



HAL
open science

Fuzz-CBA: Classification à base de règles d'association floues et systèmes de recommandation

Joel Pinho Lucas, Anne Laurent, Maria N. Moreno, Maguelonne Teisseire

► To cite this version:

Joel Pinho Lucas, Anne Laurent, Maria N. Moreno, Maguelonne Teisseire. Fuzz-CBA: Classification à base de règles d'association floues et systèmes de recommandation. LFA: Logique Floue et ses Applications, Sep 2009, Annecy, France. pp.283-290. lirmm-00430509

HAL Id: lirmm-00430509

<https://hal-lirmm.ccsd.cnrs.fr/lirmm-00430509>

Submitted on 7 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fuzz-CBA : Classification à base de règles d'association floues et systèmes de recommandation

Joel Pinho Lucas^{1,2}, Anne Laurent², María Moreno¹ et Maguelonne Teisseire^{2,3}

¹ Departamento de Informática y Automática, Universidad de Salamanca - Plaza de la Merced s/n 37008, Salamanca, Spain

² LIRMM-Univ. Montpellier 2-CNRS UMR 5506 161 rue Ada, 34095 Montpellier

³ Cemagref, UMR TETIS - Maison de la télédétection 500, rue J.F.Breton 34093 Montpellier Cedex 5

Résumé :

Les systèmes de recommandation sont de plus en plus présents, notamment sur Internet où il est maintenant usuel de proposer par exemple un hôtel ou un film à une personne en fonction de son profil et de ce qu'ont choisi les internautes lui ressemblant. Ces systèmes de recommandation s'appuient souvent sur des systèmes dédiés initialement à la classification et notamment à la classification à partir de règles d'association, ceux-ci étant capables d'expliquer les décisions. Cependant, même s'il existe des approches permettant d'extraire des règles d'association et motifs séquentiels flous quand les données sont numériques, celles-ci n'ont pas été étendues pour la classification et la recommandation. Dans cet article, nous nous intéressons donc à l'utilisation de règles d'association floues pour la définition d'un système de recommandation et nous montrons l'intérêt de cette approche à travers différentes expérimentations.

Mots-clés :

Fouille de données, Règles d'association floues, Classifieurs à base de règles, Systèmes de recommandation.

Abstract:

Recommender systems are more and more used, especially on the Internet (e.g. movie and hotel recommendation). These systems often rely on classification methods, especially on association rule-based methods. However, the methods based on association rules and sequential patterns have not been studied for recommender systems in the case of numerical attributes. We thus propose an original recommendation method based on fuzzy association rules and show its interest through experiments.

Keywords:

Data Mining, Fuzzy Association Rules, Rule-Based Classifiers, Recommendation Systems.

1 Introduction

Les systèmes de recommandation visent à proposer aux utilisateurs les produits qui leur conviennent le mieux. Selon [CKLT03] et Lee et al. [LKR01], les méthodes utilisées dans les systèmes de recommandation peuvent être di-

visées en deux catégories principales : le filtrage collaboratif et les méthodes basées sur le contenu. Les méthodes basées sur le contenu comparent les contenus des données avec des profils prédéfinis [LKR01]. Par conséquent, les systèmes de recommandation qui utilisent ce type de méthode ne prennent pas en compte les informations acquises par d'autres utilisateurs. D'un autre côté, les méthodes de filtrage collaboratif sont basées sur des opinions recueillies auprès d'autres utilisateurs [CGM⁺99]. Dans [SKKR01], le filtrage collaboratif repose sur l'algorithme du plus proche voisin. Ainsi, les produits recommandés seront ceux que les utilisateurs d'intérêts similaires ont aimé. [BHK98] a classé les méthodes de filtrage collaboratif en deux groupes : les méthodes basées sur la mémoire (memory-based methods), qui sont également dénommées méthodes basées sur l'utilisateur (user-based methods), et les méthodes basées sur les modèles (model-based methods), qui sont également dénommées méthodes basées sur les items (item-based methods). Les méthodes basées sur la mémoire utilisent les approches prédictives (par exemple les k-plus proches voisins) et un ensemble d'entraînement stockant les opinions et avis d'autres utilisateurs. Cependant les données sont souvent très peu denses. Par exemple, les systèmes issus du commerce électronique offrent généralement des millions de produits à la vente, et il n'est pas possible d'obtenir des avis sur tous les produits [CLZ05]. Par conséquent, les systèmes de recommandation n'utilisent géné-

ralement pas uniquement les méthodes basées sur la mémoire. Les méthodes de recommandation basées sur le contenu sont habituellement employées en liaison avec les méthodes de filtrage collaboratif. Les techniques d'apprentissage automatique et de fouille de données sont les méthodes les plus employées. Certains systèmes emploient également les agents [CZC⁺03, MGPL04].

Nous détaillons ci-dessous les problèmes les plus critiques liés aux méthodes employées pour les systèmes de recommandation. Notons que, comme pour tout système de classification supervisée, il existe deux types d'erreurs principales : les faux négatifs (produits non recommandés alors qu'ils auraient pu) et les faux positifs (produits recommandés mais qui ne plaisent pas). Selon [SKKR00], les faux positifs sont plus critiques car ils conduisent à l'insatisfaction des utilisateurs.

Dans cet article, nous proposons l'élaboration de modèles de recommandation en utilisant des règles d'association dans une perspective de prévision. L'utilisation de ces systèmes à base de règles d'association permet de traiter de gros volumes de données en s'appuyant sur les algorithmes développés par la communauté de fouille de données (algorithmes par niveaux, *a la* APriori notamment), tout en permettant d'expliquer la décision prise par la présence des règles prises en compte à chaque classification d'un nouvel exemple. Cependant, les données habituellement collectées dans les systèmes de recommandation contiennent des attributs numériques, ce que les méthodes existantes de classification à partir de règles d'association ne permettent pas de traiter. En effet, même s'il existe de nombreuses propositions pour traiter les règles d'association numériques et floues [SA96, CA97, DP03], ces travaux n'ont jamais été intégrés dans les systèmes de classification ou de recommandation. Nous proposons donc une méthode basée sur les règles d'association floues, et montrons son intérêt au travers d'expérimentations menées sur différentes bases de données.

Dans la suite de cet article, nous présentons tout d'abord les systèmes de classification à base de règles d'association classiques, avant de nous intéresser à la méthode proposée ici, reposant sur les règles d'association floues.

2 Classification à base de règles d'association et Utilisation de la logique floue

Comme indiqué précédemment, les algorithmes d'extraction de règles d'association peuvent être utilisés pour construire des modèles de recommandation, comme développé dans [LKR01]. Les règles d'association ont été introduites par [AIS93] visant, d'abord, à découvrir les motifs de consommation dans des bases de données commerciales. Une règle d'association décrit une relation associative entre ensembles d'attributs (itemsets) qui se trouvent fréquemment en même temps dans les transactions d'une base de données. Ainsi, il ne s'agit pas de prédire, mais plutôt de décrire les données. L'utilisation des règles d'association dans la classification a notamment été proposée dans CBA (Classification Based in Association) [LHM98], CPAR (Classification based on Predictive Association Rules) [YH03], MCAR (Multi-class Classification based on Association Rule) [TCP05], CMAR (Classification based on Multiple class-Association Rules) [LHP01] et TFPC (Total from Partial Classification) [CLZ05]. Comme les règles de classification classiques ne permettent pas de choisir une cible prédéfinie, [LHM98] a introduit des contraintes sur les termes conséquents des règles. De cette façon, le terme conséquent d'une règle d'association représente l'attribut classe.

Les systèmes de classification à base de règles d'association fonctionnent en deux étapes principales : (1) création de la base de règles d'association pour chaque classe, (2) classification d'un nouvel exemple.

La création de la base de règles correspond

à l'extraction des règles d'association de la forme $r = \text{Produit}_1, \dots, \text{Produit}_n \rightarrow \text{Classe}_i$ puis au calcul de la confiance de ces règles : $\text{conf}(r) = \frac{\text{supp}_C(r)}{\text{supp}(r)}$ avec $\text{supp}_C(r)$ le support de la règle dans la sous-base constituée des exemples correspondant à la classe considérée.

Pour tout nouvel exemple pour lequel il faudra proposer une classe cible, il s'agira alors de trouver toutes les règles qui s'appliquent et de considérer la classe majoritaire apparaissant comme cible de ces règles.

Cependant, comme précisé ci-dessus, les règles d'association considérées dans les systèmes de classification ne permettent pas de prendre en compte les bases de données numériques. Des méthodes basées sur la logique floue ont donc été proposées dans la littérature, comme par exemple dans [DHP06]. Ainsi, [Yag03] propose de représenter et de décrire des informations dans un système de recommandation basé sur le contenu. [CLL05] ont implémenté quant à eux un système de recommandation basé sur les concepts flous afin de recommander aux utilisateurs des produits qui ne sont pas consommés d'habitude, tandis que [NFJK99] s'appuie sur le clustering flou pour effectuer des recommandations en accord avec des profils utilisateurs déterminés par des techniques de clustering flou. Enfin, [dCFLH08] combine la théorie des ensembles flous avec les réseaux bayésiens.

3 Fuzz-CBA : Recommandation et Classification à base de règles d'association floues

Dans les systèmes classiques, les utilisateurs similaires sont regroupés, et on recommande à tout nouvel utilisateur les produits correspondant au groupe le plus proche. Dans l'approche que nous proposons, nous souhaitons associer l'utilisateur non pas à un seul groupe, mais à plusieurs, et lui proposer donc une liste de produits associés aux différents groupes dont il est proche.

Dans cette section, nous décrivons notre propo-

ID	âge	emploi	film	Année	note
1	23	etud.	Batman	1995	1
...

Tableau 1 – Exemple de base de données

sition Fuzz-CBA et les principales phases qui la composent : (1) création des groupes d'utilisateurs, (2) création des bases de règles d'association floues pour chaque classe, et (3) classification d'un nouvel exemple.

3.1 Création des groupes d'utilisateurs

Nous considérons des bases de données d'apprentissage décrivant les préférences des utilisateurs en fonction de certains critères, comme le montre le tableau 1.

La première phase de notre méthode est de créer des groupes d'utilisateurs. Pour cela, nous utilisons un algorithme de clustering classique, tel que les k-means. A l'issue de cette étape, nous disposons donc de différents groupes d'utilisateurs G_1, \dots, G_U décrits en fonction de leur identifiant. Il s'agit alors de décrire de quels groupes un nouvel utilisateur est le plus proche en fonction des valeurs de différents attributs.

3.2 Création des bases de règles

A partir des groupes d'utilisateurs décrits ci-dessus, nous visons à définir les règles décrivant au mieux chacun d'entre eux. La base est donc partitionnée en fonction des groupes et les règles d'association floues sont extraites. La gestion des attributs numériques est effectuée en construisant des partitions floues : (i) soit par une approche de type *equi-width* découpant l'univers en parties de tailles égales, (ii) soit par une approche de type *equi-depth* découpant l'univers en parties comprenant le même nombre d'éléments, ces partitions étant ensuite fuzzifiées en considérant des sous-ensembles flous trapézoïdaux dont le noyau est constitué par le sous-ensemble central et dont le support est augmenté par les deux sous-ensembles voi-

sins (inférieur et supérieur).

Chaque attribut A_i de la base de données est donc maintenant décrit selon différentes descriptions détaillées issues de la partition floue (e.g. âge moyen). On notera ici A_i un attribut, A_i^j un sous-ensemble flou de la partition décrite sur l'univers de A_i (domaine de définition), et $\mu_{A_i^j}$ la fonction d'appartenance de ce sous-ensemble flou. Un n-uplet de la base de données DB sera quant à lui noté L .

Il s'agit alors de suivre les algorithmes classiques de construction de règles d'association, en construisant d'abord les ensembles d'attributs-descriptions apparaissant fréquemment dans la base de données (au sens d'un support minimal). On nomme *support* le nombre d'occurrences d'un ensemble d'attributs-descriptions dans la base de données. Ce support est calculé de la manière suivante :

$$supp(A_{i1}^{j1}, \dots, A_{ik}^{jk}) =$$

$$\sum_{L \in DB} \overline{\top}(\mu_{A_{i1}^{j1}}(L[A_{i1}]), \dots, \mu_{A_{ik}^{jk}}(L[A_{ik}]))$$

où $\overline{\top}$ est la généralisation d'une t-norme au cas n-aire : $\overline{\top}(x_1, \dots, x_n) = \top(x_1, \top(x_2, \dots, \top(x_{n-1}, x_n)))$.

Notons que l'algorithme d'extraction des règles se contente de rechercher les règles de la forme $A_{i1}^{j1}, \dots, A_{ik}^{jk} \rightarrow G_u$ consistant à prédire un produit qui plaira à un utilisateur en fonction d'un ensemble de critères sur la valeur des attributs. Par exemple on aura : $age_moyen, annee_recente \rightarrow G_2$.

La dernière étape consiste à calculer à quel point chacune des règles $r = A_{i1}^{j1}, \dots, A_{ik}^{jk} \rightarrow G_u$ est vraie et prédit donc le groupe considéré. Pour ce faire, nous utilisons la mesure suivante :

$$conf(r) = \frac{supp(A_{i1}^{j1}, \dots, A_{ik}^{jk}, G_u)}{supp(A_{i1}^{j1}, \dots, A_{ik}^{jk})}$$

Seules les règles ayant une confiance forte sont

conservées. Le seuil minimal est donné par l'administrateur du système de recommandation.

Nous disposons ainsi d'un ensemble de bases de règles qu'il s'agit d'appliquer quand un nouvel exemple (par exemple un nouvel utilisateur) souhaite recevoir une recommandation.

On note $R_u = \{(r_i^u, s_i^u, c_i^u, P_i^u)\}$ l'ensemble de règles de la classe G_u où chaque règle r_i^u est donc décrite par son support s_i^u , sa confiance c_i^u et les objets à recommander P_i^u .

3.3 Classification d'un nouvel exemple, recommandation

Face à un nouvel exemple L_{new} décrit selon les attributs de la base de données DB considérée, le système que nous proposons procède selon les étapes suivantes :

(1) calculer la pertinence de chaque groupe G_u avec la formule suivante scannant toutes les règles $r \in R_u$ avec $r = A_{i1}^{j1}, \dots, A_{ik}^{jk} \rightarrow G_u$:

$$pertinence =$$

$$\sum_{r \in R_u} \top_n(\mu_{A_{i1}^{j1}}(L_{new}[A_{i1}]), \dots, \mu_{A_{ik}^{jk}}(L_{new}[A_{ik}]))$$

(2) retrouver les t groupes ayant la plus forte pertinence (nous fixons $t = 3$ dans nos expérimentations)

(3) retrouver pour chacun des t meilleurs groupes les p produits les plus fréquents pour les recommander à l'utilisateur.

Les produits recommandés à l'utilisateur sont donc ceux qui plaisent le plus dans les groupes étant le plus proches de lui au sens des groupes utilisateurs obtenus par le clustering et retrouvés par le calcul ci-dessus.

4 Expérimentations

Dans cette section, nous synthétisons les expérimentations réalisées. Tout d'abord, nous analysons le comportement des algorithmes

de classification basés sur les règles d'association. Nous comparons ensuite l'algorithme CBA avec la méthode Fuzz-CBA proposée ici. Les expérimentations ont été réalisées sur deux bases de données : MovieLens¹ (rassemblant des avis sur des films) et Book Crossing [ZMKL05] (rassemblant des avis sur des livres). Pour les deux bases de données, WEKA [WF05] a été utilisé pour effectuer le prétraitement et la transformation des données. Au total, cinq bases de données ont été utilisées : une obtenue à partir de MovieLens et quatre obtenues à partir de BookCrossing (BookCrossing USA, BookCrossing World correspondant à tous les pays sauf les Etats-Unis, et ces deux bases avec dix valeurs distinctes seulement dans les attributs de l'auteur et du Pays/Etat). Dans toutes les expérimentations, nous avons utilisé la validation croisée (avec 10 sous-ensembles) pour l'évaluation des algorithmes.

4.1 Expérimentation avec classifieurs associatifs

Dans cette expérimentation, nous avons analysé la précision des classifieurs suivants : C4.5, BayesNet, CBA, CPAR et CMAR. Les deux premiers ont été exécutés à travers WEKA et les trois autres ont été obtenus à partir de LUKS-KDD². L'objectif principal de l'expérimentation était de comparer la précision des algorithmes utilisant les données recueillies à partir des systèmes de recommandation. Le tableau 2 rapporte les résultats obtenus après l'exécution des algorithmes mentionnés ci-dessus. Chaque ligne présente la précision obtenue pour chaque base de données. Rappelons que la précision d'une classe c_i est calculée par la formule suivante :

$$precision(C_i) = \frac{\text{nbre d'objets classés dans } C_i \text{ et appartenant à } C_i}{\text{nbre d'objets classés dans } C_i}$$

et permet d'estimer à quel point, quand un objet est classé dans une classe, il lui appartient

1. <http://www.grouplens.org/>

2. <http://www.csc.liv.ac.uk/~frans/KDD/Software/>

effectivement.

Les résultats ont révélé que les classifieurs associatifs, à l'exception de CMAR dans les données de BookCrossing, ont atteint des valeurs de précision similaires aux classifieurs traditionnels. Cette expérimentation nous permet de conclure que les méthodes de classification à base de règles d'association sont efficaces pour la recommandation. Notons que la mesure de la confiance est fixée à une haute valeur.

4.2 CBA vs. Fuzz-CBA

Nous présentons ici les résultats des expérimentations menées pour comparer CBA avec la version floue proposée ici (Fuzz-CBA), en comparant les deux approches de discrétisation (equi-width (ew) et equal-depth (ed)). Nous rapportons également le nombre de règles générées. Le nombre de règles est fixé entre 80 et 140, le support à 1%, la confiance à 75 ou 80% selon la taille de la base de données. Le tableau 3 montre les résultats obtenus, qui sont le plus souvent équivalents aux résultats de CBA. Notons que les mauvais résultats avec le equi-width pour la base MovieLens peuvent être expliqués par le fait que les intervalles de données générés de manière automatique avec cette méthode ne reflètent pas le contexte réel des données. Ainsi, nous pouvons conclure que la méthode développée dans cet article est efficace pour la recommandation, même si elle ne peut l'être que si la base de données utilisée a un nombre raisonnable d'enregistrements pour construire le modèle de classification.

Cependant, nous rappelons que les utilisateurs préfèrent qu'on ne leur recommande rien plutôt que quelque chose qui ne leur plaira pas et que, par conséquent, le problème de la diminution du nombre de faux positifs est primordial. Or le tableau 4 montre que notre méthode réduit ce nombre de faux positifs, ce qui confirme l'intérêt de notre approche.

Base de données	C4.5	BayesNet	CBA	CPAR	CMAR
MovieLens	82,88%	81,95%	81,4%	74,07%	85,16%
BCrossing World	80,21%	80,87%	79,47%	73,25%	70,43%
BCrossing World 10	79,98%	80,51%	80,52%	79,86%	41,41%
BCrossing USA	81,31%	80,23%	80,28%	78,15%	77,66%
BCrossing USA 10	80,82%	81,53%	81,56%	76,71%	69,59%

Tableau 2 – Comparaison entre les classifieurs

Base de données	CBA (ew)	Fuzz-CBA(ew)	CBA (ed)	Fuzz-CBA (ed)
MovieLens	47,01% - 44,4R	17,01% - 16,4R	82,1% - 85R	82,22% - 89,7R
BCrossing World	79,77% - 85,6R	80,4% - 80R	79,92% - 94,1R	81,04% - 135,6R
BCrossing World 10	79,65% - 91,5R	75,24% - 84,1R	79,04% - 105,6R	70,78% - 117,8R
BCrossing USA	78,2% - 72,9R	78,19% - 73,9R	78,65% - 69,2R	80,1% - 102,3R
BCrossing USA 10	79,29% - 93,4R	78,82% - 88,9R	79,68% - 92,1R	78,62% - 139R

Tableau 3 – Comparaison entre CBA et Fuzz-CBA : Précision obtenue et nombre de règles (R)

Base de données	Bayesnet	C4.5	CBA	Fuzz-CBA
MovieLens	47.4%	42.6%	33.22%	32.08%
BCrossing World	45.15%	39.9%	23.03%	31.89%
BCrossing World 10	40.3%	42.45%	33.83%	32.88%
BCrossing USA	47.45%	41.55%	22.45%	20.43%
BCrossing USA 10	48.25%	44%	34.66%	28.63%

Tableau 4 – Taux de faux positifs

5 Conclusion

Dans cet article, nous proposons une nouvelle approche de recommandation reposant sur la fouille de données floue en créant une base de règles floues. Les expérimentations menées démontrent l'intérêt de cette approche, notamment en ce qui concerne le taux de faux positifs qui correspondent à des recommandations qui n'ont pas satisfait l'utilisateur.

L'originalité du système proposé réside dans l'utilisation de méthodes de classification à base de règles d'association, et de règles d'association floues en particulier, pour la recommandation, ce qui n'existe pas dans la littérature. L'utilisation de la logique floue permet de mieux traiter les bases de données qui contiennent des attributs numériques discrétisés de manière stricte dans les approches de classification à partir de règles d'association. De plus, notre approche permet de ne pas cantonner un utilisateur à un seul groupe mais de lui ouvrir des recommandations liées à plusieurs

groupes proches. Ceci réduit l'un des risques de la recommandation : le "grey sheep problem" [CGM⁺99] qui se produit lorsqu'un utilisateur atypique se voit recommander les objets les plus populaires du moment en dépit des particularités de son profil par rapport à la *masse* des utilisateurs (par exemple quand un film très populaire en vient à être recommandé à tout le monde).

De nombreuses perspectives sont associées à ce travail, notamment pour étudier les différentes manières d'extraire les règles floues (e.g. création des partitions floues), de les combiner, et de calculer les mesures de support et de confiance.

Références

- [AIS93] Rakesh Agrawal, Tomasz Imielinski, and Arun Swami. Mining association rules between sets of items in large databases. In Peter Buneman and Sushil Jajodia, editors, *Proceedings of the*

1993 *ACM SIGMOD International Conference on Management of Data*, pages 207–216, Washington, D.C., 26–28 1993.

- [BHK98] John S. Breese, David Heckerman, and Carl Kadie. Empirical analysis of predictive algorithms for collaborative filtering. pages 43–52, 1998.
- [CA97] K.C. Chan and W.-H. Au. Mining fuzzy association rules. In *CIKM '97 : Proceedings of the sixth international conference on Information and knowledge management*, pages 209–215. ACM, 1997.
- [CGM⁺99] M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes, and M. Sartin. Combining content-based and collaborative filters in an online newspaper, 1999.
- [CKLT03] Kwok-Wai Cheung, James T. Kwok, Martin H. Law, and Kwok-Ching Tsui. Mining customer product ratings for personalized marketing. *Decis. Support Syst.*, 35(2) :231–243, 2003.
- [CLL05] Yukun Cao, Yunfeng Li, and Xiaofeng Liao. Applying fuzzy logic to recommend consumer electronics. In *ICDCIT*, pages 278–289, 2005.
- [CLZ05] Frans Coenen, Paul H. Leng, and Lu Zhang. Threshold tuning for improved classification association rule mining. In *PAKDD*, pages 216–225, 2005.
- [CZC⁺03] Michael Chau, Daniel Zeng, Hsin-chun Chen, Michael Huang, and David Hendriawan. Design and evaluation of a multi-agent collaborative web mining system. *Decis. Support Syst.*, 35(1) :167–183, April 2003.
- [dCFLH08] Luis M. de Campos, Juan M. Fernández-Luna, and Juan F. Huete. A collaborative recommender system based on probabilistic inference from fuzzy observations. *Fuzzy Sets Syst.*, 159(12) :1554–1576, 2008.
- [DHP06] Didier Dubois, Eyke Hüllermeier, and Henri Prade. Fuzzy methods for case-based recommendation and decision support. *J. Intell. Inf. Syst.*, 27(2) :95–115, 2006.
- [DP03] Didier Dubois and Henri Prade. A note on quality measures for fuzzy association rules. In *Proceedings IFSA'03, 10th International Fuzzy Systems Association World Congress, number 2715 in Lecture Notes in Artificial Intelligence*, page 677. Springer-Verlag, 2003.
- [LHM98] Bing Liu, Wynne Hsu, and Yiming Ma. Integrating classification and association rule mining. In *Knowledge Discovery and Data Mining*, pages 80–86, 1998.
- [LHP01] Wenmin Li, Jiawei Han, and Jian Pei. CMAR : Accurate and efficient classification based on multiple class-association rules. In *ICDM*, pages 369–376, 2001.
- [LKR01] Chi-Hoon Lee, Y.-H. Kim, and Phill-Kyu Rhee. Web personalization expert with combining collaborative filtering and association rule mining technique. *Expert Systems and Applications.*, 21(3) :131–137, 2001.
- [MGPL04] María N. Moreno, Francisco J. García, M. José Polo, and Vivian F. López. *Using Association Analysis of Web Data in Recommender Systems*, volume 3182/2004, chapter E-Commerce and Web Technologies, pages 11–20. Springer, 2004.
- [NFJK99] Olfa Nasraoui, Hichem Frigui, Anupam Joshi, and Raghu Krish-

- napuram. Mining web access logs using relational competitive fuzzy clustering. In *In Proceedings of the Eight International Fuzzy Systems Association World Congress*, 1999.
- [SA96] Ramakrishnan Srikant and Rakesh Agrawal. Mining quantitative association rules in large relational tables. In *Proceedings of the ACM SIGMOD International Conference On Management of Data*, pages 1–12, 1996.
- [SKKR00] Badrul M. Sarwar, George Karypis, Joseph A. Konstan, and John Riedl. Analysis of recommendation algorithms for e-commerce. In *ACM Conference on Electronic Commerce*, pages 158–167, 2000.
- [SKKR01] Badrul M. Sarwar, George Karypis, Joseph A. Konstan, and John Reidl. Item-based collaborative filtering recommendation algorithms. In *World Wide Web*, pages 285–295, 2001.
- [TCP05] F. Thabtah, P. Cowling, and Y. Peng. Mcar : multi-class classification based on association rule. In *AICCSA '05 : Proceedings of the ACS/IEEE 2005 International Conference on Computer Systems and Applications*, pages 33–I, Washington, DC, USA, 2005. IEEE Computer Society.
- [WF05] Ian H. Witten and Eibe Frank. *Data Mining : Practical machine learning tools and techniques*. 2nd edition, 2005.
- [Yag03] Ronald R. Yager. Fuzzy logic methods in recommender systems. *Fuzzy Sets Syst.*, 136(2) :133–149, 2003.
- [YH03] X. Yin and J. Han. Cpar : Classification based on predictive association rules. In *SIAM International Conference on Data Mining (SDM'03)*, pages 331–335, 2003.
- [ZMKL05] Cai-Nicolas Ziegler, Sean M. McNee, Joseph A. Konstan, and Georg Lausen. Improving recommendation lists through topic diversification. In *14th International World Wide Web Conference (WWW '05)*, 2005.