

---

# Diversité de recommandations

## Application à une plateforme de blogs et évaluation

Laurent Candillier\* — Max Chevalier\*\* — Damien Dudognon\*,\*\*  
— Josiane Mothe\*\*,\*\*\*

\* *Ebuzzing - OverBlog, 1 avenue Jean Rieux, F-31500 Toulouse*  
*prénom.nom@ebuzzing.com*

\*\* *IRIT, UMR5505, CNRS, Université de Toulouse*  
*118 rte de Narbonne, F-31062 Toulouse cedex 09*  
*Prénom.Nom@irit.fr*

\*\*\* *IUFM, École interne Université Toulouse le Mirail*

---

*RÉSUMÉ. Les systèmes de recommandations (SR) ont pour objectif de proposer automatiquement à l'utilisateur des objets en relation avec ses intérêts. Dans le contexte de la recherche documentaire, les intérêts de l'utilisateur peuvent être modélisés à partir des contenus des documents visités ou des actions réalisées. Pour tendre vers des recommandations plus pertinentes, nous proposons un modèle de SR qui construit une liste de recommandations répondant à un large spectre d'intérêts potentiels. L'originalité de notre modèle est qu'il repose sur la notion de diversité, obtenue en agrégeant différentes mesures d'intérêt pour construire la liste de recommandations finale. Nous définissons également un protocole permettant d'évaluer l'intérêt de ces recommandations. Nous présentons enfin les résultats obtenus par notre SR basé sur la diversité dans le cadre de la recommandation de billets de blogs.*

*ABSTRACT. Recommender Systems (RS) aim at providing objects related to user's interests. In Information Retrieval, user's interests can be modeled from the visited content and actions. To obtain more relevant recommendations for each user, we propose a SR model that builds a list of recommendations aiming at covering a large range of interests. The SR model we propose is based on the use of different interest measures and of an aggregation function to build the final set of recommendations. In this paper, we also evaluate this model on a blog platform.*

*MOTS-CLÉS : Systèmes de recommandations, diversité, mesure d'intérêt, plateforme de contenus*  
*KEYWORDS: Recommender systems, diversity, interest measure, content platform*

---

## 1. Introduction

Le Web est une source considérable de contenus qui ne cesse d'être enrichie par de nouvelles contributions des usagers. A titre d'exemple, sur *OverBlog*<sup>1</sup>, le nombre d'articles publiés est passé de 7,2 millions en 2010 à 12,5 millions en 2011.

L'enjeu est d'aider l'utilisateur à accéder à l'information pertinente noyée dans la masse. Les systèmes de recommandations (SR) sont une solution pour répondre à ce problème. A titre d'exemple, ils peuvent fournir à l'utilisateur des documents potentiellement pertinents en lien avec le document qu'il consulte. Nous proposons un modèle de SR qui produit une liste de recommandations diversifiée permettant de répondre à différents intérêts potentiels de l'utilisateur. Ce modèle est évalué dans le cadre d'une application réelle.

Nous abordons dans la Section 2 les SR et les différentes approches orientées diversité. Dans la Section 3, nous présentons notre modèle de SR. Enfin, nous explicitons en Section 4 le protocole expérimental mis en place pour valider notre proposition ainsi que les résultats obtenus.

## 2. Systèmes de recommandations d'information et diversité

Les systèmes de recommandations (SR) sont définis comme étant "des outils logiciels et des techniques qui suggèrent aux usagers des éléments utiles" (Ricci *et al.*, 2011). Afin d'identifier les informations à recommander, différentes stratégies ont été proposées dans la littérature (Malone *et al.*, 1987), (Candillier *et al.*, 2009) : le filtrage basé sur le contenu, le filtrage collaboratif et le filtrage hybride.

Dans cet article, nous ferons référence à la notion de "mesure d'intérêt" pour caractériser les différentes mesures permettant d'identifier les informations potentiellement pertinentes à recommander pour l'utilisateur. Ainsi une mesure d'intérêt pourra correspondre à un score de similarité, à un rang ou n'importe quelle autre valeur permettant d'ordonner les informations. Pour répondre au fait qu'un utilisateur peut considérer une information comme pertinente par rapport à différents critères, chacune des mesures d'intérêt peut reposer sur une dimension caractérisant les informations telles que son contenu (Dudognon *et al.*, 2010)), sa popularité, l'organisation relative des documents (Cabanac *et al.*, 2007), les liens existants au travers d'un réseau d'auteurs (Mothe *et al.*, 2006).

Malgré la diversité de critères existant dans la littérature, la majorité des SR n'utilisent qu'une seule mesure d'intérêt pour produire des recommandations. En conséquence, ils proposent souvent des résultats très proches. Pourtant (Ziegler *et al.*, 2005) indiquent que des recommandations très similaires ne présentent que peu d'intérêt pour l'utilisateur. De plus (Santos *et al.*, 2010) ajoutent qu'en diversifiant les résultats, la probabilité de satisfaire les intérêts de l'utilisateur est maximisée.

---

1. <http://www.over-blog.com/>

L'objectif de la diversité dans les SR est double : d'une part réduire la redondance dans la liste de recommandations et d'autre part tenir compte des divers intérêts des usagers.

Plusieurs approches se sont intéressées à introduire de la diversité dans les recommandations. (Ziegler *et al.*, 2005) définissent une mesure de similarité intra-liste sur la base d'une diversité selon une seule mesure d'intérêt. D'autres approches proposent de combiner différentes mesures. Les approches de recommandation notamment multi-critères (Roy, 1996) permettent de calculer des scores d'utilité pour un objet en fonction d'un ensemble de critères. La difficulté de ces approches est qu'il est important d'obtenir la valeur des différents critères.

Une alternative à la combinaison des critères est d'employer indépendamment différentes mesures d'intérêt. *Amazon.com* propose par exemple plusieurs listes de recommandations différentes. Néanmoins, cette indépendance des mesures d'intérêt conduit parfois à une redondance de l'information.

Les techniques de fusion habituellement utilisées apportent une réponse partielle à cette problématique. (Schafer *et al.*, 2002) et (Jahrer *et al.*, 2010) proposent de fusionner plusieurs sources de recommandations et présentent un "Méta SR". Cependant, ces approches, comme *CombSUM* (Shaw *et al.*, 1994), tendent à favoriser les documents présents dans plusieurs listes au détriment des documents très pertinents pour une seule mesure. Ceci ne favorise donc pas nécessairement la diversité d'intérêts dans la liste de recommandations.

En conclusion, la diversité dans une liste de recommandations est soit difficile à apprécier (cas de la combinaison de mesures d'intérêt), soit souvent limitée à un aspect particulier (cas de la diversité de contenu). Nous avons souligné dans un travail précédent (Candillier *et al.*, 2012) que la notion de diversité d'intérêts est importante pour les usagers et que la diversité de contenu n'est pas le seul intérêt auquel il faut répondre. Dans cet article, nous proposons un modèle de SR original ayant pour objectif de recommander à l'utilisateur, pour un document visité, une liste d'informations diversifiées basée sur l'agrégation des résultats issus de plusieurs mesures d'intérêt et couvrant différentes dimensions des informations.

### **3. Système de recommandations basé sur l'agrégation de résultats diversifiés**

Les SR visent à proposer des informations complémentaires aux informations consultées. Nous proposons de définir un modèle de SR basé sur la diversité d'intérêts pour maximiser les chances de satisfaire les usagers.

Notre modèle respecte les contraintes qui correspondent à des exigences concrètes rencontrées dans les systèmes en ligne et qui sont présentées ici :

**Contrainte 1** *Le SR proposé n'utilise pas de modèle d'utilisateur. L'exploitation du modèle de recommandations est donc possible sur n'importe quelle plateforme de*

*contenus (en particulier sans identification). Il n'est donc pas possible d'avoir un retour (notes) sur les mesures et les informations recommandées.*

**Contrainte 2** *L'approche doit être générique afin de permettre l'intégration de n'importe quelle mesure d'intérêt.*

**Contrainte 3** *L'approche doit s'affranchir de toute supposition concernant les intérêts de l'utilisateur et l'importance qu'il attribue à chacun d'eux.*

**Contrainte 4** *Chaque information visitée par un utilisateur peut susciter des intérêts différents et variés. Ces intérêts peuvent être affinés dans le temps.*

Dans ce contexte, le principe retenu pour notre modèle favorisant la diversité des recommandations est l'utilisation conjointe de plusieurs mesures d'intérêt et d'une fonction d'agrégation permettant notamment de supprimer les doublons entre les différentes listes. Nous garantissons ainsi la diversité dans la liste de recommandations pour couvrir un panel plus large d'intérêts. Notre modèle comporte quatre composants principaux :

– les éléments en entrée du SR : il s'agit de la collection de documents qui constitue la source de contenus d'où sont extraites les recommandations et le document visité. Ces deux éléments sont les seules entrées du SR permettant ainsi de répondre aux contraintes 1 et 3, c'est-à-dire qu'aucune connaissance préalable de l'utilisateur n'est nécessaire ;

– les **mesures d'intérêt** : à partir de la collection et du document consulté, elles restituent un ensemble ordonné de documents à recommander. L'idée est de ne pas limiter les mesures d'intérêt pouvant être intégrées au SR. Le modèle peut en effet être étendu par de nouvelles mesures d'intérêt et satisfaire ainsi la contrainte 2 ;

– le **processus d'agrégation** : il combine les résultats issus des différentes mesures d'intérêt assurant ainsi la diversité des recommandations et la suppression des doublons. L'objectif poursuivi est de permettre d'implanter des mesures d'agrégation spécifiques pour chaque plateforme de contenus sur laquelle le SR peut être intégré. Ceci répond à la contrainte 3 ;

– un **processus d'apprentissage** : il analyse les recommandations consultées par les utilisateurs pour le document visité. Il permet au processus d'agrégation d'ajuster, dans la liste de recommandations finale, les proportions de résultats issus de chaque mesure d'intérêt. Ce dernier composant permet de répondre à la contrainte 4. Ce composant ne rentre pas dans le périmètre de cet article.

Nous avons implanté le modèle au sein de la plateforme de blogs *OverBlog* afin de recommander, suite à la visite d'un billet de blog, des billets potentiellement pertinents pour l'utilisateur. La section suivante présente l'expérimentation que nous avons menée ainsi que les résultats obtenus.

#### 4. Expérimentation

Cette évaluation porte sur l'implantation du modèle proposé sur la plateforme de blogs *OverBlog*. L'objectif est de vérifier l'intérêt des recommandations fournies par notre modèle par rapport à des mesures considérées indépendamment. La mesure d'évaluation que nous utilisons est le taux de clics sur les recommandations proposées par notre modèle. Il s'agit d'une approximation de la pertinence perçue par les usagers souvent retenue dans l'évaluation de systèmes en ligne (c'est à dire sans collection de référence) (Chapelle *et al.*, 2009). Nous présentons tout d'abord le protocole d'évaluation mis en place, puis les résultats obtenus.

Pour cette expérimentation, nous avons comparé les résultats de notre modèle par rapport aux mesures d'intérêt qui sont d'ores et déjà intégrées à *OverBlog*. Il s'agit des trois mesures suivantes (les deux premières sont des variantes de mesures de contenu) :

- **MoreLikeThis (mlt)** : recherche dans la collection les billets de blogs possédant également les termes les plus représentatifs du document visité (le contenu global des billets est analysé) ;
- **SearchSim (search)** : repose sur le contenu textuel des titres des billets de blogs uniquement. Cette mesure favorise les billets de blogs ayant des termes similaires aux termes du document visité ;
- **BlogArticle (blog)** : retourne des billets appartenant au même blog que le billet visité. Cette mesure ne tient pas du tout compte du contenu textuel.

Outre ces mesures, nous ajoutons une implantation simplifiée du modèle d'agrégation (*fused*). Il repose sur la sélection des deux résultats ayant le meilleur score pour chacune des trois mesures précédentes. Le résultat de cette agrégation offre donc une diversité des recommandations, et ce sur différentes dimensions (contenu, appartenance à un blog). Ce choix est motivé par le fait que le modèle est intégré à la plateforme en exploitation. Ainsi la complexité de calcul doit être limitée pour ne pas mettre en péril l'ensemble de la plateforme. Les recommandations sont présentées dans un bloc positionné à la suite d'un billet. Le nombre de recommandations est limité à 6 pour maintenir la qualité de service en respectant notamment un temps de réponse inférieur à la seconde. Pour limiter l'impact sur les ressources nécessaires pour le calcul des recommandations sur la plateforme en exploitation, les recommandations ont également été pré-calculées pour chaque billet.

Pour pouvoir comparer le taux de clics, nous avons décidé de diviser les blogs en quatre groupes en fonction de leur identifiant interne sur la plateforme. Un groupe de blogs est ainsi associé à une et une seule mesure (*mlt*, *search*, *blog* ou *fused*). Nous avons mis en place cette évaluation pour une durée d'un mois (août 2012). Enfin, pour éviter le biais de la sélection par l'utilisateur des premiers documents recommandés, les recommandations sont présentées dans un ordre aléatoire à chaque affichage.

La figure 1 présente les résultats obtenus au cours du mois d'août 2012 sur la plateforme *OverBlog*. Cette figure présente le taux de clics pour chacune des mesures utilisées. Durant cette période, le bloc de recommandations a été affiché 886 041 fois

et a conduit à 34475 clics, ce qui représente un taux de clics moyen est de 3.89%. Le tableau 1 présente la répartition des affichages du bloc de recommandations pour chaque mesure durant le mois d'août 2012, ainsi que leur taux de clics moyen. Nous constatons que les différentes mesures ont été présentées de manière assez équilibrées aux visiteurs de la plateforme. Nous pouvons également souligner le fait que les taux de clics moyens pour les trois mesures de la plateforme sont inférieurs à 4% alors que notre modèle conduit à un taux de clics supérieur (4.70%). Ceci démontre l'intérêt de l'approche d'agrégation par rapport aux mesures considérées séparément.

Mesure	Taux d'affichage	Taux de clics moyen
<i>mlt</i>	24.14%	3.55%
<i>search</i>	26.30%	3.95%
<i>blog</i>	23.76%	3.49%
<b><i>fused</i></b>	25.80%	<b>4.70%</b>

**Tableau 1.** Affichages et taux de clics moyen par mesure durant le mois d'août 2012.

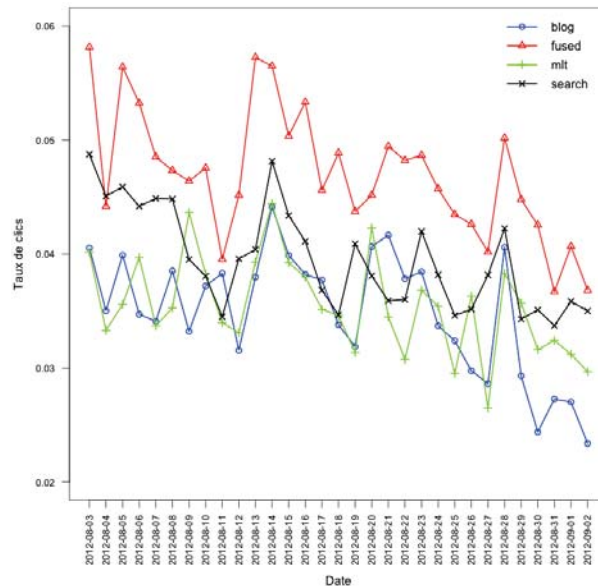
Pour affiner l'interprétation des résultats, nous présentons le taux de clics quotidien pour chaque mesure durant le mois d'août 2012 dans la figure 1 qui montre que :

- les visiteurs ont utilisé les recommandations du système et ce quelle que soit la mesure utilisée (avec un taux de clics minimum de 2.5% pour la mesure *blog*) ;
- le taux de clics journalier du modèle de fusion (*fused*) est globalement supérieur à celui de toutes les autres mesures.

Cette expérimentation souligne donc l'intérêt d'une approche d'agrégation des mesures d'intérêt par rapport à l'usage de ces mesures considérées indépendamment.

Afin de consolider les résultats obtenus, nous avons vérifié si ces interprétations étaient vérifiées au cours de la première quinzaine du mois de septembre 2012. Durant cette période, le bloc de recommandations a été présenté 584 884 fois. Le taux de clics moyen, indépendamment des mesures, s'élève à 3.46%. Ce taux est légèrement plus faible que pour le mois d'août (3.89%). La baisse globale du taux de clics moyen peut s'expliquer par le fait que les recommandations étaient similaires d'une visite sur l'autre dans le temps (car pré-calculées). L'attrait qu'elles suscitent était donc réduit pour les visiteurs qui revenaient sur le blog.

Dans cette nouvelle étude, l'analyse de la provenance des recommandations dans la mesure *fused*, qui sélectionne les deux meilleurs résultats de chacune des trois autres mesures, a montré l'importante variété des recommandations, notamment au travers de l'importance non négligeable de la mesure *blog* (24.02% des clics en moyenne). Nous avons constaté dans le même temps que les mesures de contenu *mlt* (28%) et *search* (47.98%) ont un impact plus important que la mesure *blog*. Nous avons vérifié que cette conclusion était identique sur tous les billets visités par les usagers : nous avons observé la grande différence de provenance des recommandations sur lesquelles l'utilisateur a cliqué d'un billet à l'autre.



**Figure 1.** Taux de clics quotidien au cours du mois d'août 2012

Ces expériences permettent de valider la contrainte 4.

Chaque billet visité pouvant susciter des intérêts différents, l'intégration d'un processus d'apprentissage est justifiée car elle permettrait de moduler l'agrégation des mesures d'intérêt de façon plus précise. Un exemple d'amélioration dans notre cas est de prendre en compte plus de documents issus des mesures les plus importantes pour un billet donné, c'est-à-dire ayant généré le plus de clics, tout en préservant au moins un résultat des autres mesures. De manière plus précise, pour un résultat de provenance de (*mlt* 20%, *search* 40%, *blog* 40%), et pour un bloc composé de 10 recommandations pour la mesure *fused*, 2 recommandations proviendraient de la mesure *mlt*, 4 recommandations de la mesure *blog* et 4 recommandations seraient issues de la mesure *search*.

## 5. Conclusion

Nous avons proposé dans cet article un modèle générique de SR visant à couvrir un nombre important d'intérêts des usagers. Nous ne posons aucune hypothèse préalable sur leurs intérêts et nous ne considérons que le document visité comme référence pour la phase de recommandation. La variété des intérêts que le document peut potentiellement susciter est traduite par l'emploi de plusieurs mesures d'intérêt. Ces dernières donnent lieu, après agrégation, à une liste unique et diversifiée de recommandations.

Cette diversité des intérêts couverts par la liste de recommandations permet de maximiser les chances de satisfaire l'utilisateur. L'implantation réalisée sur *OverBlog* nous a permis d'apprécier les contraintes techniques et computationnelles inhérentes à ce type de projet. Le prototype dans sa configuration actuelle supporte le passage à l'échelle.

Nos travaux s'orientent à présent vers le processus d'apprentissage, en exploitant notamment les historiques de navigation de la plateforme de blogs. Pour un document donné, les proportions seront par exemple ajustées en fonction du taux de clics des différentes mesures d'intérêt.

## 6. Bibliographie

- Cabanac G., Chevalier M., Chrisment C., Julien C., « An Original Usage-Based Metrics for Building a Unified View of Corporate Documents », *Database and Expert Systems Applications*, vol. 4653, p. 202-212, 2007.
- Candillier L., Chevalier M., Dudognon D., Mothe J., « Multiple Similarities for Diversity in Recommender Systems », *International Journal On Advances in Intelligent Systems*, vol. 5, n° 3 and 4, p. 234-246, 2012.
- Candillier L., Jack K., Fessant F., Meyer F., « State-of-the-Art Recommender Systems », *Collaborative and Social Information Retrieval and Access : Techniques for Improved User Modeling*, p. 1-22, 2009.
- Chapelle O., Zhang Y., « A dynamic bayesian network click model for web search ranking », *WWW*, p. 1-10, 2009.
- Dudognon D., Hubert G., Ralalason B. J. V., « ProxiGénéa : Une mesure de similarité conceptuelle », *VSSST*, 2010.
- Jahrer M., Töschler A., Legenstein R., « Combining predictions for accurate recommender systems », *ACM SIGKDD*, ACM, p. 693-702, 2010.
- Malone T. W., Grant K. R., Turbak F. A., Brobst S. A., Cohen M. D., « Intelligent information-sharing systems », *Communications of the ACM*, vol. 30, n° 5, p. 390-402, 1987.
- Mothe J., Chrisment C., Dkaki T., Dousset B., Karouach S., « Combining mining and visualization tools to discover the geographic structure of a domain. », *Computers, Environment and Urban Systems*, vol. 30, n° 4, p. 460-484, 2006.
- Ricci F., Rokach L., Shapira B., « Introduction to Recommender Systems Handbook », *Recommender Systems Handbook*, p. 1-35, 2011.
- Roy B., *Multicriteria Methodology for Decision Aiding*, Kluwer Academic, 1996.
- Santos R. L. T., Macdonald C., Ounis I., « Selectively diversifying web search results », *CIKM*, p. 1179-1188, 2010.
- Schafer J. B., Konstan J. A., Riedl J., « Meta-recommendation systems : user-controlled integration of diverse recommendations », *CIKM*, p. 43-51, 2002.
- Shaw J. A., Fox E. A., « Combination of Multiple Searches », *TREC*, p. 243-252, 1994.
- Ziegler C. N., McNeel S. M., Konstan J. A., Lausen G., « Improving recommendation lists through topic diversification », *WWW*, p. 22-32, 2005.