

# PersoRec : un système personnalisé de recommandations pour les folksonomies basé sur les concepts quadratiques

Mohamed Nader Jelassi\* \*\* \*\*\*, Sadok Ben Yahia\*\*, Engelbert Mephu Nguifo\* \*\*

\* Clermont Université, Université Blaise Pascal, LIMOS, BP 10448,  
F-63000 Clermont-Ferrand, France  
nader.jelassi, mephu@isima.fr

\*\*Université de Tunis El-Manar, Faculté des Sciences de Tunis, LIPAH, 2092,  
Tunis, Tunisie

nader.jelassi, sadok.benyahia@fst.rnu.tn

\*\*\* CNRS, UMR 6158, LIMOS, F-63173 Aubière

**Résumé.** Nous proposons un nouveau système appelé PersoRec afin de personnaliser les recommandations (d'amis, de tags ou de ressources) faites aux utilisateurs dans les folksonomies. La personnalisation des recommandations est réalisée en prenant en compte le profil des utilisateurs. Cette nouvelle donnée permet de proposer aux utilisateurs des tags ou/et ressources plus adaptées à leurs besoins. En plus du profil des utilisateurs, nous avons recours à leur historique de partage de tags et de ressources dans le but de regrouper les utilisateurs ayant partagé des tags et des ressources en commun tout en ayant des profils équivalents (*i.e.*, des structures appelées concepts quadratiques). Ces deux données prises en compte au moment du processus de recommandation a permis d'améliorer la qualité des recommandations faites aux utilisateurs. PersoRec est donc capable de générer une recommandation personnalisée pour chaque utilisateur selon le mode de recommandation qu'il désire (recommandation d'amis, de tags ou de ressources) et selon le profil qu'il possède.

## 1 Introduction et Motivations

Une *folksonomie* désigne un système de classification collaborative par les internautes (Mika, 2007). L'idée est de permettre à des utilisateurs de partager et de décrire des objets via des mots-clés (tags) librement choisis. Les *folksonomies* ont à tenir compte des besoins de ses utilisateurs lors de la recommandation de tags ou de ressources. Cela a incité les chercheurs à proposer des systèmes de recommandation personnalisés afin de suggérer les tags et ressources les plus appropriés aux utilisateurs et de répondre aux besoins de chaque utilisateur. En effet, le domaine de personnalisation tente de fournir des solutions afin d'aider les utilisateurs à partager les bons tags et les bonnes ressources parmi le très grand nombre de données dans les *folksonomies*. De plus, la personnalisation tente d'aider les utilisateurs à aborder le problème de surcharge d'information (Das et al., 2012). Et pour réussir ou tenter de répondre au mieux aux attentes de chaque utilisateur de la *folksonomie*, il est utile d'avoir

plus d'informations sur lui. Pour atteindre cet objectif, nous considérons une nouvelle dimension dans une *folksonomie* qui contient des informations supplémentaires sur les utilisateurs, classiquement composée de trois dimensions (utilisateurs, tags et ressources), et nous proposons une approche de regroupement des utilisateurs aux intérêts équivalents sous forme de structure appelées concepts quadratiques (Jelassi et al., 2013). Un concept quadratique illustre une conceptualisation partagée dans la *folksonomie*. Par exemple, un concept quadratique peut être : "*Jack et Kate qui sont âgés entre 18 et 25 ans ont utilisé les tags 'action' et 'aventure', parmi d'autres, pour annoter des films comme 'Indiana Jones' et 'Star Wars'.*". MOVIELENS.

## 2 Les Concepts Quadratiques

Les systèmes de recommandation ont besoin de comprendre les besoins des utilisateurs afin de leur fournir de meilleures recommandations. Pour atteindre cet objectif, nous considérons le profil des utilisateurs comme une quatrième dimension d'une folksonomie classiquement composée de trois dimensions <utilisateurs, tags, ressources> (le nouveau contexte est ainsi dénoté *v-folksonomie*). La tâche consiste ensuite à regrouper des utilisateurs ayant des intérêts communs dans des structures appelées *concepts quadratiques* (ou quadri-concepts) (Jelassi et al., 2013). Cependant, cette tâche, bien que coûteuse, est exécutée hors-ligne et une seule fois, et n'a aucune incidence sur le temps d'exécution des recommandations.

Une question se pose alors : pourquoi les quadri-concepts ? D'un côté, si on peut facilement étudier les tags utilisés par un seul utilisateur sur une ressource, il est évident de constater que la tâche devient rapidement intraitable lorsque cela implique plusieurs utilisateurs et plusieurs ressources. D'un autre côté, les tags (ou ressources) recommandés s'avèrent ne pas être très spécifiques (Jäschke et al., 2007), *i.e.*, des tags qui sont des mots "bateau" ou bien des ressources ne correspondant pas aux besoins spécifiques de l'utilisateur. Grâce aux quadri-concepts, nous pouvons résoudre ces deux problèmes. En effet, d'un côté, les quadri-concepts sont des structures regroupant les tags et ressources en commun à un ensemble maximal d'utilisateurs. D'un autre côté, dans un quadri-concept, les tags et ressources qui ont été utilisés en combinaison seront regroupés d'où un résultat plus spécifique et répondant au besoin de notre système de recommandation. Ces concepts sont une représentation réduite de la *v-folksonomie* qui peut contenir des milliers de quadruplets dans la vraie vie. Un exemple de quadri-concept serait : "*Jack et Kate qui sont âgés entre 18 et 25 ans utilisent les tags 'action' et 'aventure', parmi d'autres, pour annoter des films comme 'Indiana Jones' et 'Star Wars'.*". Une fois extraits, ces quadri-concepts sont utilisés pour notre algorithme de recommandation personnalisée. De plus, les utilisateurs des *folksonomies* sont intéressés par des recommandations multi-mode (utilisateurs, tags et ressources), ainsi, les algorithmes répondant à cette attente sont appréciés (Ricci et al., 2011). Enfin, nous nous intéressons à une minorité de nouveaux utilisateurs qui risquent de ne pas recevoir de recommandations. Cette problématique est connue comme le "démarrage à froid" (ou cold start) (Ricci et al., 2011).

### 3 PERSOREC : un système personnalisé de recommandation basé sur les quadri-concepts

Le pseudo code de PERSOREC est disponible dans les papiers (Jelassi et al., 2013) (Jelassi et al., 2014). PERSOREC prend un ensemble de quadri-concepts  $QC$  comme entrée ainsi qu'un utilisateur cible  $u$  avec son profil  $v$  et (optionnellement) une ressource  $r$  (à annoter). PERSOREC permet de générer trois ensembles : un ensemble d'utilisateurs proposés, un ensemble de tags suggérés et un ensemble de ressources recommandées. En plus de l'algorithme original, nous avons ajouté une mesure de score afin de classer les recommandations par ordre d'importance. La mesure (notée  $rec\_score$ ) correspondant à un profil  $v$  est définie comme suit :

$$rec\_score(r_i, v) = \frac{|u_i|}{|UU|} / \exists t_i \exists r_i \exists v_i \in v, (u_i, t_i, r_i, v_i) \in \mathcal{F}_v \quad (1)$$

La mesure  $rec\_score$  d'une ressource  $r_i$  correspondant à un profil  $v$  est le nombre d'utilisateurs uniques ayant le même profil  $v$  (ou au moins une information de profil  $v_i \in v$ ) et ayant partagé la même ressource  $r_i$ , divisé par le nombre total d'utilisateurs uniques dans l'ensemble des quadri-concepts (noté  $UU$ ). Par exemple, si une ressource  $r_1$  a été partagée par 7 utilisateurs différents parmi un ensemble total de 67 utilisateurs uniques, son score sera égal à 0.104. L'introduction de cette mesure de score a permis d'améliorer la qualité de nos recommandations sur des jeux de données du monde réel.

Afin d'implémenter notre système personnalisé de recommandation PERSOREC<sup>1</sup>, nous avons d'abord utilisé le langage de programmation C++ afin de générer l'ensemble de quadri-concepts à partir d'un jeu de données (Jelassi, 2014). Nous utilisons la structure *Vector* afin de stocker les quadri-concepts extraits. Ensuite, afin de générer une interface interactive de notre système de recommandation et de proposer des recommandations personnalisées aux utilisateurs, nous avons créé un site web qui gère et offre des recommandations aux utilisateurs en fonction de leurs profils. Afin d'atteindre cet objectif, nous utilisons le framework php Symfony. Symfony vise à accélérer la création et la maintenance d'applications web et de remplacer les tâches de codage répétitives.

### 4 Démonstration et exemple illustratif

La démonstration de notre système personnalisé de recommandation PERSOREC<sup>2</sup> démontre à travers deux jeux de données du monde réel, *i.e.*, MOVIELENS (<http://movielens.umn.edu/>) et BOOKCROSSING (<http://www.bookcrossing.com/>), le processus de recommandation pour un utilisateur individuel. Il est important de noter que notre système personnalisé de recommandation est générique, *i.e.*, peut être appliqué à n'importe quel jeu de données dont les données ont une structure quadratique (utilisateur, tag, ressource, profil). Dans ce qui suit, nous donnons un exemple illustratif du processus de recommandation sur le jeu de données MOVIELENS. Le processus de recommandation est basé à la fois sur le profil des utilisateurs ainsi que sur les quadri-concepts, *i.e.*, les tags et ressources partagés par

1. téléchargeable à partir de ce lien <http://www.isima.fr/jelassi/Folkrec>

2. une vidéo de démonstration est disponible sur ce lien <http://www.isima.fr/jelassi/Demo>

## Recommandation personnalisée dans les folksonomies

The screenshot shows the PersoRec interface. At the top, it says "Welcome to the Personalized recommender system PersoRec!". The user is identified as Yasmine. The main content is divided into three columns:

- Left Column:** "The Last Movie you have share:" shows "Titanic". Below it, "Friends:" lists "Fran" and "Julie".
- Center Column:** "We recommend you the movies:" features "Bridget Jones's Diary (2001)" with a score of 0.084. Below it, a smaller movie card for "Titanic (1997)" is visible.
- Right Column:** "Friends recommendation:" lists several users with their ages and locations, such as "nader (27 ans) -Tunis-" and "Mouma (26 ans) -quebec-".

FIG. 1 – Un snapshot du siteweb PERSOREC pour le jeu de données MOVIELENS. (*gauche*) le profil de Yasmine, ses films partagés et sa liste d'amis (*centre*) les recommandations de films our l'utilisateur Yasmine (*droite*) la liste d'amis proposés à Yasmine.

The screenshot shows the "You want to share the movie 'Raiders of the Lost Ark'" interface. On the left is the movie poster for "Raiders of the Lost Ark". To the right, there is a "Tags:" section with three empty input boxes. Below that, "Recommended Tags:" shows buttons for "Indiana Jones", "archaeology", and "Nazis". A "Share this movie" button is also present.

FIG. 2 – Recommandation de tags lorsque Yasmine souhaite partager le film Raiders of Lost Ark.

des utilisateurs ayant des profils et intérêts équivalents. Considérons l'utilisateur *Yasmine* qui utilise notre système.

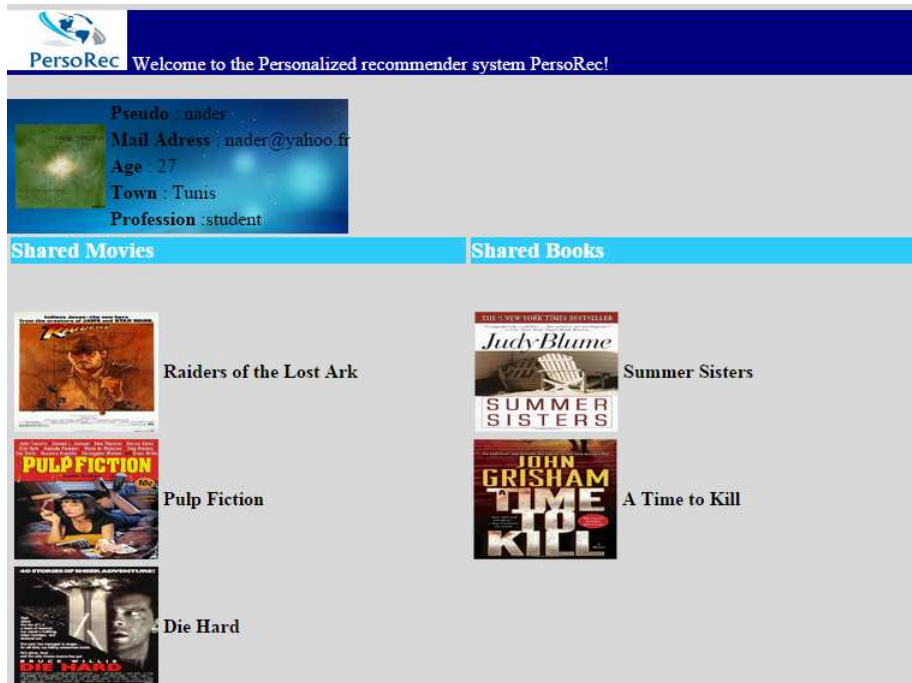


FIG. 3 – Consultation du profil de l'utilisateur nader avant son ajout à la liste d'amis.

La Figure 1 illustre un snapshot (ou instantané) du siteweb PERSOREC pour le jeu de données MOVIELENS. À gauche de la Figure, PERSOREC affiche le profil de *Yasmine* ainsi que les films qu'il a partagé et sa liste d'amis. Tandis qu'au centre de la Figure, nous affichons les recommandations de films avec leurs scores correspondants pour *Yasmine* (Voir la mesure de score dans (Jelassi et al., 2015)). À droite de l'interface, nous affichons une proposition d'une liste d'amis pour *Yasmine*. Ainsi, chaque utilisateur du siteweb PERSOREC recevra des recommandations basées sur son profil ainsi que sur les tags et ressources partagés par des utilisateurs ayant le même profil. Par ailleurs, il/elle recevra une liste d'amis proposés. Chaque utilisateur est aussi capable de partager des ressources (recommandés ou pas), d'ajouter des amis ou encore d'annoter des ressources avec des tags (librement choisis ou parmi ceux suggérés). Ainsi, par exemple, lorsque *Yasmine* souhaite partager le film *Raiders of Lost Ark*, PERSOREC lui suggère des tags qui ont été utilisés par des utilisateurs ayant le même profil, *i.e.*, *indiana jones*, *archeology* et *nazis* (*c.f.*, Figure 2). Enfin, avant d'éventuellement ajouter un utilisateur à sa liste d'amis, *Yasmine* a la possibilité de consulter son profil, *i.e.*, ses informations personnelles ou encore ses ressources partagées comme démontré par la Figure 3. La personnalisation des recommandation à travers le profil a permis d'améliorer leurs qualités par rapport aux principales approches de la littérature (Bellogín et al., 2013) (Qumsiyeh et Ng,

2012) (Kim et al., 2011).

## 5 Conclusion et Perspectives

Dans ce papier, nous présentons notre système personnalisé de recommandation et une démonstration de notre site web correspondant. En guise de perspectives, nous souhaitons améliorer notre plateforme de recommandations en intégrant des mises à jour incrémentales afin de mettre à jour les recommandations offertes aux utilisateurs.

## Références

- Bellogín, A., I. Cantador, et P. Castells (2013). A comparative study of heterogeneous item recommendations in social systems. *Inf. Sci.* 221, 142–169.
- Das, M., S. Thirumuruganathan, S. Amer-Yahia, G. Das, et C. Yu (2012). Who tags what ? an analysis framework. In *Proceedings of PVLDB* 5(11), 1567–1578.
- Jäschke, R., L. Marinho, A. A. Hotho, S.-T. Lars, et G. Stum (2007). Tag recommendations in folksonomies. In *Proc. of the 11th ECML PKDD, Warsaw, Poland*, pp. 506–514.
- Jelassi, M. N. (2014). Une nouvelle approche pour l'extraction efficace des quadri-concepts fréquents. In *RJCIA 2014*, Rouen, France.
- Jelassi, M. N., S. Ben Yahia, et E. Mephu Nguifo (2013). A personalized recommender system based on users' information in folksonomies. In *Proc. of the 22nd International Conference on World Wide Web companion, WWW '13 Companion*, pp. 1215–1224.
- Jelassi, M. N., S. Ben Yahia, et E. Mephu Nguifo (2014). Vers des recommandations plus personnalisées dans les folksonomies. In *IC 2014*, Clermont-Ferrand, France.
- Jelassi, M. N., S. Ben Yahia, et E. Mephu Nguifo (2015). Towards more targeted recommendations in folksonomies. *Social Network Analysis and Mining* 5(1).
- Kim, H. K., H. Y. Oh, J. C. Gu, et J. K. Kim (2011). Commenders : A recommendation procedure for online book communities. *Electron. Commer. Rec. Appl.* 10(5), 501–509.
- Mika, P. (2007). Ontologies are us : A unified model of social networks and semantics. *Journal of Web Semantics.* 5(1), 5–15.
- Qumsiyeh, R. et Y.-K. Ng (2012). Predicting the ratings of multimedia items for making personalized recommendations. In *SIGIR'12*, New York, NY, USA, pp. 475–484. ACM.
- Ricci, F., L. Rokach, B. Shapira, et P. B. Kantor (Eds.) (2011). *Recommender Systems Handbook*. Springer.

## Summary

Recommender systems are now popular both commercially as well as within the research community, where many approaches have been reported in the literature. We propose a new personalized recommender system namely PERSOREC that relies on structures called quadratic concepts which group users with close interests.