

## DENOMBREMENT D'HYPOTHESES EN ASSOCIATION-FUSION DE DONNEES

D. PILLON

P. BLANC-BENON

THOMSON-SINTRA ASM BP 138 06561 VALBONNE CEDEX

### RÉSUMÉ

En association-fusion de données sonar, on est confronté à un flot d'informations tel qu'un opérateur humain est incapable de traiter le problème. On introduit donc des algorithmes automatiques permettant de trouver parmi les partitions hypothèses celle qui est la plus vraisemblable. On montre comment, au moyen de techniques de dénombrement, une recherche exhaustive de la solution optimale est inenvisageable en pratique, du fait de l'explosion combinatoire. La structure en treillis de l'ensemble des partitions permet néanmoins de dégager un algorithme sous-optimal, dit de recherche hiérarchisé. Enfin une généralisation sur les mêmes principes est proposée au cas où l'on associe des sorties des sous-systèmes locaux de fusion.

### ABSTRACT

In the field of data-fusion for sonar applications, one has to cope with such overwhelming data flows that even human operator is unable to solve the problem. Automatic algorithms are so introduced unabling to find among all the hypothetical partitions the most likely one. We show how an exhaustive search of the optimal solution is not conceivable by means of exact counting of all the partitions and combinatorial reasoning. However, the lattice structure attached to the set of partitions enables to derive a sub-optimal algorithm which only examines compatible partitions in a hierarchical search. Finally, a generalisation of such problems is proposed for the fusion of outputs issued from local data-fusion centers.

### 1 - INTRODUCTION

Pour continuer à assurer la détection passive de sources de plus en plus silencieuses, on a été amené à équiper les sous-marins d'antennes de plus en plus grandes et dotées de plus en plus d'hydrophones. Une conséquence indirecte de cette augmentation des gains d'antenne a été la saturation des opérateurs humains par un très important flot d'informations.

En effet, la portée a certes augmenté face à des sources discrètes, mais aussi face aux nombreux bâtiments bruyants (cargos, pétroliers, chalutiers...) présents en permanence sur de vastes zones océaniques. Si l'on n'effectue pas judicieusement la synthèse de cette multitude de détections, l'avantage opérationnel attendu par l'emploi de tels récepteurs performants s'amenuise.

Ce problème est rendu encore plus complexe par la présence sur le même bâtiment de nombreuses antennes nécessaires pour assurer la couverture spatiale et fréquentielle complète de l'horizon de veille. De plus, chacune des antennes est suivie de plusieurs types de traitements : Bande Large (BL), Bande Etroite (BE), démodulation, adaptatif... Enfin, ce besoin de synthèse d'informations ne concerne pas uniquement la partie sonar : par exemple, lorsque le sous-marin est à l'immersion périscopique, s'ajoute une avalanche de données non-acoustiques (optique, veille électromagnétique...)

Aussi, pour donner au commandement une vue claire et synthétique de l'ensemble des sources environnantes, faut-il associer les informations censées provenir d'un même bruiteur et fusionner ces données afin d'estimer la cinématique et le type de ce contact.

Parmi les problèmes rencontrés lorsqu'on cherche à développer cette fonction d'association-fusion, sorte de prolongement des fonctions de perception élémentaires de chaque senseur, est celui dit "d'explosion combinatoire". Ce terme vague, issu du domaine des systèmes experts, est relatif au dénombrement d'hypothèses multiples à tester lors de la phase d'association.

Cette phase est une opération de décision multi-hypothèses qui permet de déterminer, au sens d'un certain critère, quelle est la meilleure partition, de  $N$  mesures en  $P$  groupes. La difficulté provient de la méconnaissance à la fois du nombre de groupes et à la fois de la composition de ces groupes. Cette étape d'association est suivie de l'estimation du meilleur représentant moyen de chacun des groupes discernés. Ce cas est par exemple celui de l'association-fusion de pistes BE issues d'une même plate-forme.

Ce que nous allons surtout faire ici c'est le dénombrement exact des hypothèses à tester dans le cas où chacune des mesures peut s'associer avec une quelconque de ses congénères. Nous présenterons et comparerons deux approches, l'une où l'on scrute exhaustivement toutes les hypothèses, l'autre, sous-optimale, où l'on exploite la structure particulière du graphe reliant ces hypothèses.

Enfin, nous étendrons cette approche au cas plus complexe où certains éléments appartiennent à des catégories a priori inassociables, le cas illustratif étant alors l'association de pistes issues de sous-systèmes locaux de fusion.

### 2 - VRAISEMBLANCE D'UNE PARTITION

Supposons que l'on ait  $N$  éléments à associer en groupes. Ces éléments peuvent être indifféremment des mesures primaires (azimut  $\beta$ , fréquence, ...), des pistes représentées par le vecteur d'état issu du pistage angulaire (par exemple  $\beta$ ,  $\dot{\beta}$ ) ou les éléments modélisant la cinématique du bruiteur. Pour alléger l'exposé et sans nuire à la généralité, nous allons considérer des variables mono-dimensionnelles  $z_i$ ,  $i \in [1, \dots, N]$ . De plus, on suppose que l'on connaît la variance (ou covariance) de chacun de ces éléments. Ici, dans notre contexte statistiquement simplifié, mais combinatoirement complexe, on supposera que les  $z_i$  ont tous le même  $\sigma$ .



Le problème décisionnel est de trouver la partition la plus vraisemblable. Rappelons qu'une partition à P classes est un ensemble de P sous-ensembles deux à deux disjoints et dont l'union est égale à l'ensemble initial des N z<sub>i</sub>. Par exemple, pour quatre mesures z<sub>1</sub> à z<sub>4</sub>, on notera (z<sub>1</sub>z<sub>2</sub>/z<sub>3</sub>/z<sub>4</sub>) la partition à trois classes où z<sub>3</sub> et z<sub>4</sub> sont des singletons. Pour un nombre de classes donné, une partition est définie par l'indication de l'appartenance d'un élément z<sub>i</sub> à une classe C<sub>i</sub> ; on notera J(i), cette surjection définissant entièrement une partition donnée.

En supposant le contexte gaussien indépendant, la vraisemblance d'une telle partition de variables monodimensionnelles de même sigma est donnée par :

$$V(Z,P,J) = (2\pi\sigma^2)^{-N/2} \cdot \exp \left[ - \sum_{i=1}^N (z_i - a_{J(i)})^2 / 2\sigma^2 \right] \quad (1)$$

avec : Z = [z<sub>1</sub>, ... , z<sub>N</sub>], les mesures à associer.

A = [a<sub>1</sub>, ... a<sub>p</sub>] les moyennes des P groupes.

Dans [1] ou [2], est donnée la formule générale pour des z<sub>i</sub> multidimensionnels. Ni le nombre P de classes, ni les représentants a<sub>k</sub> de chaque classe ne sont connus pour décider de la meilleure partition, c'est-à-dire de la surjection J qui rend (1) maximum.

### 3 - RECHERCHE EXHAUSTIVE ET PROBLEME DE DENOMBREMENT

Pour résoudre le problème précédent, on peut commencer par se fixer le nombre de classes P. Il suffit alors de calculer la vraisemblance (1) de chacune de ces partitions à P classes. Comme on ne connaît pas les a<sub>k</sub>, on leur substitue la moyenne  $\hat{a}_k$  estimée à partir de chacun des P sous-ensembles de la partition (vraisemblance généralisée) :

$$\hat{a}_k = \frac{1}{n_k} \sum z_j^{1(k)} \quad (2)$$

où : n<sub>k</sub> est le cardinal de la classe k  
z<sub>j</sub><sup>1(k)</sup> sont les éléments de cette classe

On détermine ainsi  $\hat{J}_P$ , la partition à P classes la plus vraisemblable.

Dans un second temps, on peut faire varier P de 1 à N. Le problème est que  $\hat{J}_P$  est une fonction strictement croissante en P et donc que la partition la plus vraisemblable sera celle des N singletons. Ce problème n'est autre que celui du choix optimal de l'ordre d'un modèle comme cela arrive si souvent en traitement du signal. On sait alors qu'il existe un certain nombre de tests (Akaike, Rissanen...) permettant de déterminer le nombre optimal de classe P\* qui en quelque sorte n'est autre que le nombre minimum de paramètres suffisant à la modélisation des mesures (principe de parcimonie). En pratique, tout est basé sur l'exploitation du changement de pente de  $\hat{J}_P$  en fonction de P.

Ce qui nous intéresse surtout ici, c'est le nombre de vraisemblances qu'il faut calculer pour obtenir  $\hat{J}_{P^*}$ . Tout d'abord, pour obtenir  $\hat{J}_P$ , il faut faire autant de calculs de la formule (1) qu'il y a de partitions à P classes dans un ensemble de N éléments. Ce nombre, appelé nombre de Stirling de 2ème espèce S(N,P), est donné par la formule bi-récurrente suivante :

$$S(N,P) = S(N-1, P-1) + P S(N-1, P) \quad (3)$$

avec : S(1,1) = 1

Pour le démontrer, il suffit de faire le raisonnement suivant : supposons connu le nombre de partitions de N-1 élément en P classes et ajoutons un N<sup>ième</sup> élément. Soit on le met dans l'une des classes (et il y a P façons de le faire), soit il forme à lui seul une classe (singleton). D'où la formule (3).

Cette formule est analogue à celle des combinaisons (notée ici C(N,P)) et l'on peut, comme pour le triangle de Pascal, construire un triangle de "Stirling".

1							1						
1	1						1	1					
1	2	1					1	3	1				
1	3	3	1				1	7	6	1			
1	4	6	4	1			1	15	25	10	1		
1	5	10	10	5	1		1	31	90	65	15	1	
1	6	15	20	15	6	1	1	63	301	350	140	21	1

Pascal

Stirling

Donnons aussi, sans démonstration, une formule explicite des nombres de Stirling :

$$S(N,P) = 1/P! \sum_{k=0}^P (-1)^{P-k} C(P,k) k^N \quad (4)$$

Lorsqu'on veut déterminer la partition optimale, c'est la somme cumulée des nombres de Stirling qui intervient. Notons

$$\pi(N) = \sum_{k=1}^N S(N,k), \text{ alors :}$$

$$\pi(N) = \sum_{k=0}^{N-1} C(N-1,k) \pi(k) \quad (5)$$

Pour démontrer (5), on considère les partitions sur N-1 éléments et un N<sup>ième</sup> élément que l'on rajoute. Le nombre de partitions à N éléments s'obtient en prenant une partie quelconque E<sub>k</sub> des N-1 éléments à k éléments. A ce E<sub>k</sub>, on ajoute le n<sup>ième</sup> élément, formant ainsi l'ensemble E'<sub>k</sub>. Le nombre de partitions de N éléments contenant E'<sub>k</sub> est égal au nombre de partitions de N-1 éléments contenant E<sub>k</sub>. Il y a C(N-1, k) façon de choisir ce E<sub>k</sub>. Comme k varie de 0 à N-1, on obtient la formule (5).

On peut aussi utiliser la formule approximative suivante proposée par [3] :

$$\pi(N) \approx N^N (e \cdot \text{Log } N)^{N/(1/\text{Log } N-1)} \quad (6)$$

Bien qu'inférieur à la factorielle,  $\pi(N)$  est astronomiquement grand comme le montre le tableau ci-dessous :

N	1	2	5	10	20	50
$\pi(N)$	1	2	52	115975	5.10 <sup>13</sup>	2.10 <sup>47</sup>
N!	1	2	120	3628800	2.10 <sup>18</sup>	3.10 <sup>64</sup>

Ces chiffres montrent qu'au-delà d'une dizaine d'éléments, la stratégie de scrutation exhaustive de toutes les partitions hypothèses est inenvisageable. Nous allons maintenant décrire une méthode non optimale scrutant nettement moins de partitions.

#### 4 - STRUCTURE EN TREILLIS ET ALGORITHME DE RECHERCHE HIERARCHISE

Les partitions peuvent être reliées entre elles par une forme particulière de graphe, dit treillis qui suppose l'existence d'une relation d'ordre entre partitions. Plus précisément, il s'agit de la relation d'ordre partiel "... est plus fine que ...": une partition plus fine qu'une autre est telle que chacune de ses classes est incluse dans une classe de sa partition supérieure. On mettra donc ces deux partitions (dites compatibles) en relation l'une avec l'autre dans le graphe (figure 1). Ce graphe est dit treillis car toute paire d'éléments a une borne supérieure (i.e. le plus petit des majorants) et une borne inférieure unique (voir figure 2, leur construction sur un exemple de 2 partitions à 3 classes). Ces graphes particuliers que sont les treillis bénéficient d'un certain nombre de propriétés [4] mais c'est surtout cette notion d'ordre partiel qui va permettre de réduire drastiquement le nombre de partitions à tester. En effet, on va se limiter à parcourir un chemin particulier à travers le treillis.

Il consiste à commencer par la partition la plus fine (les singletons) et à ne tester que des partitions compatibles entre elles à chaque niveau dans le treillis, c'est-à-dire des partitions ayant le même antécédent. L'algorithme de recherche correspondant est dit hiérarchique : il revient à choisir, parmi les partitions compatibles celle qui a la plus grande vraisemblance, puis de réitérer le même processus de choix à partir de la partition sélectionnée. On progresse ainsi en réduisant petit à petit le nombre de classes jusqu'à ce que le test d'arrêt n'intervienne pour déterminer  $P^*$ . Notons, au passage, que la vraisemblance d'une partition se calcule aisément à partir de la vraisemblance de la partition compatible plus fine juste antécédente [2].

En suivant un tel cheminement (voir exemple figure 1) de partitions compatibles entre elles, on ne calcule au maximum plus que  $\sim N^3/6$  vraisemblances. En effet, à chaque étape on ne doit tester que  $C(k,2) = k(k-1)/2$  partitions compatibles. On le fait depuis  $k = N$  jusqu'à  $k = 2$ .

Le nombre maximal de partitions compatibles testées est donc égal à :

$$\sum_{k=2}^N k(k-1)/2 = N(N-1)(N+1)/6 \quad (7)$$

en se rappelant que  $\sum_{k=1}^N k^2 = N(N+1)(2N+1)/6$  (8)

Cette méthode hiérarchisée est sous optimale par rapport à l'approche exhaustive présentée au paragraphe précédent, mais on a constaté sur simulation que les écarts entre les solutions obtenues par ces deux méthodes sont très faibles. Ceci est peut-être dû au nombre important de chemins différents qui permettent, dans le treillis, d'aller de la partition initiale des singletons à la partition solution. Bien sûr, cette comparaison ne peut se faire qu'avec un nombre limité d'éléments car, comme on l'a vu au paragraphe précédent, dès que  $N$  dépasse la douzaine, l'approche exhaustive est inenvisageable sur les calculateurs actuels.

#### 5 - GENERALISATION : PARTITIONS AVEC INTERDITS

Considérons le cas pratique où la fusion est faite par plusieurs systèmes locaux qui transmettent leur résultat de synthèse à un centre global de fusion. Comme on ne va pas remettre en cause le travail effectué au niveau de chaque centre local, il s'agit de faire la fusion intercentre. On considèrera ici les cas 2 ou 3 centres et donc le problème d'association dans lesquels les classes ne comportent (respectivement) que 2 ou 3 éléments au maximum.

Soient  $N$ ,  $M$  et  $L$  éléments à associer issus de 3 sous-systèmes et  $D(N,M)$  et  $T(N,M,L)$  les nombres de partitions pour 2 et 3 origines respectivement ( $N \geq M \geq 1$  sans restriction de généralité). Le nombre  $k$  de classes d'une partition à deux origines est compris entre  $N$  et  $N+M$ . Le nombre total  $D(N,M)$  s'écrit donc :

$$D(N,M) = \sum_{k=N}^{N+M} d_k \quad (9)$$

où  $d_k$  désigne le nombre de telles partitions à  $k$  classes. Ces partitions ont au plus 2 éléments par classe. Soit donc  $i$  le nombre de singletons restant sur les  $M$  éléments de la seconde origine,  $i$  varie entre 0 et  $M$ , donc :

$$D(N,M) = \sum_{i=0}^M d_{N+i} \quad (10)$$

En dénombrant les façons de choisir les  $i$  singletons sur les  $M$  seconds éléments, les  $N-(M-i)$  singletons sur les  $N$  premiers et en tenant compte des  $(M-i)!$  permutations dans les  $M-i$  paires formées, il vient :

$$d_{N+i} = C(M,i) C(N,N-M+i)(M-i)! \quad (11)$$

$$d'où : D(N,M) = \sum_{i=0}^M C(M,i) A(N,i) \quad (12)$$

contrairement au cas de  $\pi(N)$ , il existe pour les  $D(N,M)$  une formule récurrente :

$$D(N,M) = MD(N-1, M-1) + D(N,M-1) \quad (13)$$

En quelque sorte, pour  $\pi(k)$ , l'ordre de la récurrence est  $k$  car il y a  $k$  origines distinctes puisqu'il n'y a pas d'interdit. Si l'on compare  $D(N,M)$  à  $\pi(N+M)$ , on peut constater un gain allant de 50 pour 10 éléments à plus  $10^{18}$  pour 50 éléments. Ces chiffres montrent l'intérêt combinatoire de la structure hiérarchisée faite de centres locaux suivis d'un centre global. La figure 3 montre un exemple de fusion globale de 4 pistes issues de 2 sous-systèmes à  $N$  et  $M$  pistes, ainsi que le sous-treillis extrait du treillis complet de la figure 1.

Si l'on pense au cas à 3 origines distinctes, on peut montrer de façon analogue que :

$$T(N,M,L) = \sum_{i=0}^L d_{M+i} D(N,M+i) \quad (14)$$

ce qui donne la formule de récurrence d'ordre 3 suivante :

$$T(N,M,L) = ML T(N-1, M-1, L-1) + LT(N-1, M, L-1) + T(N, M, L-1) \quad (15)$$

Là aussi, la réduction du nombre d'hypothèses est importante, que ce soit dans l'approche hiérarchisée ou exhaustive.

#### 6 - CONCLUSION

On peut considérer, après avoir traité des problèmes de dénombrement, que la comparaison des performances entre les approches exhaustive et non exhaustive, comme par exemple ici l'approche hiérarchisée, reste un problème ouvert. Il y a en effet nettement trop d'hypothèses pour pouvoir tester sur simulation la méthode exhaustive. On ne pourra que se contenter de comparer entre elles des approches non-exhaustives sous optimales.



REFERENCES

- [1] K.F. GONG, M.L. GRAHAM, A.G. LINDGREN, "Data-Fusion in a Multisensor Multicontact Environment" 20th Asilomar Conf. on Signals, Systems & Computers, Pacific Grove, (CA - USA) November 1986.
- [2] P. BLANC-BENON, D. PILLON "Multi-tracks Association for Underwater Passive Listening" NATO ASI Conf., Kingston (Canada), July 1988.
- [3] N. de BRUIJN "Asymptotic methods in analysis", North-Holland, 1961.
- [4] M. BARBUT, B. MONJARDET "Ordre de Classification" Hachette Paris 1970.

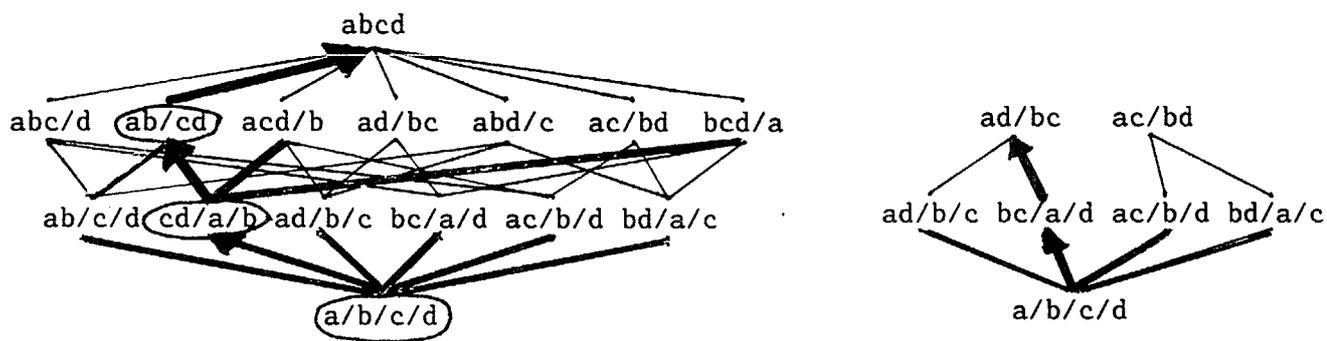


Figure 1 - Treillis des partitions (4 éléments) et cheminement entre partitions compatibles

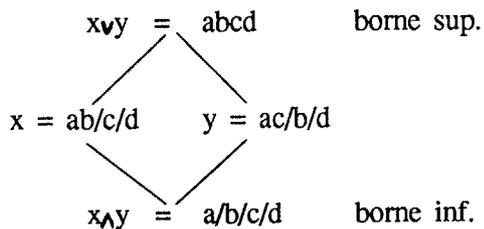


Figure 2 - Bornes supérieure et inférieure de deux partitions

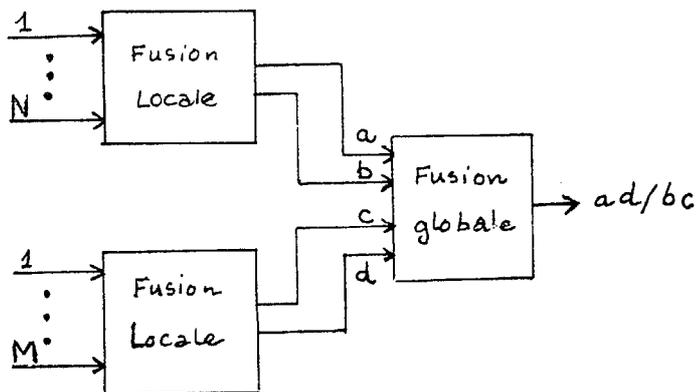


Figure 3 - Fusion globale / locale