

A HYBRID MATCHMAKING ENGINE USING ADVANCED RECOMMENDER SYSTEM

Under the guidance of Dr. Pramod Kumar Singh

K. Shashank (2014IPG-47)

P. Akhilesh Naik (2014IPG-111)

Jk. Snehith (2014IPG-121)

October 29, 2017

ABV INDIAN INSTITUTE OF INFORMATION TECHNOLOGY AND MANAGEMENT

- Why we choose this topic ?
- Recommendation systems in e-commerce.

- Defining recommender system.
- Content based, Collaborative filtering
- A glance of Proposed recommender system.

OBJECTIVE

- The objective of the following study is to bring contribution to the efficiency of recommender systems used in matrimonial sites that are well used in online dating sites in other parts of world
- Enable users to get their match in a more accurate way using content based scoring algorithms and reciprocal recommendations.
- To solve the cold-start problems by being able to provide recommendations immediately based on new user's profiles and preference.
- In the later part we use collaborative filtering which uses the interactions of the similar active users, including the people they like/dislike and are liked/disliked by, to result in recommendations.

LITERATURE REVIEW

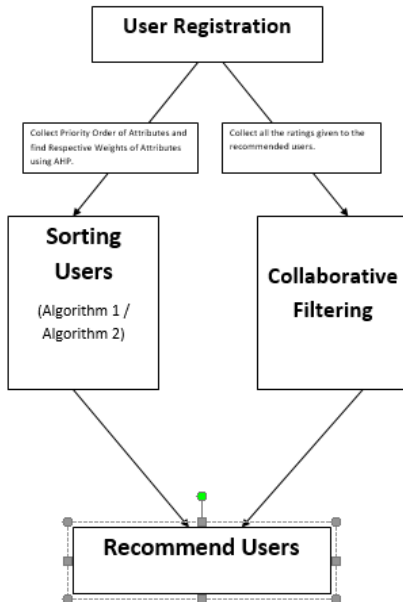
contributors	Features
Recon systems	<ul style="list-style-type: none">· Combination of content-based RS and reciprocal algorithm· User profile and interactions with other user as parameters - implicit preferences· Matching with complementary profiles
Brozosky, petrcek	<ul style="list-style-type: none">· Two recommender systems developed based on collaborative filtering

contributors	Features
K. Joshi and S. Kumar	<ul style="list-style-type: none">· Use of AHP in matrimony to calculate weights of different attributes
Sampath.M.K	<ul style="list-style-type: none">· Use of collaborative filtering in matrimonial sites
S. Khavate and M. A. Potey	<ul style="list-style-type: none">· Use of partial fuzzy scoring system for indian matrimony

MATRIMONIAL SITE AND DOMAIN OVERVIEW

- The matrimonial site - Its intent, purpose and the requirement of the user.
- Creates a profile for himself/herself or the person for whom he/she would want to find a match.
- User's Profile are further divided into account information, contact information, personal profile and preferred partner's profile.
- Following the completion of user profile now partner preferences are taken respectively. On completion the user becomes an active user of the site.
- The preferred partner profile consist of user's excepted range of age, range of height, range of income, religion, caste, education, occupation, complexion, body-type, diet, smoke and drink.

- Now recommendations are given to the active user based on the partner's preferences.
- Some of the attributes preferences may have **null values**.



THE CONCEPT OF AHP

- Complex multi-criteria decision making framework
- HOW IT WORKS?
- Pair-wise comparison
- MAKING COMPARISON MATRIX
- Pair-wise comparisons are analyzed
- Spectrum of values from (1 to 9) and ($\frac{1}{9}$ to $\frac{1}{2}$) are available
- $A_{ij} = 1/A_{ji}$

- PRIORITY AND EIGEN VECTORS
- Normalized eigen vector is known as priority
- Each value in the matrix is divided by the sum its column.
- Normalized eigen vector is obtained by Averaging the rows.
-

- HOW TO COMPUTE AHP IN FULL HIERARCHY

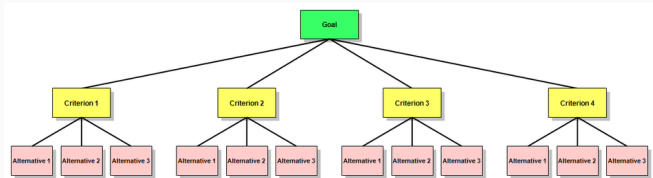


Figure: AHP HIERARCHY

$$W_i = W_i' * W_{\text{PriorNode}}$$

Algorithm takes the the priority order of the user as a, b, c, d which technically mean $a > b > c > d$ then the corresponding comparison matrix of size 4×4 is as follows :

Table 2.1: Comparison matrix.

	a	b	c	d
a	1	2	3	4
b	1/2	1	2	3
c	1/3	1/2	1	2
d	1/4	1/3	1/2	1

corresponding weights:

$\text{Weight}_a = 0.46$, $\text{weight}_b = 0.27$, $\text{weight}_c = 0.15$, $\text{weight}_d = 0.09$

SCORING ALGORITHM-1

- Classification of attributes into High important and Low important groups
- High importance attributes: Religion, Caste, Occupation, Diet, Smoke, Drink
- Low importance attributes: Age, Height, Education, Income, Body Type Complexion
- Shortlisting of recommendations

$$\text{SimilarityIndex(SI)} = 1 - \text{dist} * \text{factor}$$

- **dist** is distance between preference and profile value
- **factor** = $1/\text{no of possible values for that attribute}$
- For example, if user **A**'s **Complexion** preference is **Wheatish**

$$\text{factor} = 1/4 = 0.25$$



Table: SI of complexion attribute for user A

User	Complexion	Similarity
C	Wheatish	$1-0=1$
B	Dark	$1-1*0.25=0.75$
D	Fair	$1-1*0.25=0.75$
E	Very Fair	$1-2*0.25=0.5$

- Calculating SI for all Low importance attributes

$$\text{Score} = \sum_{i=1}^6 \text{similarity}_i * \text{weight}_i$$

- Recommendations in the descending order of scores

SCORING ALGORITHM-2

Here firstly we intake

- User Profile
- User Preferential List
 - **Most Preferred:** Diet: Vegetarian
 - **Preferred:** Diet: Eggitarian
 - **Least Preferred:** Diet: Non Vegetarian
- User Priority Order
- Order of Criteria

USER PROFILE

	Educational Qualification	Occupation	Earnings	Family status	Location	Age	Height	Complexion	Body	Diet
2	B.E	Banking	4	Middle class	Rural	24	5'7	Fair	Slim	Eggitarian
3	MBA	Logistics	6	Rich class	Urban	27	5'3	Very Fair	Heavy	Vegetarian
4	MCA	Software	4	Middle class	Semi Urban	26	5'2	Whitish Brown	Slim	Vegetarian
5	MD/MS	Teaching	7	Middle class	Semi Urban	29	5'2	Dark	Athletic	Eggitarian
6	BAMS	Looking for a job	0	Middle class	Rural	25	5'8	Very Fair	Slim	Vegetarian
7	Chartered accountant	ML/LLM	0	Middle class	Rural	25	5'8	Very Fair	Slim	Non Vegetarian

Figure: User Profiles

USER PREFERENTIAL LIST

	Educational Qualification	Occupation	Earnings	Family status	Location	Age	Height	Complexion	Body type	Diet
M P	M.com	Software	6,7	Rich class	Metro	21,22,23,24,25	5'7,5'8,5'9	Very Fair, Fair	Slim	Vegetarian
P	Chartered Accountant	Businessman	4,5	Upper middle class	Urban, Semi Urban	19,20,26,27	Anything	Whitish, Whitish Brown	Athletic	Eggitarian
L P	Anything	Anything	0,1,2,3	Middle class	Rural	18,28, 29	Anything	Dark	Heavy	Non Vegetarian

Figure: User Preferential List

USER PRIORITY ORDER AND ORDER OF CRITERIA...

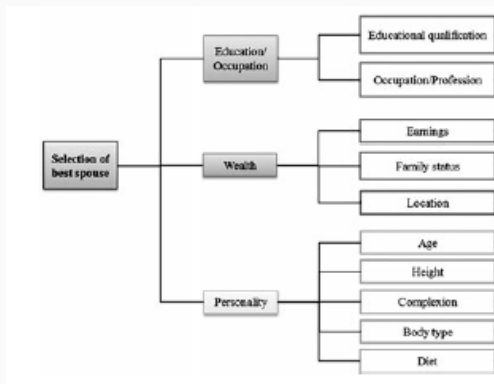


Figure: HIERARCHY OF ATTRIBUTES

- (Education, Wealth, Personality)
- Educational Qualification, Occupation, Earnings, Family status, Location, Age, Height, Complexion, Body type, Diet

Levels are introduced

- Level 1 (1, 100, 10000)
- Level 2 (1, 10, 100)
- Level 3 (1, 4, 16)
- Level 4 (1, 2, 4)

Level X $(1, \alpha, \beta)$

- If $\text{Value}(i_{\text{bride}}) = \text{Value}(i_{\text{bridegroom}})$ in
- **Most Preferred Category** : $\text{Score}_i = \beta * W_i$
- **Preferred Category** : $\text{Score}_i = \alpha * W_i$
- **Least Preferred** : $\text{Score}_i = 1 * W_i$

$$\text{Score} = \sum_{i=1}^{10} \text{score}_i$$

- Level 1 (1, 100, 10000)
- Most Preferred (MP)
- Level 2 (1, 10, 100)
- Most Preferred + Best(Preferred)
- Level 3 (1, 4, 16)
- Most preferred + Better(Preferred) + Best(Least Preferred)
- Level 4 (1, 2, 4)
- MP + Almost(Preferred) + Better(Least Preferred)

SCORING ALGORITHM 2 CONT...

- 'Age': 0.06, 'Family status': 0.08, 'Body type': 0.016, 'Diet': 0.01, 'Height': 0.04, 'Educational Qualification': 0.35, 'Complexion': 0.02, 'Location': 0.04, 'Occupation': 0.17, 'Earnings': 0.16

Table: Sample scores of candidates

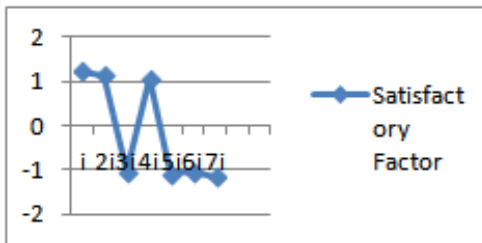
	User 1	User 2	User 3
MP	Diet	0	0
P	0	10	EQ
LP	9	0	9

- Level 1: $1 > 2 > 3$
- Level 2: $2 > 1 > 3$
- Level 3: $2 > 3 > 1$

SCORING ALGORITHM 2 CONT...

$$(\text{Satisfactoryfactor})_i = a_i/b_i$$

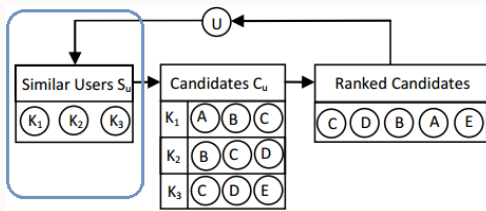
- a_i = Max (No. of liked users in interval I , No. of disliked users in the interval I)
- b_i = Min (No. of liked in the interval I , no. of disliked users in the interval I)
- i = interval
- $\Sigma(\text{satisfactory factor})_i > 0$



COLLABORATIVE ALGORITHM

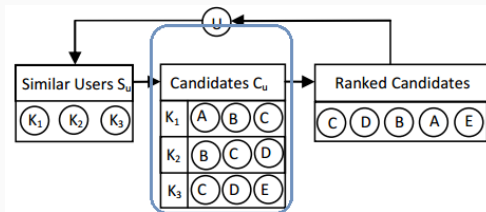
- Generating similar users based on user preferences
- Generating candidates to be considered for recommendation based on user interaction
- Assigning scores to candidate users

GENERATING SIMILAR USERS



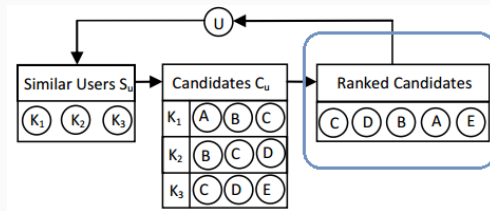
- A list of K most similar users S_u to U in active user's gender group
- Based on distance metrics
- Fill the list with users whose preferences are similar to active user

GENERATING CANDIDATE USERS



- Candidates in unordered fashion
- Reciprocal interest
- For every user in similar users list (S_u), we retrieve the all the users that he/she has reciprocal interest with

SCORES TO CANDIDATE USERS



- A candidates might have received positive interest from more than one S_u user (multiple reciprocity)
- For each candidate user X in the candidate list, we calculate the number of times X has a positive interaction with a user in similar list S_u
- Similarly negative interaction
- Score = positive - negative

Table: Sample scores of candidates

Candidate	Positive responses	Negative responses	Score
A	2	7	-5
B	6	5	1
C	12	6	6
D	6	2	4

- Higher the score, more he is liked reciprocally
- Recommended in the descending order of scores

RESULTS AND DISCUSSION

- Implementation of algorithm on sample data
- The LOW importance preferences of a prospective bride is as follows: age - 24 to 26, height - 5.9 to 6, education - Diploma, annual income - 8 lakhs to 10 lakhs., body type - Average, complexion - Wheatish.

Table: Profiles of active bride grooms

U-ID	Age	Height	Education	Income	Body	Complexion
21	26	5.7	Masters	4L	Athletic	Dark
22	24	5.5	Doctors	7L	Slim	Dark
23	30	5.1	Doctors	14L	Avg	Wheatish
24	22	4.10	HSC	8L	Avg	Fair
25	26	5.2	Bachelors	3L	Hvy	VFair

- Users are recommended in the decreasing order of scores

Table: Recommendations for Active Bride

User-ID	Score
21	49.3
24	48.4
22	46.45
23	40.2
25	44.65

- Level 1:

$$3 > 4 > 6 > 5 > 7 > 2$$

- Level 2:

$$3 > 4 > 7 > 5 > 2 > 6$$

- Level 3:

$$3 > 4 > 7 > 2 > 5 > 6$$

- Level 4:

$$3 > 4 > 7 > 2 > 5 > 6$$

CONCLUSION

Our partially developed algorithm have shown that the two proposed recommender systems not only gives a new way of gaining information from users but also explains the mutual acceptance by providing matches with much similarity between them and not only are we providing N-recommendations but the order of each match that is provided to active user is taken into account, its an improvement in the field known less important.

THANK YOU