

Problématique

- Les systèmes d'accès à l'information ne tiennent pas suffisamment compte :
 - Des préférences de l'utilisateur
 - Du contexte de l'utilisateur

Objectifs

- Améliorer la pertinence des SRI
 - Avec les méthodes de la RI collaborative
 - En exploitant les interactions des utilisateurs
- Prendre en compte les profils utilisateurs
 - Pour personnaliser les résultats

Proposition

- Construction d'un profil utilisateur P(U) fondé sur l'historique des requêtes:

$$P(U) = \{(q_1^U, D_{q_1}^U), (q_2^U, D_{q_2}^U), \dots, (q_N^U, D_{q_N}^U)\}$$

- q_i^X : i-ème requête de l'utilisateur X.
- $D_{q_i}^X$: ensemble des documents qui sont sélectionnés comme pertinents pour la requête q_i de l'utilisateur X.
- Construire une communauté évolutive à partir des profils similaires
 - Proposition de 3 méthodes de calcul de similarité entre profils utilisateurs (fig. 1) fondées sur:
 - les requêtes (fig. 1.a)
 - les documents (fig. 1.b)
 - les paires requête – documents (fig. 1.c)
- Exploiter les requêtes similaires de cette communauté (fig. 2)

$$S_q(q_1, q_2) = \frac{\sum_{i=1}^K cw_i(q_1) \times cw_i(q_2)}{\sqrt{\sum_{i=1}^S w_i^2(q_1)} \times \sqrt{\sum_{i=1}^T w_i^2(q_2)}}$$

Calcul de similarité de requêtes [4]

$$S(P_q(X), P_q(Y)) = \frac{\sum_{i=1}^N \sum_{j=1}^M s_q(q_i^X, q_j^Y)}{N \times M}$$

Fig. 1a : Similarité de profils fondés sur les requêtes

$$S_d(d_1, d_2) = \cosine(\vec{d}_1, \vec{d}_2) = \frac{\vec{d}_1 \cdot \vec{d}_2}{\|\vec{d}_1\| \times \|\vec{d}_2\|}$$

Calcul de similarité de document

$$S(P_d(X), P_d(Y)) = \frac{\sum_{i=1}^N \sum_{j=1}^M s_d(d_i^X, d_j^Y)}{N \times M}$$

Fig. 1b : Similarité de profils fondés sur les documents

$$S_{qd}((q_1, D_{q_1}), (q_2, D_{q_2})) = \left(\frac{\alpha_q^2}{S_q^2(q_1, q_2)} + \frac{\alpha_d^2}{S_d^2(D_{q_1}, D_{q_2})} \right)^{-1}$$

Calcul de similarité de requête – documents [1]

$$S(P(X), P(Y)) = \frac{\sum_{i=1}^N \sum_{j=1}^M S_{qd}((q_i^X, D_{q_i}^X), (q_j^Y, D_{q_j}^Y))}{N \times M}$$

Fig. 1c : Similarité de profils fondés sur les paires requête-documents [1]

Fig. 1 : Trois méthodes de calcul de similarité entre les utilisateurs

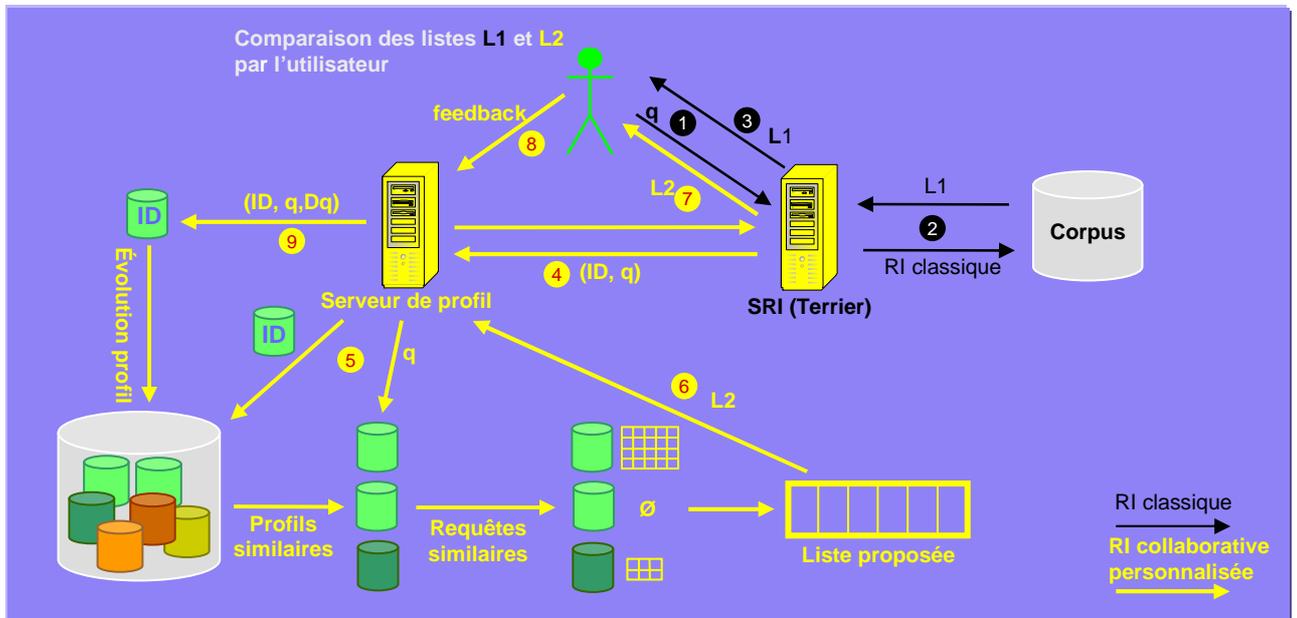


Fig. 2 : Exploitation de l'historique des requêtes utilisateurs [3]

Perspectives

- Intégrer les autres composantes du profil utilisateur
 - Caractéristiques utilisateur : fonction, expertise
 - Centres d'intérêts
- Adaptation des documents au contexte de l'utilisateur
 - Adaptation physique : format du texte, taille [2]
 - Adaptation sémantique : résumé sémantique

Publications

- [1] Naderi H., Rumpler B., "PERCIRS: a PERSONalized Collaborative Information Retrieval System", INFORSID, 1-3 June 2006, Hammamet Tunisia.
- [2] Naderi H., Rumpler B., Pinon J.M., "A Two Layered personalized Information Retrieval System", MAPS'06, in conjunction with ICPS'06 IEEE, 29 June 2006, Lyon, France.
- [3] Naderi H., Rumpler B., Pinon J.M., "An Efficient Collaborative Information Retrieval System by Incorporating the User Profile", AMR 2006, Geneva, Switzerland, July 2006.
- [4] Wen J., Nie J., and Zhang H., "Clustering user queries of a search engine", 10th International WWW Conference, W3C, 2001.