permettent d'obtenir une estimation de temps de parcours moyen, et un échantillon réduit de véhicules traceurs, qui recueillent les temps de parcours qu'ils ont réalisés. Dans cet article, nous abordons le problème de l'estimation du temps de parcours comme un problème typique de fusion de classifieurs en utilisant la théorie des croyances pour sa finesse de modélisation de la connaissance. Dans le cadre de cette théorie, nous avons présenté deux approches afin de modéliser les informations boucles magnétiques et véhicules traceurs en fonctions de croyance.

La première technique repose sur le concept de fusion de classifieurs. Dans ce cas, chaque source d'information est considérée comme un classifieur. Pour chacun de ces classifieurs, nous définissons une matrice de confusion qui reflète le pouvoir de discrimination des sources. Cette matrice nous permet par la suite de construire les fonctions de croyance.

La seconde approche emploie une technique classique de classification basée sur le calcul de distance pour construire les fonctions de croyance.

L'utilisation de ces deux approches dans le cadre de notre application réelle s'est avérée efficace comparativement à des approches mono-capteur. Cela est d'autant plus vrai pour une technique originale de fusion de classifieurs que nous avons mise en oeuvre (Méthode n°3)

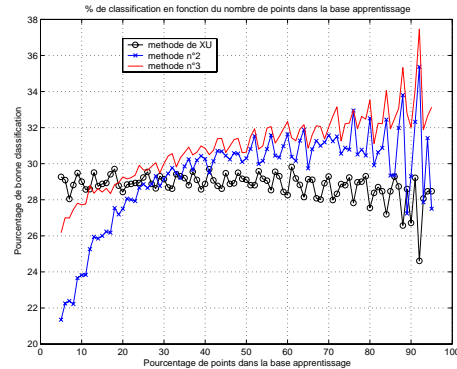


FIG. 4 – Comparaison du pourcentage de bonne classification en fonction du nombre de points dans la base d'apprentissage pour les 3 méthodes présentées.

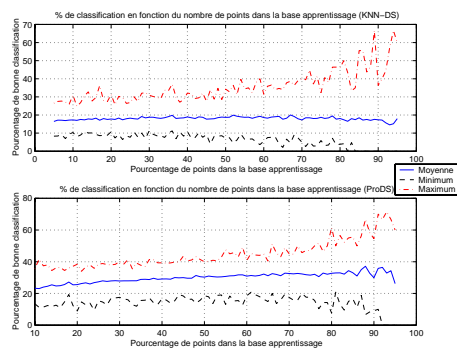


FIG. 5 – Pourcentage de bonne classification en fonction du nombre de points dans la base d'apprentissage pour les KNN-DS et ProDS.

et pour une approche classique de classification par la théorie de l'évidence (ProDS). Il est évident que l'on ne peut pas généraliser ces résultats car ils ont été obtenus dans le cadre d'une application spécifique. Toutefois, ces premiers travaux nous permettent de penser que l'emploi de techniques de fusion de données basées sur la théorie de l'évidence tendra à améliorer l'estimation du temps de parcours.

Les perspectives liées à ce travail concernent tout d'abord l'apprentissage adaptatif. Pour cette perspective, il s'agira de construire une base d'apprentissage pour chaque créneau horaire de la journée (l'amplitude du créneau horaire restant à définir). Ensuite, selon l'heure de la journée où l'on désire connaître le temps de parcours, on utilise la base d'apprentissage adéquate. Ce principe permettra de prendre en compte l'aspect dynamique et donc évolutif de l'application.

La dernière perspective concerne l'emploi des différents développements récents sur la prise en compte des aspects continus dans le cadre de la théorie de l'évidence. Ces travaux

permettraient d'estimer les temps de parcours non plus de manière discrète comme cela est le cas actuellement mais de façon continu permettant ainsi une estimation plus précise.

Références

- Aissa, A. B., N. E. E. Faouzi, et E. Lefevre (2004). Classification multisource via la théorie des fonctions de croyance. In *Workshop Fouille de Données Complexes Dans un Processus D'extraction de Connaissances, 4ème Journée d'Extraction Des Connaissances, EGC'2004*, pp. 31–44.
- Bonvalet, F. et Y. Robin-Prévallée (1987). Mise au point d'un indicateur permanent des conditions de circulation en ile-de-france. *Transport Environnement Circulation* (84-85).
- Dempster, A. (1968). A generalization of bayesian inference. *Journal of Royal Statistical Society, Serie B* 30, 205–247.
- Denoeux, T. (1995). A k-nearest neighbour classification rule based on Dempster-Shafer theory. *IEEE Transactions on Systems, Man and Cybernetics* 25(5), 804–813.
- Denoeux, T. (2000). A neural network classifier based on Dempster-Shafer theory. *IEEE Transactions on Systems, Man and Cybernetics Part A* 30(2), 131–150.
- Faouzi, N. E. E. (2000). Fusion de données pour l'estimation des temps de parcours via la théorie de l'évidence. *Recherche Transport Sécurité* (68), 15–30.
- Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton, New Jersey : Princeton University Press.
- Smets, P. et R. Kennes (1994). The transferable belief model. *Artificial Intelligence* 66(2), 191–234.
- Xu, L., A. Krzyzak, et C. Y. Suen (1992). Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition. *IEEE Transactions on Systems, Man and Cybernetics* 22(3), 418–435.
- Zouhal, L. M. (1997). *Contribution à l'application de la théorie des fonctions de croyance en reconnaissance des formes*. Ph. D. thesis, Université de Technologie de Compiègne.
- Zouhal, L. M. et T. Denoeux (1998). An evidence-theoretic k-NN rule with parameter optimization. *IEEE Transactions on Systems, Man and Cybernetics - Part C* 28(2), 263–271.

Summary

The paper address the road travel time on an urban axis by classification method based on evidence theory. The information used to estimate the travel time are complementary and redundancy. It is then necessary to implement a strategy of multi-sensors data fusion. The selected approach is the evidence theory. This theory takes into account more naturally the imprecision and uncertainty of the used data. Two strategies are implemented. The first one is a approach by classifier fusion where each information source are considered as a classifier. The second approach is a classification approach based on distance to modeling the belief function. Results of these approaches, on data collected on an urban axis of Toulouse, show the advantages of fusion in the framework of this application.