

## KONSEP MULTICRITERIA COLLABORATIVE FILTERING UNTUK PERBAIKAN REKOMENDASI

Wiranto<sup>1</sup>, Edi Winarko<sup>2</sup>

<sup>1</sup>Jurusan Teknik Informatika, Universitas Sebelas Maret

<sup>2</sup>Program Studi Ilmu Komputer, Universitas Gajah Mada

E-mail: [remove\\_hyperlink@snati2010](mailto:remove_hyperlink@snati2010), [nama\\_penulis@snati2010](mailto:nama_penulis@snati2010)

### ABSTRAKS

Untuk membantu pencari informasi yang belum memiliki referensi diperlukan alat bantu recommender system. Pengembangan recommender system sebagian besar dilakukan dengan menggunakan pendekatan berbasis collaborative filtering. Sistem berbasis collaborative filtering akan bekerja dengan cara mempelajari kebiasaan para pencari informasi dan membangun profil pencari informasi, kemudian memberikan rekomendasi.

Pendekatan collaborative filtering klasik diterapkan pada kasus pemilihan item yang hanya memiliki satu kriteria. Sementara itu, banyak kasus yang tidak bisa dimodelkan dengan satu kriteria. Oleh karena itu konsep collaborative filtering perlu dikembangkan untuk pemilihan item yang memiliki banyak kriteria agar rekomendasi yang dihasilkan memiliki kualitas lebih baik dan relevan dengan kebutuhan pengguna.

*Kata Kunci: recommender system, collaborative filtering, kemiripan, multi kriteria*

### 1. PENDAHULUAN

Ledakan informasi pada jaringan internet menyebabkan masyarakat mengalami kesulitan mendapatkan informasi yang cepat dan relevan dengan kebutuhannya. Untuk itu diperlukan alat bantu yang dapat mengakses dan menemukan informasi seperti yang dikehendaki.

Dalam proses pencarian informasi terdapat dua tipe pencari. Tipe pertama adalah pencari yang sudah memiliki referensi yang jelas informasi apa yang hendak dicari. Tipe seperti ini cukup dibantu dengan mesin pencari (*search engine*). Sedangkan tipe kedua adalah seseorang yang tidak berbekal referensi, namun hanya memiliki topik tertentu. Pencari tipe kedua ini dapat dipastikan mengalami kesulitan untuk mendapatkan informasi yang dimaksud. Oleh karena itu, dia tidak cukup dibantu hanya dengan *search engine* tetapi memerlukan suatu sistem lain yang disebut dengan *recommender system* atau sistem penyedia rekomendasi.

*Recommender system* merupakan model penyelesaian masalah yang menerapkan teknik-teknik tertentu pada pembuatan rekomendasi untuk pemilihan suatu informasi, produk dan jasa (Goldberg,2001). Terdapat dua pendekatan dalam mengembangkan suatu *recommender system* yaitu *content-based* dan *collaborative filtering* (Kangas,2002).

Pendekatan *content-based* akan menyediakan rekomendasi dengan cara membandingkan representasi konten yang terkandung dalam suatu item dengan representasi konten yang diinginkan pengguna. Untuk itu, harus dilakukan proses *indexing*, yaitu melakukan ekstraksi informasi yang dikandung item.

Sedangkan sistem yang dibangun dengan pendekatan *collaborative filtering* akan bekerja dengan cara mempelajari kebiasaan para pencari informasi dan membangun profil pencari informasi, kemudian memberikan rekomendasi. Ide awal yang melandasi munculnya teknik *collaborative filtering* adalah perbandingan rasa suka atau tidak suka seseorang, misalnya A, terhadap sekumpulan produk, dengan orang lain, misalnya B, untuk memprediksi pilihan A untuk produk yang lain (Goldberg,2001 dan Kangas,2002).

Pendekatan *collaborative filtering* terbagi menjadi dua, yaitu pendekatan berbasis pengguna (*user based*) dan pendekatan berbasis item (*item based*). Pendekatan *user based* memanfaatkan histori mengenai pilihan pencari informasi, sedangkan *item based* memanfaatkan histori nilai rating dari informasi yang dicari. Untuk meningkatkan akurasi dan efisiensi, telah dikembangkan sistem dengan pendekatan *hybrid*, yaitu mengkombinasikan pendekatan *user based* dan *item based* (Vozalis,2003).

Penelitian dengan topik *recommender system* relatif belum banyak dilakukan mengingat topik ini termasuk disiplin baru dalam bidang ilmu komputer. Dari beberapa penelitian yang sudah dilakukan pun, pendekatan *collaborative filtering* hanya diterapkan pada kasus-kasus item yang memiliki kriteria tunggal. Sementara itu, banyak permasalahan yang tidak bisa dimodelkan dengan kriteria tunggal. Berdasarkan realitas tersebut, pada penelitian ini akan dilakukan eksplorasi pendekatan *collaborative filtering* untuk pemilihan item yang memiliki multi kriteria

## 2. PENULISAN

Prinsip kerja dari algoritma *collaborative filtering* adalah memberikan rekomendasi atau prediksi item berdasarkan pada opini pengguna-pengguna yang mempunyai kemiripan. Opini yang diberikan bisa bersifat eksplisit maupun implisit (Sarwar,2001).

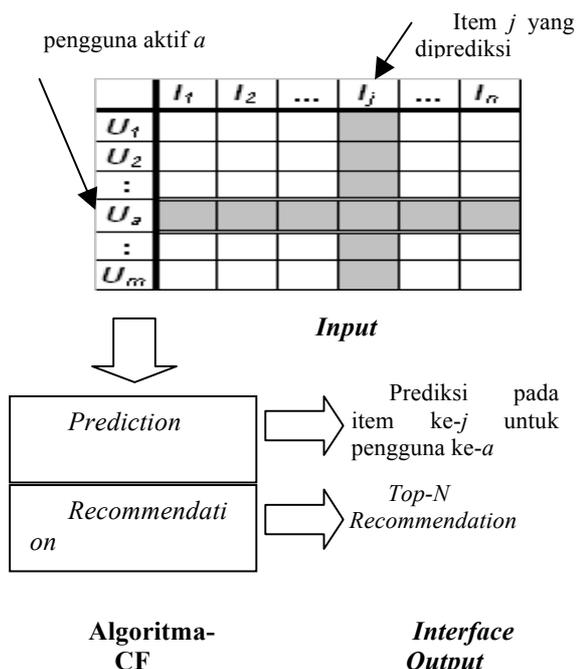
Tujuan dari algoritma *collaborative filtering* adalah menganjurkan item baru atau memprediksikan evaluasi item kepada pengguna khusus berdasarkan pada ketertarikan sebelumnya dari pengguna dan opini dari pengguna-pengguna lain yang mempunyai ketertarikan yang mirip. Opini bisa diberikan secara eksplisit oleh pengguna berupa nilai *rating* atau bisa juga secara implisit dihasilkan dari riwayat pembelian, dengan melakukan analisis terhadap *logs*, *navigation history* atau dengan cara yang lain (Sarwar,2001).

Ada dua hal utama yang dilakukan *recommender system* yang menggunakan algoritma *collaborative filtering*, yaitu:

- Prediction*, melakukan prediksi opini yang akan diberikan oleh pengguna dalam skala bilangan yang sama.
- Recommendation*, memberikan rekomendasi berupa daftar item dengan nilai prediksi tertinggi. Hal yang perlu dicatat adalah item-item yang direkomendasikan belum pernah dibeli, dilihat atau di-*rating* oleh pengguna tersebut. Antarmuka ini pada *recommender system* yang menggunakan algoritma *collaborative filtering* disebut dengan *Top-N recommendation* (Sarwar,2001).

Gambar 1 menunjukkan skema kerja dari *recommender system* yang menggunakan pendekatan *collaborative filtering* dengan *profile* berbentuk *user-item ratings matrix* berukuran  $m \times n$ , dimana  $m$  adalah jumlah pengguna dan  $n$  adalah jumlah item sebagai *input* (Sarwar,2001).

Dalam implementasinya, sebenarnya terjadi proses komputasi yang cukup rumit sebab sebelum dilakukan prediksi harus terlebih dahulu dibentuk sebuah *neighborhood* sebagai dasar perhitungan prediksi.

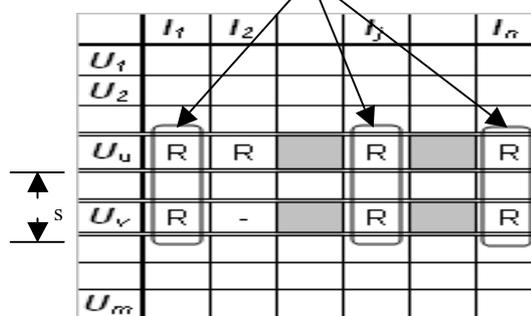


Gambar 1 Proses Collaborative Filtering

Pendekatan *collaborative filtering* pada dasarnya dibagi menjadi dua kategori yaitu *user based collaborative filtering* disebut juga *memory-based*, dan *item based collaborative filtering* yang disebut juga *model-based* (Sarwar, 2001).

Pada pendekatan *user based collaborative filtering*, proses pembentukan *user-neighborhood* dilakukan dengan terlebih dahulu melakukan perhitungan kemiripan antara pengguna aktif dengan pengguna lainnya (*user-user similarity*). Perhitungan hanya diberlakukan pada himpunan item-item yang telah di-*rating* oleh kedua pengguna yang dibandingkan seperti ditunjukkan Gambar 2.

*User-user similarity* hanya dihitung dari item-item yang telah di-*rating* oleh pengguna  $u$  dan pengguna  $v$ .



Gambar 2. Pembentukan User-Neighborhood

Untuk menghitung *user-user similarity* bisa digunakan persamaan *cosine-based similarity* seperti dalam persamaan (1). Pengguna yang

dibandingkan misalnya  $u$  dan  $v$ , dianggap sebagai vektor baris dengan anggotanya adalah nilai *rating* yang terdapat pada baris tersebut. Dua vektor dikatakan sama jika membentuk sudut  $0^\circ$  atau nilai cosinusnya 1. Dengan kata lain dua pengguna dikatakan mirip jika nilai cosinus dari perhitungan mendekati 1.

$$s(u, v) = \cos(\vec{R}(u, *), \vec{R}(v, *)) = \frac{\vec{R}(u, *) \cdot \vec{R}(v, *)}{\|\vec{R}(u, *)\| \cdot \|\vec{R}(v, *)\|} \quad (1)$$

di mana :

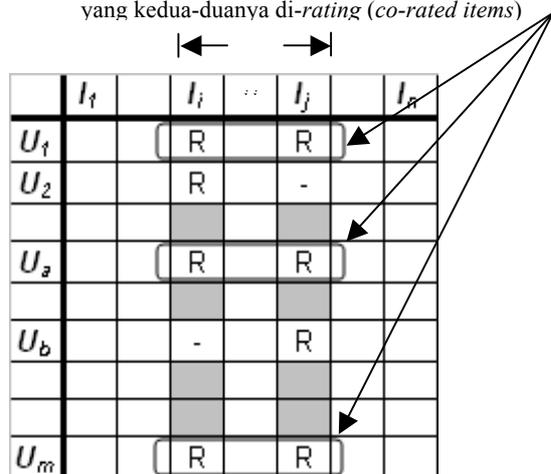
$\vec{R}(u, *)$  adalah vektor baris dengan anggota nilai *rating* pada pengguna  $u$ .

$\cos(\vec{R}(u, *), \vec{R}(v, *))$  adalah nilai cosinus sudut yang dibentuk vektor baris *rating* pengguna  $u$  dan  $v$ .

Sedangkan pendekatan *item based collaborative filtering* menjadikan item-item yang telah di-*rating* oleh pengguna menjadi dasar perhitungan. Algoritma ini melakukan perhitungan bagaimana kemiripan item-item yang telah di-*rating* dengan item-item lain dan selanjutnya dipilih sekelompok item yang mempunyai kemiripan dengan item yang sudah di-*rating*. Nilai kemiripan tersebut dicatat untuk dijadikan nilai bobot untuk memprediksi nilai *rating* pada item target.

Dasar perhitungan kemiripan antara dua item  $i$  dan  $j$  adalah terlebih dahulu mengisolasi pengguna-pengguna yang telah menilai keduanya kemudian teknik *item similarity* diterapkan untuk memperoleh nilai kemiripan seperti ditunjukkan Gambar 3.

Item-Item similarity dihitung hanya pada item-item yang kedua-duanya di-*rating* (co-rated items)



Gambar 3. Pembentukan *Item-Neighborhood*

Nilai *item-item similarity* bisa dihitung dengan menggunakan persamaan *cosine-based similarity* seperti halnya menghitung *user-user similarity*. Yang berbeda adalah orientasinya, yaitu

membandingkan antara dua item, misalnya item  $i$  dan item  $j$ . Sehingga item-item tersebut dianggap sebagai vektor kolom dengan anggota nilai *rating* pada kolom tersebut (Deshpande, 2004). Persamaan (2) menunjukkan hal tersebut.

$$s(i, j) = \cos(\vec{R}(*, i), \vec{R}(*, j)) = \frac{\vec{R}(*, i) \cdot \vec{R}(*, j)}{\|\vec{R}(*, i)\| \cdot \|\vec{R}(*, j)\|} \quad (2)$$

di mana :

$\vec{R}(*, i)$  adalah vektor kolom yang anggotanya nilai *rating* pada item  $i$ .

Untuk memperjelas proses yang terjadi dalam *collaborative filtering*, diberikan ilustrasi sederhana seperti Gambar 4. Diasumsikan ada lima pengguna yaitu  $u_1, \dots, u_5$  dan lima item  $i_1, \dots, i_5$ . Yang ingin dilakukan adalah memprediksi nilai *rating* yang diberikan pengguna  $u_1$  terhadap item  $i_5$ . Secara klasik, akan dicari beberapa pengguna yang sudah memberi nilai *rating* untuk  $i_5$  dan paling mirip dengan  $u_1$ . Dalam kasus ini ditemukan yang paling dekat dengan  $u_1$  adalah  $u_2$  dan  $u_3$ . Karena kedua pengguna sudah memberikan nilai *rating* 7 pada  $i_5$ , maka sistem akan memberikan nilai prediksi  $u_1$  untuk  $i_5$  sebesar 7 juga.

	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$
$u_1$	4	6	4	6	?
$u_2$	4	6	4	6	7
$u_3$	4	6	4	6	7
$u_4$	5	5	5	5	4
$u_5$	5	5	5	5	4

Gambar 4. Ilustrasi *Collaborative Filtering* Klasik

### 3. KOLABORATIVE FILTERING MULTI KRITERIA

Munculnya gagasan *collaborative filtering* multi kriteria didasari kenyataan bahwa banyak item yang nilai *rating*-nya tidak bisa diwakili hanya dengan satu kriteria saja. Sebagai contoh, dalam suatu sistem *e-commerce*, produk dinilai oleh pelanggan berdasarkan beberapa kriteria, seperti kualitas barang, harga, pengiriman, layanan purna jual dan lain sebagainya (Adomavicius, 2007). Oleh karena itu, penggunaan teknik *collaborative filtering* klasik (dengan satu kriteria) dalam beberapa aplikasi sering tidak relevan dan tidak mengakomodasi opini yang diberikan oleh pengguna.

Untuk mempermudah pemahaman terhadap konsep *collaborative filtering* multi kriteria diberikan ilustrasi seperti Gambar 5. Nampak dalam gambar bahwa setiap pengguna memberikan nilai rating untuk empat kriteria yang berbeda. Dari empat nilai rating tersebut kemudian dilakukan perhitungan nilai rating keseluruhan (*overall*). Misalnya nilai rating yang diberikan  $u_1$  untuk  $i_1$  adalah 2, 2, 6, 6 dan nilai rating keseluruhan adalah 4, yang merupakan hasil rata-rata dari nilai keempat kriteria tersebut. Fungsi rata-rata hanya sebagai contoh sederhana saja.

	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$
$u_1$	4 <sub>(2,2,6,6)</sub>	6 <sub>(4,4,8,8)</sub>	4 <sub>(2,2,6,6)</sub>	6 <sub>(4,4,8,8)</sub>	?
$u_2$	4 <sub>(6,6,2,2)</sub>	6 <sub>(8,8,4,4)</sub>	4 <sub>(6,6,2,2)</sub>	6 <sub>(8,8,4,4)</sub>	7
$u_3$	4 <sub>(6,6,2,2)</sub>	6 <sub>(8,8,4,4)</sub>	4 <sub>(6,6,2,2)</sub>	6 <sub>(8,8,4,4)</sub>	7
$u_4$	5 <sub>(2,2,8,8)</sub>	5 <sub>(3,3,7,7)</sub>	5 <sub>(2,2,8,8)</sub>	5 <sub>(3,3,7,7)</sub>	4
$u_5$	5 <sub>(2,2,8,8)</sub>	5 <sub>(3,3,7,7)</sub>	5 <sub>(2,2,8,8)</sub>	5 <sub>(3,3,7,7)</sub>	4

Gambar 5. Ilustrasi *Collaborative Filtering* Multi Kriteria

Kalau pada teknik *collaborative filtering* klasik sebelumnya pengguna  $u_2$  dan  $u_3$  sebagai pengguna yang paling mirip dengan  $u_1$ , maka pada konsep multi kriteria hal itu tidak berlaku. Hal ini disebabkan oleh perbedaan yang sangat jauh nilai rating multi kriteria antara  $u_2$  dan  $u_3$  dengan  $u_1$  meskipun nilai rating keseluruhan sama persis. Sebagai contoh, nilai rating yang diberikan  $u_1$  untuk  $i_1$  adalah 4<sub>(2,2,6,6)</sub> dan nilai rating yang diberikan  $u_2$  untuk  $i_1$  adalah 4<sub>(6,6,2,2)</sub>. Jadi meskipun nilai keseluruhannya sama, yaitu 4 namun nilai dari keempat kriterianya saling bertolak belakang.

Berdasarkan pendekatan *collaborative filtering* multi kriteria, pengguna  $u_4$  dan  $u_5$  justru dinilai memiliki kemiripan dengan  $u_1$ . Meskipun nilai keseluruhan tidak mirip, yaitu 4 dan 5, namun nilai dari masing-masing kriterianya justru lebih mendekati. Sehingga nilai prediksi  $u_1$  untuk item  $i_5$  ditetapkan 4, sama dengan nilai yang diberikan oleh pengguna  $u_4$  dan  $u_5$ .

Dengan demikian, terlihat sangat jelas bahwa konsep *collaborative filtering* multi kriteria akan memberikan hasil rekomendasi yang lebih relevan kepada pengguna dibanding dengan *collaborative filtering* klasik. Persoalannya adalah bagaimana menemukan pendekatan yang tepat untuk menghitung nilai rating keseluruhan.

Apabila pendekatan yang digunakan adalah perhitungan kemiripan, bisa dilakukan dengan dua cara, yaitu :

- Menerapkan fungsi penghitungan kemiripan antar pengguna atau antar item, seperti yang ditunjukkan persamaan (1) dan (2) untuk setiap kriteria. Misalnya suatu item memiliki  $k$  kriteria, maka setiap pengguna akan memberikan nilai rating untuk  $k$  kriteria yaitu  $r_1, \dots, r_k$  dan nilai rating untuk keseluruhan  $r_0$ . Perhitungan kemiripan dilakukan untuk  $k$  kriteria dan rating secara keseluruhan.  $s_1(u, v)$  merepresentasikan kemiripan untuk kriteria pertama,  $s_2(u, v)$  merepresentasikan kemiripan untuk kriteria kedua dan seterusnya. Sementara kemiripan secara keseluruhan  $s_0(u, v)$  dihitung dengan menggunakan fungsi tertentu, misalnya fungsi rata-rata seperti persamaan (3).

$$s_{avg}(u, v) = \frac{1}{k+1} \sum_{i=0}^k s_i(u, v) \quad (3)$$

- Menggunakan metrik jarak multidimensi, di mana setiap nilai rating kriteria  $r_0, r_1, \dots, r_k$  membentuk ruang berdimensi  $k+1$ . Dalam pendekatan ini berlaku hubungan terbalik antara jarak dan kemiripan, artinya semakin pendek jarak antara dua pengguna maka semakin besar kemiripannya. Untuk menghitung metrik jarak dua pengguna dilakukan tiga tahap. Tahap pertama, menghitung jarak di antara dua pengguna untuk masing-masing kriteria, misalnya  $d_{rating}(R(u, i), R(v, i))$  di mana  $R(u, i) = (r_0, r_1, \dots, r_k)$  dan  $R(v, i) = (r'_0, r'_1, \dots, r'_k)$  dengan menggunakan rumus Euclid seperti yang ditunjukkan persamaan (4).

$$\sqrt{\sum_{i=0}^k |r_i - r'_i|^2} \quad (4)$$

Tahap kedua, menghitung jarak keseluruhan antara dua pengguna, dengan menggunakan rumus persamaan (5).

$$d_{user}(u, v) = \frac{1}{|I(u, v)|} \sum_{i \in I(u, v)} d_{rating}(R(u, i), R(v, i)) \quad (5)$$

Tahap ketiga adalah menghitung tingkat kemiripan dua pengguna dengan menggunakan rumus persamaan (6).

$$s(u, v) = \frac{1}{1 + d_{user}(u, v)} \quad (6)$$

#### 4. PENDEKATAN FUNGSI AGREGASI

Sebelumnya sudah dijelaskan pendekatan yang digunakan untuk menghitung nilai rating keseluruhan berbasis kemiripan. Ada cara lain yang bisa digunakan yaitu dengan menggunakan pendekatan fungsi agregasi. Gagasan ini berangkat dari asumsi bahwa sistem rating dengan multi kriteria merepresentasikan preferensi pengguna untuk beberapa komponen yang berbeda. Sehingga rating keseluruhan suatu item merupakan fungsi agregasi dari rating seluruh kriteria, yang bisa dituliskan dengan persamaan (7).

$$r_0 = f(r_1, \dots, r_k)$$

Dengan kata lain, pendekatan fungsi agregasi didasari asumsi bahwa terdapat relasi tertentu antara rating keseluruhan dengan rating multi kriteria. Sebagai contoh, dalam *recommender system* untuk pemilihan *movie* yang menempatkan kriteria alur cerita sebagai prioritas. Maka suatu *movie* yang memiliki rating alur cerita yang tinggi pasti akan disukai pengguna dan mempengaruhi rating keseluruhan secara signifikan.

Pendekatan fungsi agregasi dilakukan dengan tiga tahapan, yakni memprediksi rating setiap kriteria, melakukan estimasi relasi rating keseluruhan dengan rating multi kriteria dan memprediksi rating keseluruhan.

Pada tahapan pertama, yakni memprediksi rating setiap kriteria, diawali dengan melakukan dekomposisi multi kriteria menjadi kriteria tunggal. Artinya permasalahan multi kriteria sudah diubah menjadi permasalahan *collaborative filtering* klasik sejumlah  $k$ . Selanjutnya bisa dilakukan prediksi nilai rating dengan menggunakan pendekatan kemiripan seperti rumus persamaan (1) atau (2).

Tahapan kedua, yaitu melakukan estimasi relasi antara rating keseluruhan dengan rating multi kriteria, yang ditulis dengan  $r_0 = f(r_1, \dots, r_k)$ . Hal ini bisa dilakukan dengan menggunakan masukan pakar yang menguasai domain permasalahan atau memanfaatkan teknik statistika seperti regresi linier, atau bisa juga menggunakan teknik *machine learning* seperti jaringan syaraf tiruan.

Tahapan terakhir adalah memprediksi nilai rating keseluruhan  $r'_0$  secara langsung dengan menggunakan nilai rating multi kriteria yang dihasilkan tahapan pertama dan fungsi  $f$  yang diestimasi pada tahapan kedua :

$$r_0 = f(r_1, \dots, r_k)$$

#### 5. KESIMPULAN

Konsep *collaborative filtering* berbasis multi kriteria memberikan kesempatan kepada pengguna *recommender system* untuk menilai suatu item

secara variatif. Agar nilai rating keseluruhan mencerminkan nilai rating multi kriteria maka perlu dikembangkan pendekatan fungsi agregasi yang tepat, sehingga rekomendasi yang dihasilkan akan memiliki kualitas lebih baik.

#### PUSTAKA

- Adomavicius, G and Kwon Y., 2007, *New Recommendation Techniques for Multicriteria Rating Systems*, Intelligent Systems, IEEE Computer Society.
- Deshpande, M., and Karypis, G., 2004, *Item-Based Top-N Recommendation Algorithms*, ACM Transaction on Information Systems.
- Goldberg, K., Roeder, T., Gupta, D., and Perkins, C., 2001, *Eigentaste : A Constant Time Collaborative Filtering Algorithms*, Information Retrieval Journal.
- Kangas, S., 2002, *Collaborative Filtering and Recommendation Systems*, LOUHI , Project Research Report, VTT Information Technology.
- Sarwar, BM., Karypis, G., Konstan, JA., and Riedl, TJ, 2001, *Item-Based Collaborative Filtering recommendation algorithms*, Proceeding of 10th International World Wide Web Conference, ACM Press.
- Vozalis, M., and Margaritis, KG., 2003, *On the Combination of user based and item-based collaborative filtering*, Technical Report, University of Macedonia