

# 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).  $\alpha$  et  $\beta$  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  $\alpha$  et  $\beta$  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

exemple par une moyenne. Une autre interprétation, plus sévère, exige que les sources amies soient cohérentes, et pénalise d'éventuelles contradictions internes : un opérateur conjonctif, tel que le minimum, peut également être envisagé.

Lors de la fusion des résultats des différents groupes, une redondance peut être interprétée comme un renforcement : des groupes de sources indépendants voire hostiles accordant une confiance élevée à une même information conduisent à une confiance globale plus élevée encore. De même, une unanimité sur une confiance faible doit pouvoir amener à une valeur finale plus faible encore. Enfin, si un désaccord est observé, un comportement de compromis peut être adopté. Les opérateurs d'agrégation dits à attitude variable et à renforcement total (Detyniecki, 2000), comme la moyenne symétrique par exemple, présentent ces propriétés simultanément. Des variantes pondérées tenant compte des tailles des groupes de sources semblent de plus appropriées.

### 3 Conclusions et perspectives

Outre des expérimentations permettant de valider l'approche proposée pour la prise en compte des relations entre sources émettant des éléments informationnels, les travaux en cours visent à enrichir encore la fusion, selon la temporalité des éléments d'information : celle-ci permet par exemple de tenir compte de leur ancienneté ou de dynamiques de confirmation/infirmité qui influencent l'évaluation de la confiance.

**Remerciements** Ces travaux ont été réalisés dans le cadre du projet CAHORS, financé par l'ANR - CSOSG'08.

### Références

- Bärecke, T., T. Delavallade, M.-J. Lesot, F. Pichon, H. Akdag, B. Bouchon-Meunier, P. Capet, et L. Cholvy (2010). Un modèle de cotation pour la veille informationnelle en source ouverte. In *Actes du 6ème colloque Veille Stratégique Scientifique & Technologique*.
- Besombes, J. et A. Revault d'Allonnes (2008). An extension of STANAG2022 for information scoring. In *Proc. of the Int. Conf. on Information Fusion*, pp. 1635–1641.
- Cholvy, L. (2004). Information evaluation in fusion : a case study. In *Proc. of IPMU'04*.
- Detyniecki, M. (2000). *Mathematical Aggregation Operators and their Application to Video Querying*. Ph. D. thesis, University of Pierre and Marie Curie.

### Summary

The fusion step of the cotation process aggregates the confidence values associated to individual assertions. We propose a fusion procedure that exploits relationships between sources, more precisely whether they are friendly or hostile. From the representation of the source network as an edge-valued graph, we propose a decomposition method to identify affinity sub-graphs linked by hostility relationships, and a fusion method that exploits this partition.