# Providing Justifications in Recommender Systems

Panagiotis Symeonidis, Alexandros Nanopoulos, and Yannis Naolopoulos

# Outline

1. New Ideas / Importance
2. Related work
3. Difference from related work
4. Problem description
5. Proposed Approach
6. Results

# New Ideas / Importance

## New Ideas:

1.Provide good recommendations- and also *justify* the recommendations.
2.Use user profile *and* item profile to make recommendations.
3.Use partial matching to match a user to a group of users resulting in better recommendations.

## Why is this important?

1. Idea #1 increases user acceptance and trust in the system by letting him know why he is being recommended an item. (Survey: Most users would like explanations to be added to recommender systems)
2. Idea#2 and #3 results in better recommendations.

# Related Work

1. **Collaborative Filtering(CF)**: Memory based algorithms use the entire database of users to find the k most similar users. Model based algorithms develop a model (ex. Clustering) and then recommend based on their model. Little research on recommendation justification.

2. **Content-based Filtering (CB):** Use the content of the items for recommendation. Ex. Use machine learning to categorize books, news articles. More research on recommendation justification.

3. Several attempts made at combining CB and CF.

# Differences from Related Work

1. Existing CF methods cluster users and items separately thus missing out on the duality between users and items.

2. Others have used bi-clustering. However they were more concerned with execution time. Recommendation accuracy was low.

3. Unlike other approaches the dependence between item features and user ratings is exploited.

# Problem description

1. Rating profile
2. Item profile
3. Feature profile
4. Explain coverage

# Rating profile

User-Item matrix(R):

| | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ |
|---|---|---|---|---|---|---|---|
| $U_1$ | 5 | - | 2 | - | 1 | - | - |
| $U_2$ | 2 | - | 4 | 1 | 4 | 3 | - |
| $U_3$ | 4 | - | 2 | - | 2 | - | 5 |
| $U_4$ | - | 3 | 1 | 4 | - | 5 | 2 |
| $U_5$ | - | 2 | 4 | 2 | 5 | 1 | - |
| $U_6$ | 5 | 1 | - | 1 | - | - | 3 |
| $U_7$ | - | 2 | 5 | - | 4 | 1 | - |
| $U_8$ | 1 | 4 | - | 5 | 4 | 3 | - |

The rating profile of user $U_1$:

$$R(U_1) = \{I_1, I_2, I_3\}$$

# Item profile

Item-Feature Matrix (F):

| | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| $I_1$ | 1 | 0 | 0 | 0 |
| $I_2$ | 0 | 1 | 1 | 0 |
| $I_3$ | 0 | 0 | 1 | 1 |
| $I_4$ | 0 | 1 | 0 | 1 |
| $I_5$ | 0 | 1 | 1 | 1 |
| $I_6$ | 1 | 0 | 1 | 1 |
| $I_7$ | 1 | 0 | 1 | 0 |

The rating profile of item $I_2$:

$$F(I_2) = \{f_2, f_3\}$$

# Feature profile

User-Item matrix(R):

| | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ |
|---|---|---|---|---|---|---|---|
| $U_1$ | 5 | - | 2 | - | 1 | - | - |
| $U_2$ | 2 | - | 4 | 1 | 4 | 3 | - |
| $U_3$ | 4 | - | 2 | - | 2 | - | 5 |
| $U_4$ | - | 3 | 1 | 4 | - | 5 | 2 |
| $U_5$ | - | 2 | 4 | 2 | 5 | 1 | - |
| $U_6$ | 5 | 1 | - | 1 | - | - | 3 |
| $U_7$ | - | 2 | 5 | - | 4 | 1 | - |
| $U_8$ | 1 | 4 | - | 5 | 4 | 3 | - |

$$P_\tau = 2$$

Item-Feature Matrix (F):

| | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| $I_1$ | 1 | 0 | 0 | 0 |
| $I_2$ | 0 | 1 | 1 | 0 |
| $I_3$ | 0 | 0 | 1 | 1 |
| $I_4$ | 0 | 1 | 0 | 1 |
| $I_5$ | 0 | 1 | 1 | 1 |
| $I_6$ | 1 | 0 | 1 | 1 |
| $I_7$ | 1 | 0 | 1 | 0 |

# Feature profile

User-Item matrix(R):

| | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ |
|---|---|---|---|---|---|---|---|
| $U_1$ | 5 | - | 2 | - | 1 | - | - |
| $U_2$ | 2 | - | 4 | 1 | 4 | 3 | - |
| $U_3$ | 4 | - | 2 | - | 2 | - | 5 |
| $U_4$ | - | 3 | 1 | 4 | - | 5 | 2 |
| $U_5$ | - | 2 | 4 | 2 | 5 | 1 | - |
| $U_6$ | 5 | 1 | - | 1 | - | - | 3 |
| $U_7$ | - | 2 | 5 | - | 4 | 1 | - |
| $U_8$ | 1 | 4 | - | 5 | 4 | 3 | - |

$$P_\tau = 2$$

Item-Feature Matrix (F):

| | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| $I_1$ | 1 | 0 | 0 | 0 |
| $I_2$ | 0 | 1 | 1 | 0 |
| $I_3$ | 0 | 0 | 1 | 1 |
| $I_4$ | 0 | 1 | 0 | 1 |
| $I_5$ | 0 | 1 | 1 | 1 |
| $I_6$ | 1 | 0 | 1 | 1 |
| $I_7$ | 1 | 0 | 1 | 0 |

# Feature profile

User-Item matrix(R):

| | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ |
|---|---|---|---|---|---|---|---|
| $U_1$ | 5 | - | 2 | - | 1 | - | - |
| $U_2$ | 2 | - | 4 | 1 | 4 | 3 | - |
| $U_3$ | 4 | - | 2 | - | 2 | - | 5 |
| $U_4$ | - | 3 | 1 | 4 | - | 5 | 2 |
| $U_5$ | - | 2 | 4 | 2 | 5 | 1 | - |
| $U_6$ | 5 | 1 | - | 1 | - | - | 3 |
| $U_7$ | - | 2 | 5 | - | 4 | 1 | - |
| $U_8$ | 1 | 4 | - | 5 | 4 | 3 | - |

$$P_\tau = 2$$

Item-Feature Matrix (F):

| | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| $I_1$ | 1 | 0 | 0 | 0 |
| $I_2$ | 0 | 1 | 1 | 0 |
| $I_3$ | 0 | 0 | 1 | 1 |
| $I_4$ | 0 | 1 | 0 | 1 |
| $I_5$ | 0 | 1 | 1 | 1 |
| $I_6$ | 1 | 0 | 1 | 1 |
| $I_7$ | 1 | 0 | 1 | 0 |

# Feature profile

User-Item matrix(R):

| | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ |
|---|---|---|---|---|---|---|---|
| $U_1$ | 5 | - | 2 | - | 1 | - | - |
| $U_2$ | 2 | - | 4 | 1 | 4 | 3 | - |
| $U_3$ | 4 | - | 2 | - | 2 | - | 5 |
| $U_4$ | - | 3 | 1 | 4 | - | 5 | 2 |
| $U_5$ | - | 2 | 4 | 2 | 5 | 1 | - |
| $U_6$ | 5 | 1 | - | 1 | - | - | 3 |
| $U_7$ | - | 2 | 5 | - | 4 | 1 | - |
| $U_8$ | 1 | 4 | - | 5 | 4 | 3 | - |

$P_\tau = 2$

Item-Feature Matrix (F):

| | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| $I_1$ | 1 | 0 | 0 | 0 |
| $I_2$ | 0 | 1 | 1 | 0 |
| $I_3$ | 0 | 0 | 1 | 1 |
| $I_4$ | 0 | 1 | 0 | 1 |
| $I_5$ | 0 | 1 | 1 | 1 |
| $I_6$ | 1 | 0 | 1 | 1 |
| $I_7$ | 1 | 0 | 1 | 0 |

$F_3 = 1$
$F_4 = 1$

# Feature profile

User-Item matrix(R):

| | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ |
|---|---|---|---|---|---|---|---|
| $U_1$ | 5 | - | 2 | - | 1 | - | - |
| $U_2$ | 2 | - | 4 | 1 | 4 | 3 | - |
| $U_3$ | 4 | - | 2 | - | 2 | - | 5 |
| $U_4$ | - | 3 | 1 | 4 | - | 5 | 2 |
| $U_5$ | - | 2 | 4 | 2 | 5 | 1 | - |
| $U_6$ | 5 | 1 | - | 1 | - | - | 3 |
| $U_7$ | - | 2 | 5 | - | 4 | 1 | - |
| $U_8$ | 1 | 4 | - | 5 | 4 | 3 | - |

$$P_\tau = 2$$

Item-Feature Matrix (F):

| | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| $I_1$ | 1 | 0 | 0 | 0 |
| $I_2$ | 0 | 1 | 1 | 0 |
| $I_3$ | 0 | 0 | 1 | 1 |
| $I_4$ | 0 | 1 | 0 | 1 |
| $I_5$ | 0 | 1 | 1 | 1 |
| $I_6$ | 1 | 0 | 1 | 1 |
| $I_7$ | 1 | 0 | 1 | 0 |

$F_3 = 1$
$F_4 = 1$

# Feature profile

User-Item matrix(R):

| | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ |
|---|---|---|---|---|---|---|---|
| $U_1$ | 5 | - | 2 | - | 1 | - | - |
| $U_2$ | 2 | - | 4 | 1 | 4 | 3 | - |
| $U_3$ | 4 | - | 2 | - | 2 | - | 5 |
| $U_4$ | - | 3 | 1 | 4 | - | 5 | 2 |
| $U_5$ | - | 2 | 4 | 2 | 5 | 1 | - |
| $U_6$ | 5 | 1 | - | 1 | - | - | 3 |
| $U_7$ | - | 2 | 5 | - | 4 | 1 | - |
| $U_8$ | 1 | 4 | - | 5 | 4 | 3 | - |

$$P_\tau = 2$$

Item-Feature Matrix (F):

| | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| $I_1$ | 1 | 0 | 0 | 0 |
| $I_2$ | 0 | 1 | 1 | 0 |
| $I_3$ | 0 | 0 | 1 | 1 |
| $I_4$ | 0 | 1 | 0 | 1 |
| $I_5$ | 0 | 1 | 1 | 1 |
| $I_6$ | 1 | 0 | 1 | 1 |
| $I_7$ | 1 | 0 | 1 | 0 |

$F_2 = 1$
$F_3 = 2$
$F_4 = 2$

# Feature profile

User-Feature matrix(P):

| | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| $U_1$ | 1 | 0 | 0 | 0 |
| $U_2$ | 1 | 1 | 3 | 3 |
| $U_3$ | 2 | 0 | 1 | 0 |
| $U_4$ | 1 | 2 | 2 | 2 |
| $U_5$ | 0 | 1 | 2 | 2 |
| $U_6$ | 2 | 0 | 1 | 0 |
| $U_7$ | 0 | 1 | 2 | 2 |
| $U_8$ | 1 | 3 | 3 | 3 |

The feature profile of user $U_5$:

$$P(U_5) = \{(f_2,1),(f_3,2),(f_4,2)\}$$

# Explain coverage

Explain coverage is used to measure the quality of justification

User-Feature matrix(P):

| | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| $U_1$ | 1 | 0 | 0 | 0 |
| $U_2$ | 1 | 1 | 3 | 3 |
| $U_3$ | 2 | 0 | 1 | 0 |
| $U_4$ | 1 | 2 | 2 | 2 |
| $U_5$ | 0 | 1 | 2 | 2 |
| $U_6$ | 2 | 0 | 1 | 0 |
| $U_7$ | 0 | 1 | 2 | 2 |
| $U_8$ | 1 | 3 | 3 | 3 |

Item-Feature Matrix (F):

| | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| $I_1$ | 1 | 0 | 0 | 0 |
| $I_2$ | 0 | 1 | 1 | 0 |
| $I_3$ | 0 | 0 | 1 | 1 |
| $I_4$ | 0 | 1 | 0 | 1 |
| $I_5$ | 0 | 1 | 1 | 1 |
| $I_6$ | 1 | 0 | 1 | 1 |
| $I_7$ | 1 | 0 | 1 | 0 |

$$\frac{\sum_{\forall (f_i, c_{f_i}) \in J(u)} \min\{c_{f_i}, P(u, f_i)\}}{\sum_{\forall f_i \in F} P(u, f_i)}$$

$$\frac{\#\, features \quad in \quad the \quad rec. \quad list}{\#\, tot. \quad features}$$

$I_2$ and $I_4$ for $U_8$:
(0+2+1+1)/(1+3+3+3)= 4/10

$I_5$ and $I_6$ for $U_8$:
(1+1+2+2)/(1+3+3+3)= 6/10

# Problem description

1. Rating profile
2. Item profile
3. Feature profile
4. Explain coverage

# Proposed Approach

1. Grouping users and items
2. Feature weighting
3. Neighborhood formation
4. Generation of recommendation and justification lists

# Grouping users and items

| | $I_4$ | $I_2$ | $I_6$ | $I_5$ | $I_3$ | $I_1$ | $I_7$ |
|---|---|---|---|---|---|---|---|
| $U_3$ | - | - | - | 2 | 2 | 4 | 5 |
| $U_6$ | 1 | 1 | - | - | - | 5 | 3 |
| $U_1$ | - | - | - | 1 | 2 | 5 | - |
| $U_5$ | 2 | 2 | 1 | 5 | 4 | - | - |
| $U_7$ | - | 2 | 1 | 4 | 5 | - | - |
| $U_2$ | 1 | - | 3 | 4 | 4 | 2 | - |
| $U_8$ | 5 | 4 | 3 | 4 | - | 1 | - |
| $U_4$ | 4 | 3 | 5 | - | 1 | - | 2 |

$b_1$: $U_{b_1} = \{U_3, U_6, U_1\}$, $I_{b_1} = \{I_1, I_7\}$

$b_2$: $U_{b_2} = \{U_5, U_7, U_2, U_8\}$, $I_{b_2} = \{I_5, I_3\}$

$b_3$: $U_{b_3} = \{U_2, U_8\}$, $I_{b_3} = \{I_6, I_5, I_3\}$

$b_4$: $U_{b_4} = \{U_8, U_4\}$, $I_{b_4} = \{I_4, I_2, I_6, I_5\}$

xMotif algorithm is used to find bi-clusters in the user and item data

# Feature weighting

1. Feature frequency (FF) is the number of times a feature occurs in a user's profile. Here FF (u,f)= P(u,f)

2. User frequency (UF) is the number of users in which the feature occurs at least once.

3. Inverse user frequency (IUF) is: $IUF(f) = \log\left(\frac{|U|}{UF(f)}\right)$

4. Calculate weighted value of feature f for user u as: W (u, f) = FF(u,f)* IUF (f)

5. Finally- generate a profile for each bicluster.

# Neighborhood formation

1. Find k biclusters closest to u.

2. The two choices for similarity measure are:

$$\mathrm{sim}_I(u,b) = \frac{\sum_{\forall i \in \mathcal{I}_u \cap \mathcal{I}_b} R(u,i) R_B(b,i)}{\sqrt{\sum_{\forall i \in \mathcal{I}_u \cap \mathcal{I}_b} R(u,i)^2} \sqrt{\sum_{\forall i \in \mathcal{I}_u \cap \mathcal{I}_b} R_B(b,i)^2}} \qquad \mathrm{sim}_F(u,b) = \frac{\sum_{\forall f \in \mathcal{F}_u \cap \mathcal{F}_b} W(u,f) W_B(b,f)}{\sqrt{\sum_{\forall f \in \mathcal{F}_u \cap \mathcal{F}_b} W(u,f)^2} \sqrt{\sum_{\forall f \in \mathcal{F}_u \cap \mathcal{F}_b} W_B(b,f)^2}}$$

$\mathrm{sim}_I$ captures rating behavior and can accurately predict which items will be rated positively by the user.

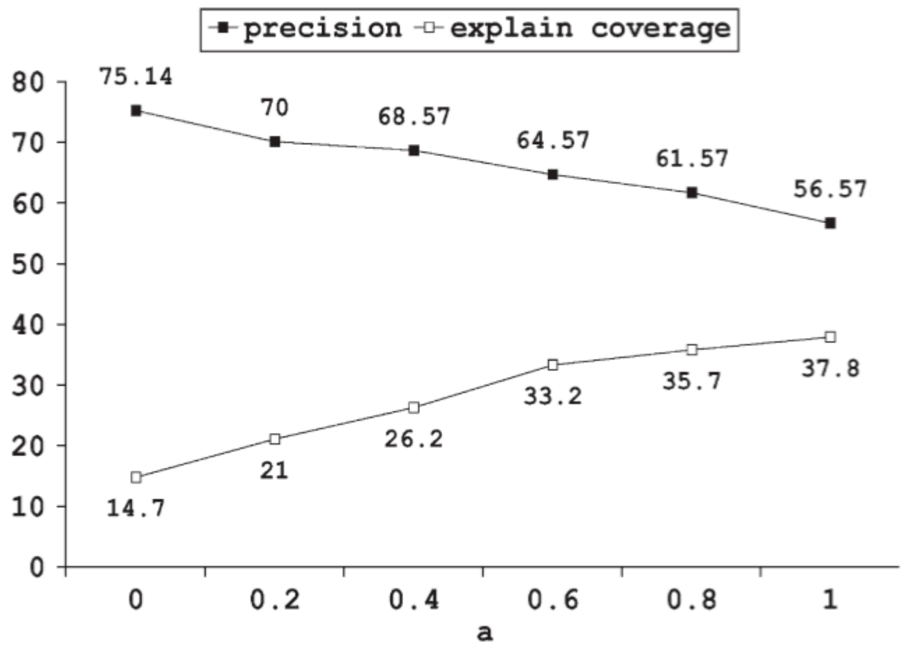
$\mathrm{sim}_F$ captures the feature profile of the users and increases the explain coverage of justification. We can combine the two as:

$$\mathrm{sim}(u,b) = (1-a) \cdot \mathrm{sim}_I(u,b) + a \cdot \mathrm{sim}_F(u,b)$$

# Generation of recommendation and justification lists

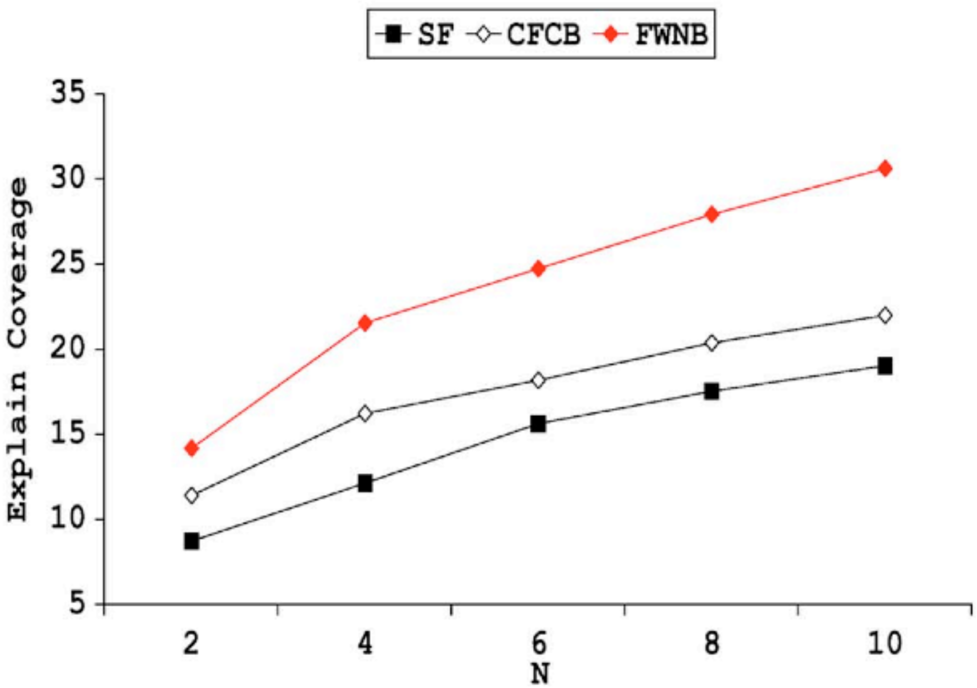
1. Identify the items in the user's neighborhood that are highly preferred by other users.
2. Contain significant features according to the weighted bicluster feature profile.

# Results



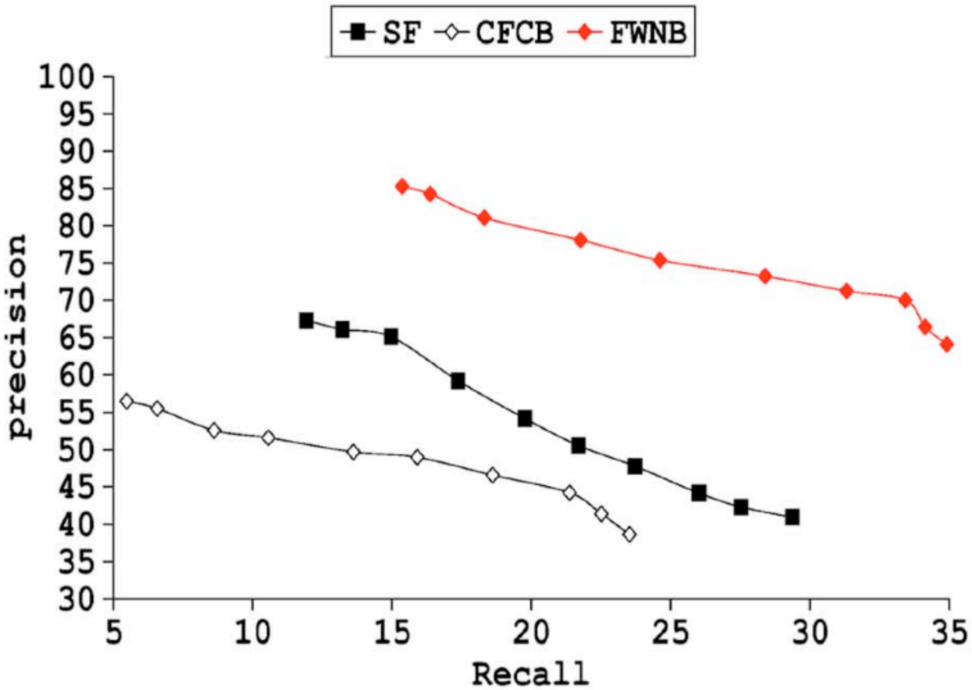

As the value of ‘a’ increases the explain coverage increases and the precision decreases. This is due to the way we combine $sim_I$ and $sim_F$.

# Results




Explain coverage increases with N. Note that the proposed method has highest explain coverage throughout.

# Results



(b)

The proposed algorithm has both higher precision and higher recall.

Precision = # relevant recommendations / #recommended items

Recall = #relevant recommendations / #total relevant items for the user

Thank You