
COCofil

Une plateforme de filtrage collaboratif orientée vers la communauté

Nathalie Denos* — Catherine Berrut* — Lizbeth Gallardo-Lopez* — An-Te Nguyen*

* Laboratoire CLIPS-IMAG

385 rue de la Bibliothèque, BP 53, 38041 Grenoble cedex

{nathalie.denos, catherine.berrut, gallardo-lopez, an-te.nguyen}@imag.fr

RÉSUMÉ. Face à la quantité et la rapidité d'apparition de nouvelles informations au quotidien, l'utilisateur peut s'en remettre non seulement à des systèmes de filtrage d'information par le contenu mais aussi à des systèmes de filtrage collaboratif. Ainsi, il peut recevoir des recommandations grâce aux évaluations de la communauté des personnes partageant les mêmes centres d'intérêt. Cependant, l'utilisateur ne perçoit pas toujours favorablement le rapport coût-bénéfice que ce type de système apporte. Alors que la plupart des travaux dans la littérature focalisent sur les fonctions internes du système, nous présentons dans cet article une plateforme de filtrage collaboratif tournée vers l'innovation en matière de fonctionnalités interactives visant à améliorer ce rapport coût-bénéfice. Cela est fait notamment au travers de fonctionnalités orientées vers la notion de communauté. La possibilité de paramétrer personnellement et finement le système (confidentialité, choix des fonctionnalités intéressantes) est également assurée, car elle devient indispensable dans ce nouveau cadre.

ABSTRACT. With the explosive growth of the quantity of new information available everyday, more and more efficient tools are needed for information retrieval. In order to get relevant information, users can rely not only on cognitive filtering systems but also on collaborative filtering systems, also called recommender systems. The latter distribute recommendations based on the evaluations of a community of people with shared interests. However, one of the major problems of collaborative filtering is that the user does not always have the feeling to be a winner with respect to the cost-benefit ratio (evaluation efforts vs. recommendations benefits). Our research is not focused on internal aspects (improve prediction algorithms), as most related work are. Instead, we present a community-oriented collaborative filtering platform that provides numerous interactive functionalities aiming to improve the cost-benefit ratio. Also, the new platform allows users to personalize the production and the visualization of recommendations as well as his privacy choices.

MOTS-CLÉS : Filtrage collaboratif, plateforme, communauté, paramétrage.

KEYWORDS : Collaborative filtering, recommender systems; community, platform, configuration.

1. Introduction

1.1. *Qu'est-ce que le filtrage collaboratif ?*

Le filtrage collaboratif a pour principe d'exploiter les évaluations que des utilisateurs ont faites de certains documents, afin de recommander ces mêmes documents à d'autres utilisateurs. Les utilisateurs évaluent généralement les documents sur une échelle allant par exemple de 1 à 5.

Ce procédé de filtrage collaboratif est réalisé à l'aide d'un profil associé à chaque utilisateur. Ce profil est constitué de paires (document, évaluation), et il s'enrichit progressivement au fur et à mesure que l'utilisateur évalue de nouveaux documents. Il est important de noter que dans le principe du filtrage « purement collaboratif », le profil ne comporte pas d'informations sur le contenu des documents évalués : seul l'identificateur du document est conservé dans le profil. Le système met donc à jour les profils des utilisateurs au fur et à mesure qu'il reçoit leurs évaluations.

Etant donné ces profils, le système prédit l'intérêt qu'un document D peut présenter pour un utilisateur donné U en s'appuyant sur les évaluations que d'autres utilisateurs $\{U_i\}$ ont faites de ce même document, et en intégrant dans le calcul la proximité des utilisateurs U_i avec U via la comparaison de leurs profils. Lorsque l'intérêt prédit dépasse un certain seuil, le système recommande le document à l'utilisateur.

Ainsi, tous les utilisateurs peuvent tirer profit des évaluations des autres en recevant des recommandations pour lesquelles les utilisateurs les plus proches ont émis un jugement de valeur favorable, et cela sans que le système dispose d'un processus d'extraction du contenu des documents. De plus, les jugements de valeur des utilisateurs intègrent non seulement la dimension thématique mais aussi d'autres facteurs relatifs à la qualité des documents.

Le filtrage collaboratif diffère du filtrage cognitif, qui s'appuie sur le contenu des documents (thèmes abordés) pour les comparer à un profil lui-même constitué de thèmes. Ainsi, un utilisateur recevra des recommandations du système même s'il est le seul inscrit, pour peu qu'il ait décrit son profil en donnant un ensemble de thèmes qui l'intéressent. En revanche, il sera soumis à l'effet « entonnoir » car son profil évoluera naturellement par restriction progressive sur les thèmes recherchés. Il sera également victime d'un effet de masse, car ne bénéficiant pas des jugements de qualité que d'autres utilisateurs ont pu faire sur les documents qu'il reçoit, c'est lui qui devra procéder à l'écrémage des documents reçus, écrémage qui fait intervenir d'autres critères que celui de la thématique.

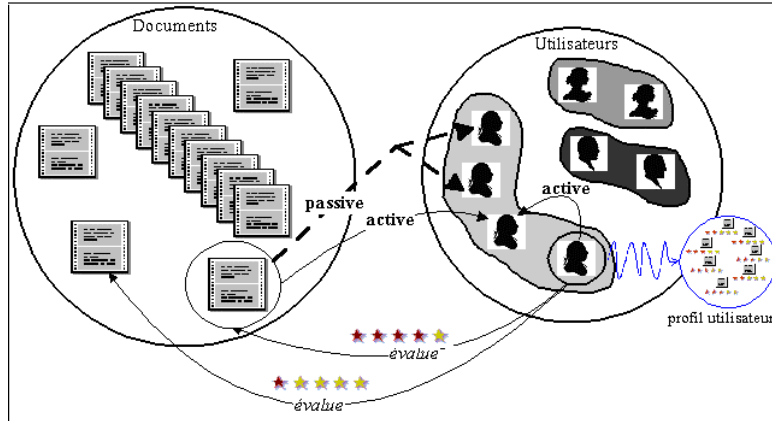


Figure 1. Schéma de filtrage collaboratif

Constatant les avantages et inconvénients de chacune de ces deux approches, on comprend que de nombreux systèmes reposent sur leur combinaison, ce qui en fait des systèmes de filtrage dits « hybrides ».

Au-delà de ces possibilités offertes par la combinaison des approches, certaines limitations du filtrage collaboratif peuvent être dépassées en développant de nouveaux principes d'interaction, comme par exemple le principe du « filtrage actif » qui contribue à limiter l'effet du « démarrage à froid », en permettant aux utilisateurs qui se connaissent de se recommander directement des documents, comme ils le feraient de manière naturelle entre collègues ou amis (Maltz et al., 1995).

1.2. Problématique abordée

Nous nous intéressons ici au filtrage « purement collaboratif », et à la manière de l'améliorer en intensifiant certains phénomènes grâce à de nouvelles fonctionnalités interactives.

Le filtrage collaboratif est un processus qui implique fortement les utilisateurs, puisque ses performances dépendent de la bonne utilisation du système : les utilisateurs doivent chacun émettre suffisamment d'évaluations pour dépasser le problème du démarrage à froid ; ils doivent être en suffisamment grand nombre pour atteindre une certaine masse critique au-delà de laquelle les calculs de prédiction prennent toute leur valeur ; les évaluations doivent concerner des ensembles de documents qui se recoupent au maximum, afin de permettre au système de comparer les profils ; etc.

Mais ce vœu pieux d'une bonne utilisation du système se heurte au problème du rapport coût-bénéfice dont les utilisateurs tiennent compte de façon consciente ou inconsciente : lorsqu'il évalue des documents, l'utilisateur se demande si ces efforts seront payés en retour à court ou à moyen terme. Les coûts assumés par l'utilisateur recouvrent l'effort d'évaluation des documents stricto sensu, mais aussi l'effort de compréhension de la tâche à accomplir, l'effort de prise en main de l'outil qui ne s'intègre pas toujours harmonieusement avec les outils auxquels l'utilisateur est habitué. Le faible bénéfice ressenti tient entre autres au fait que les recommandations ne véhiculent pas les informations permettant de comprendre les raisons de la prédiction, ni celles permettant de découvrir les membres de la communauté les plus proches.

Les travaux sur le filtrage collaboratif se focalisent généralement sur les fonctions internes (mise à jour des profils et calcul des prédictions) tant pour les efforts d'amélioration des techniques existantes, que pour l'évaluation des performances des systèmes (Breese et al., 1998 ; Fischer et al., 2000 ; Herlocker et al., 1999 ; Resnick et al., 1994). Il émerge actuellement des travaux alternatifs pour améliorer et mesurer les performances (Canny, 2002 ; Cöster et al., 2002 ; Jin et al., 2003).

Nous nous intéressons particulièrement à l'approche consistant à tirer profit de la présence des utilisateurs dans le système en s'intéressant à leur organisation en communautés. En effet, certaines limitations de l'approche collaborative peuvent être dépassées en amplifiant la connaissance réciproque des utilisateurs, et plus globalement leur conscience de la communauté à laquelle ils appartiennent.

Par exemple, le démarrage à froid peut être compensé par une période de « filtrage actif » où les utilisateurs qui se connaissent s'échangent des recommandations. Ou encore, lorsqu'un utilisateur n'est pas satisfait des recommandations que lui fournit le système, les approches classiques limitent les actions de l'utilisateur en présentant le processus de filtrage comme une boîte noire. Si le système lève le voile sur la formation des communautés d'utilisateurs, cela permet d'offrir à l'utilisateur des moyens pour rectifier son profil de façon plus radicale et éclairée.

1.3. Objectifs de COCoFil et solutions

L'objectif de la plateforme COCoFil (Community-Oriented Collaborative Filtering) que nous présentons dans cet article est d'intégrer toutes idées alternatives existantes ou nouvelles, s'articulant autour de la notion de communauté, afin d'en valider le principe et d'en évaluer l'utilité.

Pour atteindre cet objectif, nous avons fait un certain nombre de choix. Nous en dressons la liste ci-dessous en précisant chaque fois leurs apports :

- la conception du système est *modulaire* pour faciliter l'intégration de nouvelles fonctionnalités,
- le *maximum d'informations* circulant dans le système a été stocké de façon systématique afin de faciliter l'ajout de fonctionnalités utilisant des informations même rarement exploitées, ainsi que le traçage de paramètres pour l'évaluation,
- un *paramétrage* systématique des fonctionnalités a été réalisé à tous les niveaux pour permettre d'activer ou de désactiver une fonctionnalité à des fins d'évaluation, et aussi pour permettre à l'utilisateur de personnaliser l'interface du système (habitudes et besoins),
- des *fonctionnalités internes standard* (calcul de prédiction) ont été développées pour focaliser les recherches sur les fonctionnalités interactives,
- des *fonctionnalités innovantes orientées vers les communautés* d'utilisateurs ont été intégrées pour tirer un meilleur bénéfice des possibles relations entre utilisateurs du système.

C'est en cela que COCoFil n'est pas un simple système de filtrage collaboratif, mais une véritable plateforme ouverte d'essai et d'évaluation, actuellement particulièrement orientée vers la communauté. En effet elle intègre des fonctionnalités destinées à mieux exploiter la notion de communauté d'utilisateurs. Elle a déjà servi de plateforme pour concevoir et valider des fonctionnalités qui par la suite ont été intégrées de façon sélective dans un portail communautaire destiné aux chercheurs en physique des hautes énergies (projet européen TIPS¹).

2. Description globale de la plateforme COCoFil

La plateforme COCoFil comporte, outre l'espace destiné à l'utilisateur, un espace destiné à l'administration du système. Cet espace « administrateur » permet de gérer les divers paramètres du système, et d'intégrer des tableaux de bord destinés à surveiller l'état du système pendant son exploitation.

L'espace « utilisateur » est structuré selon les grandes familles d'activités que l'on peut y pratiquer. Ces activités sont conçues pour reprendre et étendre les fonctionnalités habituelles des outils courants. De plus, des passerelles sont établies entre les activités pour améliorer l'accessibilité des fonctionnalités.

La Figure 2. schématise les principales fonctionnalités de COCoFil, en représentant, dans la partie supérieure, les *actions* (ovales) que l'utilisateur U peut faire avec ses propres *données* (rectangles en gras sur le dessus de la pile) ou les *données des autres* utilisateurs (autres rectangles de la pile). Les fonctionnalités internes du système sont situées dans la partie inférieure de la figure. Les passerelles sont représentées par des flèches en pointillés.

¹ <http://tips.sissa.it> ; <http://torii.info>

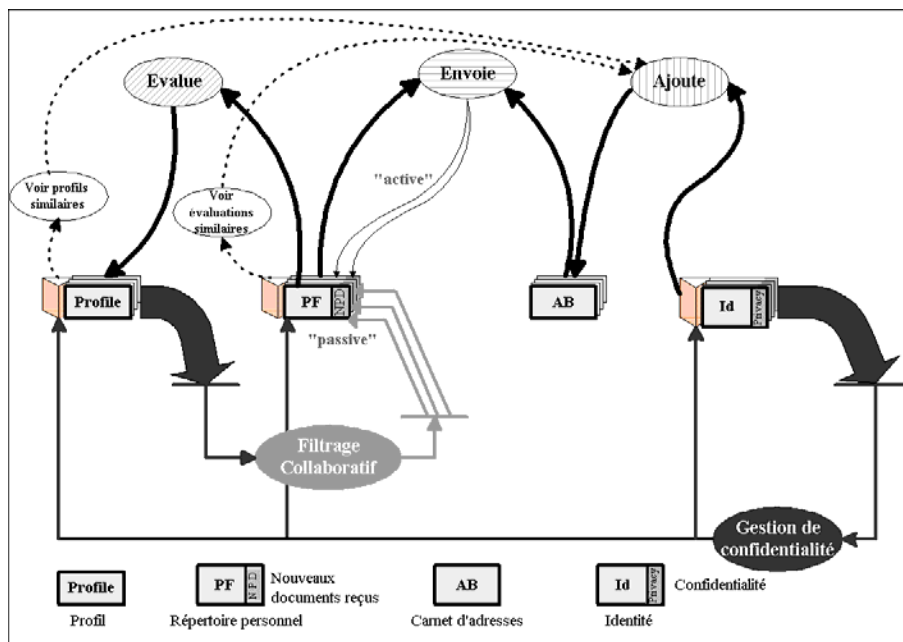


Figure 2. Schéma des principales fonctionnalités « utilisateur » et des passerelles

On trouve ainsi un répertoire personnel (PF pour « Personal Folder »), qui contient les documents que l'utilisateur y a mis, ainsi qu'un dossier réservé aux nouveaux documents recommandés (NPD pour « New Pushed Documents »). Ce répertoire personnel est enrichi de la possibilité de donner une note (Eval) à chaque document.

Par ailleurs, un carnet d'adresses (AB pour « Address Book ») permet à chacun de répertorier (Ajout) les adresses électroniques des utilisateurs à qui il souhaite envoyer des documents (filtrage actif).

Chaque utilisateur peut décider des informations qu'il souhaite divulguer ou non, qu'il s'agisse d'informations relatives à son identité, ou même des évaluations qu'il fait pour chaque document. Le système propage ces contraintes de confidentialité (Privacy), limitant ainsi certaines fonctionnalités de nature à faire se connaître les utilisateurs entre eux.

Ces dernières fonctionnalités, particulièrement orientées vers la communauté, permettent notamment de connaître les utilisateurs qui ont attribué une note semblable à U pour un document donné, ou ceux qui d'une manière plus globale ont un profil similaire à celui de l'utilisateur U.

Lorsqu'on consulte les utilisateurs ayant fait une évaluation similaire ou ayant un profil similaire, est offerte la possibilité d'ajouter directement les utilisateurs en question au carnet d'adresses. Cela constitue un exemple de passerelle entre fonctionnalités, qui contribue à concrétiser la notion de communauté.

Une autre caractéristique de cette plateforme est la richesse des données qu'elle mémorise. A titre d'exemple, l'historique du contenu des répertoires personnels est mémorisé, de façon à permettre d'intégrer une fonctionnalité de type « évaluation implicite » (on interprète les actions sur les documents du répertoire personnel comme des évaluations de ces documents).

Nous voyons donc, avec cette vue synthétique des caractéristiques originales de COCoFil, que cette plateforme comporte des fonctionnalités rares, voire inédites dans les systèmes de filtrage collaboratif existants. Ces fonctionnalités sont, pour la plupart, tournées vers la notion de communauté, et ont pour but de permettre aux utilisateurs d'intervenir dans le processus de recommandation directement et de se découvrir entre membres d'une même communauté. Le lecteur intéressé trouvera à l'URL <http://tips.imag.fr/cocofil/> une comparaison plus technique et détaillée des fonctionnalités de COCoFil avec celles de deux systèmes de filtrage collaboratif bien connus : celui d'Amazon et celui de MovieLens.

3. Fonctionnalités d'une plateforme de filtrage collaboratif orientée vers la communauté

Dans cette partie, nous présentons plus en détail les fonctionnalités de la plateforme de filtrage collaboratif COCoFil.

3.1. Identification et confidentialité

Afin d'obtenir un compte d'accès, l'utilisateur doit fournir certaines informations personnelles telles que nom, prénom, adresse électronique, au moment de l'inscription volontaire.

Toutes ces informations sont utiles non seulement pour l'identification mais aussi pour la crédibilité des recommandations dans le filtrage collaboratif actif. En effet, l'identification et au moins la pseudonymie enrichissent la perception des recommandations circulant dans la communauté. Par exemple, il semble naturel qu'un utilisateur s'intéresse plus aux suggestions de personnes identifiées qui lui ont souvent adressé des recommandations pertinentes par le passé ; de même, il aura sans doute tendance à négliger les recommandations anonymes.

Cela ne veut pourtant pas dire que la confidentialité n'est pas assurée dans COCoFil. L'utilisateur a toujours la possibilité d'indiquer qu'il souhaite

l'anonymat : ses informations personnelles ne seront alors pas divulguées. Aussi, il peut choisir de conserver l'anonymat pour ses évaluations (voir Figure 3), sachant qu'il pourra très simplement, lors d'une évaluation particulière, choisir l'option inverse.

Ces possibilités de configuration de la confidentialité sont d'autant plus importantes que la plateforme intègre des fonctionnalités destinées à rendre plus perceptible la communauté, comme nous le verrons plus loin.

Dans le projet TIPS, on a constaté que la crainte de ne pas maîtriser la confidentialité de ses propres évaluations jouait un rôle dissuasif pour certains utilisateurs. Cela n'empêchera pas de mettre éventuellement en place des mesures d'encouragement à quitter l'anonymat, si l'on estime que la performance du système peut en être améliorée.

Options for the evaluation	
Tick options you want become default	
Add to Personal Folder	<input type="checkbox"/>
Anonymous evaluation	<input type="checkbox"/>
Not include this document for social filtering information	<input type="checkbox"/>
Set up of My identity	
Tick features you accept to transmit to others persons	
Middle name	<input checked="" type="checkbox"/>
First name	<input checked="" type="checkbox"/>
E-mail	<input checked="" type="checkbox"/>
Login	<input checked="" type="checkbox"/>
Organization	<input checked="" type="checkbox"/>
Fields of research	<input checked="" type="checkbox"/>
Topics of interest	<input checked="" type="checkbox"/>
Expert competence	<input type="checkbox"/>
Anonymousness	<input type="checkbox"/>

Figure 3. Paramètres ajustables pour chaque évaluation de document

3.2. Filtrage collaboratif passif

Comme nous l'avons indiqué dans l'introduction, bien que des travaux récents dans la littérature focalisent sur le calcul de corrélation et de prédiction pour le filtrage collaboratif passif, nous ne nous intéressons pas à cet aspect. Ainsi, notre plateforme adopte des techniques classiques de filtrage collaboratif passif grâce à une approche modulaire.

Dans ce sens, le moteur de filtrage n'est plus considéré que comme une boîte noire, et peut se déclencher soit périodiquement, soit après chaque évaluation, ou encore à la demande de l'utilisateur ou de l'administrateur.

Une interface a été définie afin de permettre, autant que possible, de modifier la technique de filtrage passif sans que cela ait d'impact majeur sur le reste des fonctionnalités de la plateforme.

Enfin, il est prévu qu'un utilisateur puisse évaluer un document qui ne présente pas d'intérêt pour son profil de filtrage ; cela peut arriver par exemple dans le cadre d'une activité de relecture académique, faite sur un domaine de compétence qui n'est pas ou plus le domaine d'intérêt actuel de l'utilisateur. L'utilisateur peut indiquer qu'une telle évaluation ne doit pas intervenir dans son profil collaboratif (voir encadré Figure 3).

3.3. Filtrage collaboratif actif

Lorsqu'un utilisateur trouve des documents plus ou moins intéressants pour certains autres utilisateurs qu'il connaît, il peut les leur recommander : c'est le filtrage actif, qui est encouragé dans la plateforme avec un bouton « Send » accessible depuis la vue détaillée de tout document.

Nous verrons plus loin que la plateforme intègre des outils complémentaires à celui-ci, afin de permettre aux utilisateurs de faire du filtrage actif dans les meilleures conditions : effort limité via un groupe de destinataires par défaut (section 3.5), possibilité d'enrichir son carnet d'adresses grâce aux passerelles entre le carnet d'adresses et les fonctionnalités de perception des autres (section 3.4), etc..

3.4. Perception des autres

Chaque utilisateur peut consulter la liste des personnes qui ont des *profils similaires* au sien, et ainsi connaître de nouvelles personnes ayant les mêmes centres d'intérêt que lui (voir Figure 4). On trouve aussi la fonctionnalité de type « passerelle », qui permet à partir de cette liste, de compléter simplement son carnet d'adresses. On peut espérer que cette fonctionnalité permet à l'utilisateur de mieux cibler ses recommandations directes (filtrage actif).

Bien sûr, cette liste est soumise au filtre des choix d'anonymat faits par les utilisateurs ! Une personne ayant demandé l'anonymat n'apparaîtra dans aucune liste. Pourtant, son profil sera toujours prise en compte dans le calcul de prédiction du moteur.

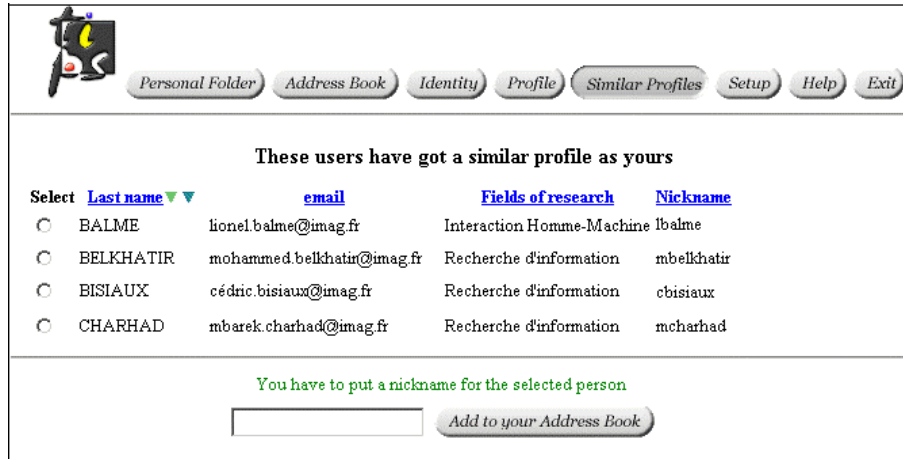


Figure 4. Liste des personnes ayant des profils similaires

Dans le même esprit de perception de la communauté, un utilisateur peut connaître la *liste des personnes ayant fait la même évaluation* que lui pour un document particulier (voir Figure 5). Cette fonctionnalité contribue également à de meilleures recommandations par filtrage actif. En effet, si un utilisateur constate qu'un nouveau document est très proche d'un autre qu'il a déjà jugé pertinent, il peut décider, au lieu de le recommander à sa liste de destinataires par défaut, de s'inspirer de la liste des personnes ayant fait la même évaluation que lui sur l'ancien document.

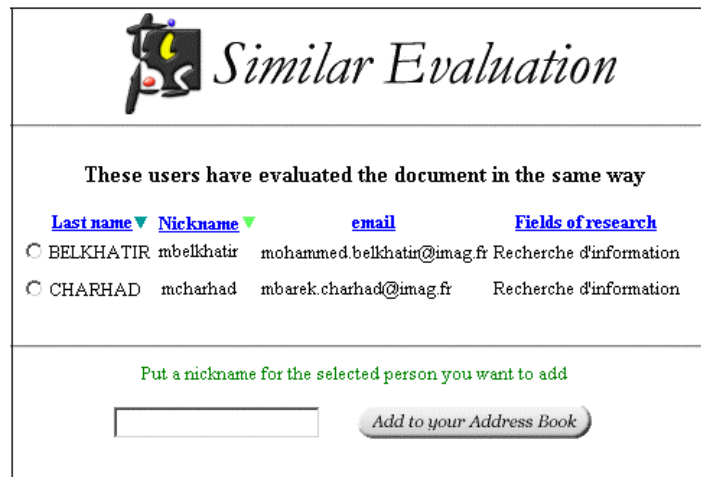


Figure 5. Liste des personnes ayant fait la même évaluation du document

A nouveau, cette fonctionnalité est soumise aux règles d'anonymat, non seulement cette fois pour ce qui est de l'identité des personnes, mais aussi par rapport à l'anonymat des évaluations : les évaluations anonymes n'apparaîtront pas.

3.5. Carnet d'adresses

Chaque utilisateur de la plateforme COCoFil possède un *carnet d'adresses* qui enregistre des informations utiles (nom, prénom, surnom, adresse électronique, etc.) des personnes « intéressantes », et évidemment « connues », avec lesquelles il souhaite échanger régulièrement des recommandations. L'utilisateur peut ajouter, modifier et supprimer des contacts dans son carnet (voir Figure 6).

To order the contacts' persons click on links in the heading

Select	Nickname	Last name	First name	Organisation	Email	Fields of research	Topics of interest
<input type="checkbox"/>	mcharhad	CHARHAD	Mbarek	CLIPS	mbarek.charhad@imag.fr	Recherche d'information	Indexation
<input type="checkbox"/>	ndenos	DENOS	Nathalie	CLIPS	nathalie.denos@imag.fr	Recherche d'information	Filtrage collaboratif
<input type="checkbox"/>	lkefi	KEFI	Leila	CLIPS	leila.kefi@imag.fr	Recherche d'information	Indexation

Click on groups' links to have more informations

[default](#)

[Thésards](#)

[Directrices](#)

You can act on the selected contacts

You can add contacts

Figure 6. Carnet d'adresses

En particulier, un carnet d'adresses dans COCoFil peut être organisé en groupes par son propriétaire afin de faciliter le filtrage collaboratif actif. Cette possibilité soulage considérablement la tâche d'envoi des recommandations du fait que l'utilisateur peut envoyer des recommandations à un groupe de personnes au lieu de le faire individuellement.

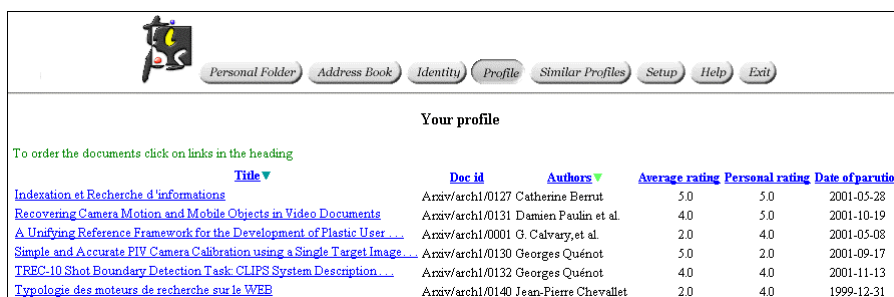
Il existe en outre un groupe spécial dans chaque carnet d'adresses, nommé « Default group » qui contient des destinataires par défaut permettant une

recommandation en un seul clic : l'effort à produire pour le filtrage actif est ainsi réduit au minimum pour ce groupe.

3.6. Consultation du profil

En général, un profil de filtrage collaboratif représente le centre d'intérêt de l'utilisateur via les évaluations d'objets (documents) faites par le passé. Un profil est donc une liste de tuples composés de l'identificateur de l'objet, la valeur de l'évaluation personnelle, et la date de cette évaluation.

La plateforme permet aux utilisateurs de consulter leur profil : chaque tuple est éventuellement complété par des informations supplémentaires explicatives, comme la prédiction du système et l'évaluation moyenne de la communauté (voir Figure 7).



Title ▼	Doc id	Authors ▼	Average rating	Personal rating	Date of parution
Indexation et Recherche d'informations	Arxiv/arch1/0127	Catherine Berrut	5.0	5.0	2001-05-28
Recovering Camera Motion and Mobile Objects in Video Documents	Arxiv/arch1/0131	Damien Paulin et al.	4.0	5.0	2001-10-19
A Unifying Reference Framework for the Development of Plastic User...	Arxiv/arch1/0001	G. Calvary, et al.	2.0	4.0	2001-05-08
Simple and Accurate PTZ Camera Calibration using a Single Target Image...	Arxiv/arch1/0130	Georges Quénot	5.0	2.0	2001-09-17
TRBC-10 Shot Boundary Detection Task: CLIFS System Description...	Arxiv/arch1/0132	Georges Quénot	4.0	4.0	2001-11-13
Typologie des moteurs de recherche sur le WEB	Arxiv/arch1/0140	Jean-Pierre Chevillet	2.0	4.0	1999-12-31

Figure 7. Profil utilisateur

Nous soulignons particulièrement notre approche de paramétrage qui permet à l'utilisateur un affichage flexible de son profil grâce à la possibilité de trier les évaluations selon plusieurs caractéristiques telles que titre, évaluation personnelle, évaluation moyenne de la communauté, chronologique, etc.

Par ailleurs COCoFil est très souple quant à la modification des profils par l'utilisateur : l'utilisateur peut non seulement consulter son profil, mais aussi réévaluer et même supprimer des évaluations de son profil.

3.7. Gestion personnelle des documents et recommandations

Dans notre plateforme, l'utilisateur a la possibilité d'organiser ses documents selon une structure arborescente de répertoires, qu'il s'agisse de ceux qui lui sont parvenus par recommandation, ou d'autres qu'il a intégré dans son espace personnel

dans l'idée, par exemple, de les évaluer plus tard. Chaque répertoire peut être trié selon plusieurs critères (voir Figure 8).

Chaque utilisateur possède un répertoire prédéfini « New Pushed Documents », qui ne peut être ni supprimé, ni renommé, et où le système place les recommandations. A chacune de ces recommandations, est associée l'indication de sa provenance (filtrage passif ou actif), ce qui constitue un début d'explication. L'utilisateur peut supprimer ou déplacer ces recommandations dans des sous-répertoires préalablement créés.

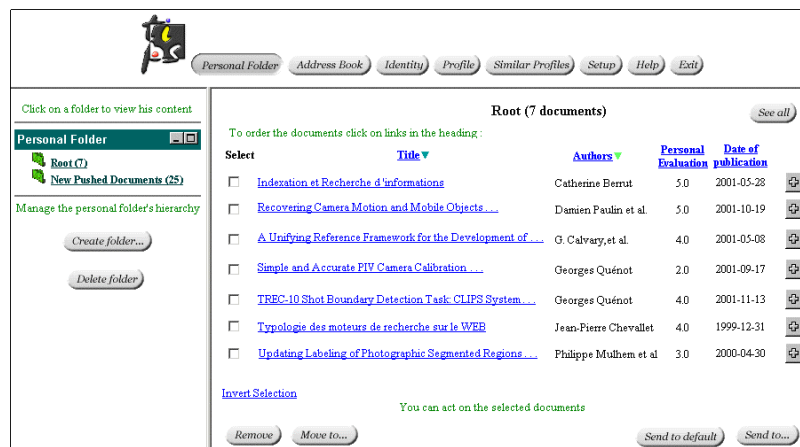


Figure 8. Gestion personnelle des recommandations

L'historique des mouvements des documents d'un répertoire à l'autre est conservé, ce qui permet d'envisager, le cas échéant, la prise en compte d'évaluations implicites : par exemple, on estimera qu'une recommandation jetée à la poubelle correspond à une évaluation très mauvaise, alors qu'une recommandation déplacée vers un répertoire créé personnellement peut être interprétée comme une bonne évaluation.

3.8. Paramétrage utilisateur

Une caractéristique importante de notre approche est le paramétrage du filtrage collaboratif. La plateforme COCoFil permet à l'utilisateur de configurer certains aspects de son utilisation :

- pour *la visualisation* : choisir la façon d'afficher les contenus dans ses répertoires personnels, son carnet d'adresses, son profil, et les listes de personnes de profils similaires et de personnes ayant fait la même évaluation, etc.

– pour *l'évaluation* : choisir le niveau de confidentialité par défaut pour les évaluations, ainsi qu'éventuellement un anonymat total ; s'ajoute à ces paramètres par défaut la possibilité d'ajuster le niveau de confidentialité au coup par coup lors des évaluations.

– pour *le filtrage collaboratif* : décider au coup par coup si une évaluation doit être prise en compte dans le profil ou non.

– pour *le niveau d'exploitation* : choisir le régime de débutant ou d'expert avec l'association des étiquettes explicatives aux boutons sur l'écran, la confirmation de l'exécution de certaines fonctionnalités, etc.

4. Exploitation de la plateforme

4.1. Terrain d'essai

La plateforme COCoFil a pour vocation de servir de terrain d'essai pour concevoir de véritables systèmes de filtrage collaboratif particuliers. Elle a donc vocation à intégrer des fonctionnalités très nombreuses, mais potentiellement « désactivables ». Par exemple, COCoFil a servi à maquetter les fonctionnalités interactives du service de filtrage collaboratif à intégrer dans le portail Torii (projet TIPS). Seul un petit nombre de fonctionnalités a été intégré, et cela relativement simplement grâce à la conception modulaire de la plateforme.

4.2. Ajout de nouvelles fonctionnalités

Par ailleurs, cette plateforme ouverte permet d'intégrer de nouvelles fonctionnalités dans la mesure où l'intégralité des données circulant dans le système est tracée dans les bases de données. Par exemple, le fait de conserver la trace de l'historique des déplacements d'un répertoire dans l'autre permet d'envisager assez simplement l'ajout de fonctionnalités d'évaluation implicite.

4.3. Suivi des performances

Dans ce même ordre d'idée, pour étudier le comportement du système pendant son exploitation, et à terme améliorer la qualité des performances du système, un tableau de bord a été développé et intégré dans la partie « Administrateur » de COCoFil [Gallardo-Lopez et al., 2003].

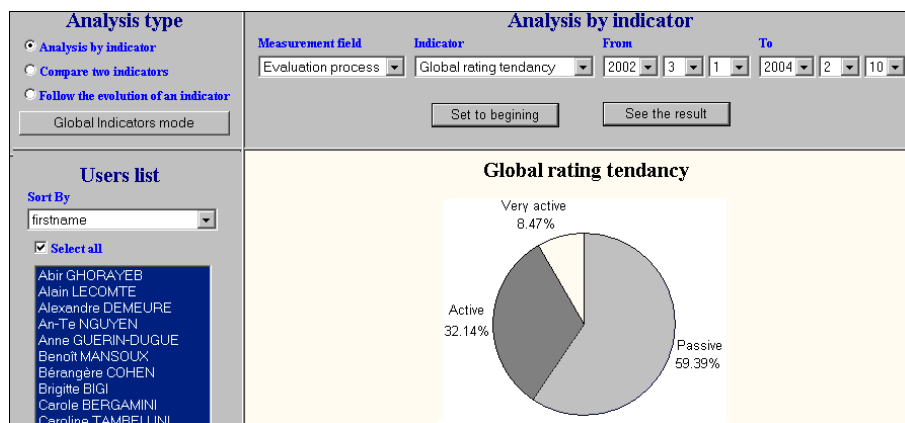


Figure 9. Tableau de bord - indicateur de tendance à évaluer

Ce tableau de bord contient plusieurs indicateurs rendant plus facile la tâche d'analyse de la performance actuelle du système, et permettant à l'administrateur (ou éventuellement à un processus automatique) de prendre de meilleures décisions à court et à long terme afin d'améliorer cette performance.

Par exemple, la classification des utilisateurs selon l'indicateur de « tendance à évaluer » fournit à l'administrateur une vue globale du processus d'évaluation du système tel qu'il se déroule actuellement (voir Figure 9). En constatant la proportion importante d'utilisateurs « passifs », l'administrateur peut soit mettre en question la précision des recommandations, soit mettre en œuvre des mesures destinées à encourager ces utilisateurs passifs à évaluer de plus en plus de documents.

5. Architecture de COCoFil

5.1. Architecture globale

En ce qui concerne les données, on trouve quatre bases de données dans la plateforme COCoFil (voir Figure 10).

La première, *Document*, est la plus grande, qui enregistre la collection de documents. Néanmoins, le processus d'alimentation de cette base de données à partir d'une source dynamique de documents n'est pas encore implémenté. On pense l'appuyer sur le protocole « OpenArchive » donnant accès aux archives ouvertes de prépublication, comme cela a été fait pour le portail Torii du projet TIPS.

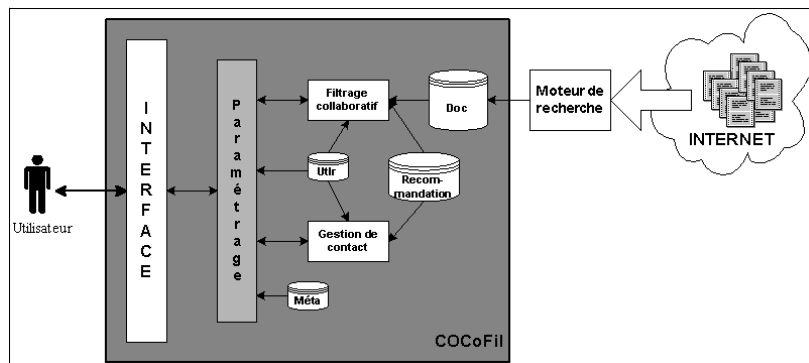


Figure 10. Architecture globale de COCoFil

La seconde base de données *Recommandation* comprend toutes les recommandations faites par le système, qu'elles aient fait l'objet d'évaluation ou non dans les répertoires personnels. Le fonctionnement des deux modules Gestion de contact, notamment la fonctionnalité d'envoi des recommandations par filtrage actif, et le filtrage collaboratif proprement dit, est décrit plus loin dans la même section. Il se fonde sur le contenu de cette même base de données.

Par ailleurs, les informations personnelles des utilisateurs, y compris des carnets d'adresses, sont stockées dans la base de données *Utilisateur*. Cette base de données sert d'une part à l'identification de l'utilisateur et d'autre part à la réalisation des filtres passif et actif.

Finalement, le paramétrage de la plateforme COCoFil est implémenté dans la base de données *Méta*. Dans la première version de COCoFil, cette base comporte une seule table dont chaque colonne caractérise un aspect de paramétrage comme l'affichage du répertoire personnel, anonymat des évaluations, profil similaire, carnet d'adresses, transmission des recommandations, niveau d'exploitation, etc. Nous envisageons déjà une conception et une implémentation plus générale de cette base de données afin de s'adapter à plusieurs contextes.

Les fonctionnalités de COCoFil sont organisées en trois modules : Filtrage collaboratif, Gestion de contact et Paramétrage (voir Figure 10).

Le module Filtrage collaboratif traite d'abord la gestion personnelle des recommandations (organiser, évaluer et visualiser). Il exploite ensuite les recommandations évaluées dans la base de données *Recommandation* pour effectuer le filtrage collaboratif passif. Et enfin, ce module procure également aux utilisateurs des listes de personnes ayant des profils similaires et des personnes ayant fait la même évaluation.

Quant au module Gestion de contacts, il assure l'identification dans la plateforme COCoFil, c'est-à-dire qu'il permet d'une part aux utilisateurs de saisir

leurs informations personnelles, et d'autre part au système d'identifier l'utilisateur lors de son accès. Grâce à ce module, l'utilisateur peut en outre organiser son carnet d'adresses et échanger des recommandations avec d'autres dans le cadre du filtrage collaboratif actif.

Finalement, le module Paramétrage joue le rôle d'une couche intermédiaire qui intervient essentiellement dans le fonctionnement des deux modules Filtrage collaboratif et Gestion de contact. Il agit sur la visualisation des recommandations, des profils, etc., et particulièrement sur la production et l'évaluation des recommandations afin de rendre le filtrage plus adaptatif au contexte d'utilisation.

5.2. Développement de COCoFil

La première version de COCoFil a été développée par l'équipe MRIM du laboratoire CLIPS-IMAG à Grenoble, dans l'environnement constitué par des logiciels gratuits, répandus, et multi plateformes tels que le serveur Web Apache, l'interpréteur TomCat, le système de gestion de bases de données MySQL, et le langage de programmation Java (voir Figure 11).

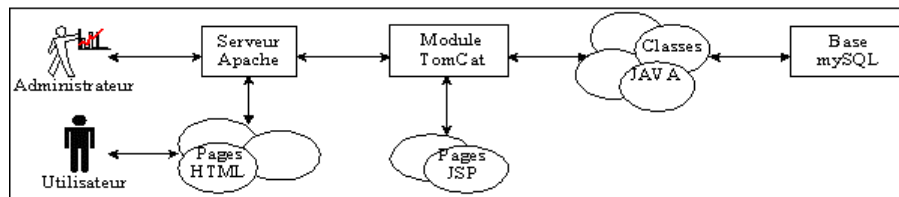


Figure 11. Fonctionnement du système

6. Conclusion et perspectives

Dans cet article, la plateforme de filtrage collaboratif orientée vers la communauté COCoFil est présentée d'abord de façon générale, et ensuite par le détail de ses fonctionnalités. A travers cette présentation, le lecteur perçoit d'une part l'approche modulaire qui facilite aussi bien l'ajout de nouvelles fonctionnalités que l'intégration avec d'autres systèmes, et d'autre part le paramétrage qui permet une utilisation plus efficace et adéquate via la personnalisation.

La plateforme COCoFil sert actuellement de terrain pour le développement d'un tableau de bord permettant de surveiller l'état du système en cours d'exploitation. L'objectif est d'intégrer des fonctionnalités interactives susceptibles d'améliorer la

qualité du service rendu par le système en cours d'exploitation, et d'évaluer l'effet produit avec le tableau de bord. Nous envisageons également d'étendre ses fonctionnalités afin d'intégrer les bénéfices de l'approche cognitive.

Remerciements

Les auteurs tiennent à remercier Delphine Million, Fleur Picollet et Jean-Gilles Plotto qui ont réalisé la majeure partie du développement de la première version de la plateforme COCoFil, et ont grandement contribué à sa conception et à sa documentation.

7. Bibliographie

- Breese J. S., Heckerman D., Kadie C., « Empirical Analysis of Predictive Algorithms for Collaborative Filtering », *Proceeding of the 14th Conference on Uncertainty In Artificial Intelligence (UAI'98)*, Jul 24-26, 1998, Madison, Wisconsin, USA.
- Canny J., « Collaborative filtering with privacy via factor analysis », *Proceedings of the 25th International Conference on Research and Development in Information Retrieval (SIGIR'02)*, Aug 11-15, 2002, Tampere, Finland.
- Cöster R., Svensson M., « Inverted file search algorithms for collaborative filtering », *Proceedings of the 25th International Conference on Research and Development in Information Retrieval (SIGIR'02)*, Aug 11-15, 2002, Tampere, Finland.
- Fisher D., Hildrum K., Hong J., Newman M., Thomas M., Vuduc R., « SWAMI : A Framework for Collaborative Filtering Algorithm Development and Evaluation », *Proceedings of the 23th Conference on Research and Development in IR (SIGIR'00)*, Jul 24-28, 2000, Athens, Greece.
- Gallardo-Lopez L., Berrut C., Denos N., « Une approche pour le contrôle de la qualité des Systèmes de Filtrage Collaboratif », *Manifestation de Jeunes Chercheurs STIC (MAJESTIC'03)*, Oct 29-31, 2003, Polytechnique de Marseille, France.
- Herlocker J. L., Konstan J. A., Brochers A., Riedl J., « An Algorithm Framework for Performing Collaborative Filtering », *Proceedings of the 22nd Conference on Research and Development in IR (SIGIR'99)*, Aug 15-19, 1999, California, USA.
- Jin R., Si L., Zhai C., Callan J., « Collaborative Filtering with Decoupled Models for Preferences and Ratings », *Proceedings of the 12th International Conference on Information and Knowledge Management (CIKM'03)*, Nov 3-8, 2003, Louisiana, USA.
- Maltz D., Ehrlich K., « Pointing the way : active collaborative filtering », *Proceeding of the Annual ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'95)*, May 07-11, 1995, Denver, Colorado, USA.
- Resnick P., Iacovou N., Suchak M., Bergstrom P., Riedl J., « GroupLens : An Open Architecture for Collaborative Filtering of Netnews », *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW'94)*, Oct 22-26, 1994, NC, USA.