

Prise en compte du réseau de sources pour la fusion d'informations

Thomas Bärecke*, Marie-Jeanne Lesot*
Herman Akdag*, Bernadette Bouchon-Meunier*

* LIP6 - Université Pierre et Marie Curie-Paris6, UMR7606
4 place Jussieu 75252 Paris cedex 05
prénom.nom@lip6.fr

1 Introduction

La cotation d'information vise à évaluer la qualité d'une information, et en particulier la confiance qu'on peut lui accorder (Cholvy, 2004; Besombes et Revault d'Allonnes, 2008; Bärecke et al., 2010). Elle se déroule en deux étapes : évaluation individuelle des éléments informationnels fournis par diverses sources, puis fusion de ces confiances en un résultat global.

Nous considérons cette seconde étape et proposons d'exploiter une connaissance a priori, fournie par un expert par exemple, sur les relations, d'affinité ou d'hostilité, entre sources : le principe sous-jacent considère que des sources indépendantes, voire hostiles, fournissant une même information lui donnent plus de poids que des sources en relation d'affinité, qui produisent naturellement une information plutôt redondante. La méthode proposée consiste à 1) identifier une partition des sources en sous-réseaux de sources amicales, indépendants ou liés entre eux par des relations d'hostilité, 2) effectuer une fusion partielle à l'intérieur de chaque sous-réseau, 3) fusionner les résultats fournis par les sous-réseaux.

2 Méthode proposée : fusion par partition

Partitionnement Etant donné deux sources v et w , on note vHw (resp. vAw) si elles sont en relation d'hostilité (resp. d'affinité). α et β étant deux poids, le coût d'une partition $\mathcal{C} = \{C_1, \dots, C_n\}$ des sources peut être défini comme $f(\mathcal{C}) = \alpha \sum_{i=1}^n |\{(v, w) \in C_i^2 | vHw\}| + \beta \sum_{(i,j) | j > i} |\{(v, w) \in C_i \times C_j | vAw\}|$. Le premier terme impose de minimiser les relations hostiles entre sources d'un même sous-groupe, le second de minimiser les relations amicales entre sources affectées à des sous-groupes différents, les paramètres α et β permettant de pondérer l'importance relative des deux critères. La partition optimale peut être obtenue par l'algorithme A* adapté à ce problème de coupe minimale, ou des stratégies de type "diviser pour régner" ou des méta-heuristiques comme les algorithmes évolutionnaires ou le recuit simulé.

Fusion Au sein d'un sous-groupe, les sources sont alors globalement en relation d'affinité : l'unanimité est attendue et ne doit pas renforcer la confiance. Aussi, nous proposons d'effectuer une agrégation de type compromis (Detyniecki, 2000) des confiances individuelles, par