

# Fusion numérique d'informations multi-sources et extraction de connaissances : Application à l'ingénierie du trafic

NOUR-EDDIN EL FAOUZI

Laboratoire d'Ingénierie Circulation-Transports  
INRETS-ENTPE  
25, avenue François Mitterrand  
69675 Bron Cedex.  
elfaouzi@inrets.fr

**Résumé.** La fusion de données permet d'enrichir une série de données parcellaires, souvent issues de sources multiples, en combinant les informations qu'elles contiennent afin d'améliorer la qualité des connaissances extraites de ces données. L'un des objectifs de la fusion de données multi-sources, ainsi définie, est l'extraction de connaissances (KDD).

Dans cet article, plusieurs schémas de fusion de données sont proposés, exploitant à la fois la complémentarité et la redondance de ces informations multi-sources. Ces schémas sont ensuite mis en œuvre pour l'élaboration d'un indicateur de la qualité de la circulation du trafic (temps de parcours) sur la base de données issues de capteurs de trafic traditionnels et de véhicules traceurs (véhicules équipés de capteurs embarqués).

## 1 Introduction

Les nouvelles technologies de l'information et les progrès qu'ont connu les systèmes de recueil et de collecte de données ont favorisé l'émergence de nouvelles sources de données (nouveaux capteurs de mesure dotés de grande précision, capteurs embarqués à bord de mobile, localisation satellitaire,...). Celles-ci permettent de disposer d'informations de plus en plus riches et complexes, de nature et de fiabilité diverses avec de fortes exigences tant en terme de qualité que de performances. Ainsi, le décideur comme le chercheur se trouve confronté, de plus en plus, à des flux de données sans cesse croissants, souvent hétérogènes, parfois contradictoires dont il est bien difficile de faire la synthèse. La réponse à ce besoin a récemment ouvert un nouveau champ, globalisant les traitements des données disponibles, connu sous les noms de « data mining » (Han et Kamber 2001) et d'*extraction automatique des connaissances* (KDD) (Klosgen et Zytkow 2002). Ce champ se propose de fournir des techniques et des outils d'assistance à l'analyse et à l'extraction automatique des connaissances, prenant simultanément en compte l'ensemble des différentes sources d'informations disponibles.

La fusion de données dont il s'agit dans ce travail doit être distinguée de la fusion de bases de données ou de fichiers (Saporta 2002). Cette dernière consiste à agréger les bases de