

Designing Personalized Recommendation in E-Commerce Site Based on Content-Based and Collaborative Filtering

Ayad R. Abbas¹ and Shaymaa Ashor^{2*}

¹Department of Computer Science, University of Technology, Baghdad-Iraq.

²Department of Computer Science, University of Technology, Baghdad-Iraq.

*Corresponding Author: shaymaashor1111@gmail.com

Abstract

After the great development in the use of the internet and widespread use of e-commerce sites, the problem of the large number of products offered appeared, thus causing confusion for customers in choosing the suitable products for their needs. To overcome this problem, e-commerce sites using the recommendation systems to nominate products to customers personally according to their needs. The proposed system combines two types of recommendation systems they are content-based filtering which uses cosine similarity function to find the similarities between the products described by texts and collaborative filtering, which uses correlation similarity function to nominate products similar in the evaluation to the products that have been previously purchased and evaluation by customers.

Several problems are solved by the proposed system in a simple way and at the same time with high efficiency and accuracy, such as the cold start, scalability, synonymy and spared the data problem. [DOI: [10.22401/JNUS.20.2.19](https://doi.org/10.22401/JNUS.20.2.19)]

Keywords: E-commerce, Personalized Recommendation System, Content-based, cosine similarity, correlation similarity.

Introduction

The vast improvement in the internet area was met with a huge improvement in the e-commerce area. Those sites have been established in a way to play one of the lead roles in people's lives by introducing different products, this thing has caused confusion to the customers in deciding what products suit them, and people who especially face that are customers who don't have practice. This way, a pressing need appeared in the system which makes it simpler for the customer to decide what product suits him best. [1]

E-commerce web sites used recommendation and suggestion methods significantly in different applications in order to attract the highest possible number of customers. By this boosting their income because the information given to the customers prior or during the period of purchase can directly influence the consumer's purchase choice. [2]

Recommendation methods are software programs spreads highly utilized by Web sites. It is used to filter huge amounts of information for suggesting suitable items to the customers. Those items may be movies, news, books, documents, games, music, web pages or products, etc. [3] [4]

There are different types of recommendation systems, the most frequently used are: content-based filtering (CBF), Collaborative filtering (CF) which divides into two types (user-based, item-based), demographic, knowledge-based filtering and finally hybrid approach which combines two or more recommendation systems to address the problems that exist in these systems, such as cold start, scalability, sparsity and Gray sheep problem that appear in the CF or the problem of lack of diversity in the presentation of products suffered by CBF [3-6]. Since any system must take into account two basic points, cost and benefit from the system. [7]

All RSs need information about the customers, items or both of them, for producing a table of recommendations [18].

Fig.(1) illustrates RSs that use feedback methods which are divided into the three types: Explicit feedback Where it is In this method the customers evaluate items explicitly through the interface system after the system asks for the customer to do that, for improving the suggested items, Implicit feedback where it is In this method information is gathered without the customer's interference through observing a customer's behavior by the

system, the last one is Hybrid feedback where it is In this method integration between the former two types. This information about the user's interests is stored in the users' profiles. Furthermore, it is useful for submitting

recommendations and suggestions then they are processed by the RS in generating a table of recommendations. [19]

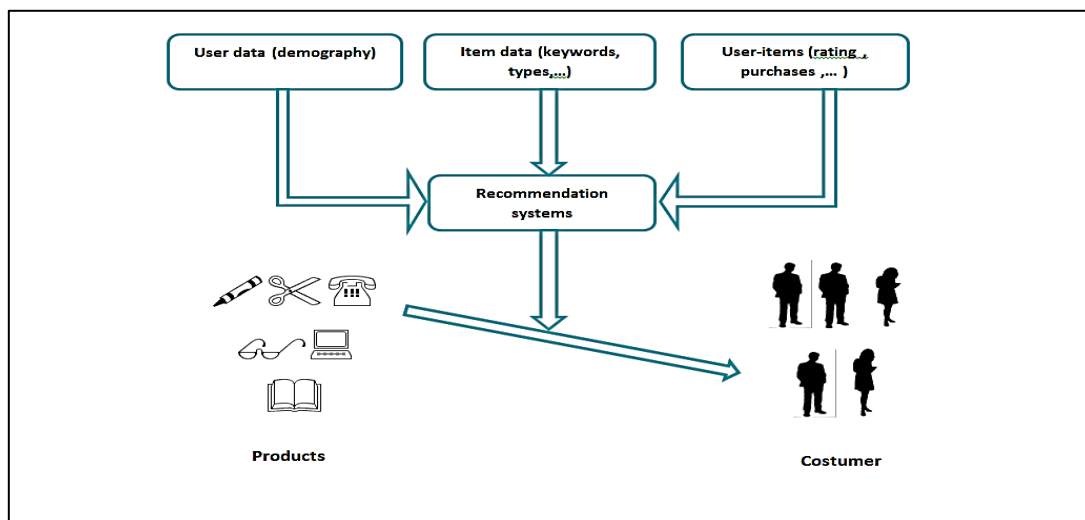


Fig.(1): The recommendation process.

The proposed system integrates both CBF and item based CF in order to address existing problems. CBF is use cosine-based similarity function in order to find the similarity between two products that described in the data base by texts. If the vector $\vec{i}=\{x1, y1\}$, vector $\vec{j}=\{x2, y2\}$ then the resemblance between a product (i) and a product (j) might be depicted in the equation below:

$$S(\vec{i}, \vec{j}) = \cos(\vec{i}, \vec{j}) = \frac{\vec{i} \cdot \vec{j}}{\|\vec{i}\| \times \|\vec{j}\|} = \frac{x1x2+y1y2}{\sqrt{x1^2+y1^2} \sqrt{x2^2+y2^2}} \dots\dots\dots (1)$$

While in the item-based CF, customers (c) ranked product (i) and product (j), then the pearson's correlation will be calculated by:

$$s(i, j) = \frac{\sum_{c \in C} (r_{c,i} - \bar{r}_i)(r_{c,j} - \bar{r}_j)}{\sqrt{\sum_{c \in C} (r_{c,i} - \bar{r}_i)^2} \sqrt{\sum_{c \in C} (r_{c,j} - \bar{r}_j)^2}} \dots\dots\dots (2)$$

So, $r_{c,i}$ is the ranking of customers c on the product (i), \bar{r}_i is the mean ranking of the ith product by every customer, C is a Collection of customers who ranking products i, j together. [8-10]

Obtaining Prediction is the most vitally crucial stages in a CF system. For item-based CF prediction, the prognosis on a certain product (i) for a customer (c) can be calculated by utilizing the summation of the rankings of the customer to products weighted by

Similarity between various products, as the following formula:

$$P_{c,j} = \frac{\sum_{i \in I} r_{c,i} \times s(i,j)}{\sum_{i \in I} |s(i,j)|} \dots\dots\dots (3)$$

Where the summations are overall other rated products $i \in I$ for customer c, S(i,j) is the similarity between products i and j, $r_{c,i}$ Is the ranking for customer c on product i. [20-21]

1.2. Related works

Many studies have been made to address the problems of RS in several respects such as:

1. [13]. The author used web usage mining to generate a dynamic RS in real time to customers in e-commerce sites, regardless of whether the customers are registered or unregistered. Web logs are analyzed in order to predict the behavior of customers and their interests. Item based recommendation technique is used for unregistered customer and user based recommendation technique is used for registered user. Nomination for the accuracy of that system for registered customers ranged from 80-85, but for unregistered customers ranged from 65-70.
2. [14]. The author integrates both CF and Demographic RS to solve the problem of

cold start, it is one of the challenges faced by CF.

Therefore the personal information of each of the new customer and customers' existing in the database (DB) is compared in order to find the similarity between them. K-means clustering algorithm is the data mining technique used by CF and Demographic RS to find similarity between users.

3. [15]. The author integrates each of the user-based CF and item-based CF at the same time, then combines two predicted results from both item-item similarities and user-user similarities in order to address the sparsity of data problem and increase accuracy.
4. [16]. The author combines CF and Non-personalized in order to propose products to customers. Non-personalized nominated products to all customers in general, without exception, are nominating products that have the highest average ratings. While CF, used Pearson-correlation algorithm for user-based collaborative techniques to measure the similarity between active user and other users.
5. [17]. The author combines between the two techniques, CF and knowledge based recommender system. The new customer enters his needs without the need for a previous customer record using a simple model, while using three approaches for submission of proposals to customers who are already registered in the system. The first recommended approach depends on a user's history to determine customer's favorites by submitting suggestions similar to items that the customer is looking for previously. The second Recommendation approach suggests suitable products to the customer based on finding similarities in interests between the active customer and other customers registered in the system. The last type, Recommendation approach submitted suggestions to the customer based on customers' previous preferences for the kitchen and location to propose appropriate restaurant to the customer.

Structuring of the Proposed System

The proposed system integrates the two types of RS: CBF and item-based CF to build an efficient system that addresses most of the problems of RS.

It consists of two main stages, the first stage called CBF stage, which especially for new customer (unregistered) who does not have purchasing records in the DB, while the second stage is for the customer who already registered in the system means he/she has a previous purchasing records.

In the first stage the proposed system uses CBF to nominate products to new customers. CBF in the proposed system used categories of products to nominate products. The categories of products described by a string of words (e.g. Office Products, Office & School Supplies, Paper, Photo Paper), which are similar to the products chosen by the customer in the past. It capable of dealing with the problem of new products that are not rated by any customer till now

In the second stage called CF stage, which the system uses the item-based CF to nomination products to customers registered in the system and who have previously record purchase in the DB.

E-commerce sites such as Amazon prefer to used item-based CF instead of user-based CF because the number of products is less than the number of customers, so that find the similarities between the products is faster and less expensive than finding similarities between customers.[3][11]

The following steps illustrate the general steps of the proposed system, see Fig.(2):

Step_1: The first stage begins when the new customer signs up to the system, then login using his/her name and password. After that the system will allocate an identification number (ID) to the customer and store all customer information in the customer profile DB.

Step_2: The customer selects a particular product from the products list, and then the system uses CBF to generate a list of the recommendations to the customer.

Step_3: The CBF connects to the DB to extract product information, in order to calculate the similarity degree between the

current product chosen by the customer and the products existing in the DB.

Step_4: The CBF generates a list of recommendations of products which get high percentage of similarity with a product that has already been selected by a new customer.

Step_5: The system provides a list of recommendations to the new customer.

Step_6: The customer chooses one product from a recommended list according to his/her preferences, then the system imposes on the customer to rate the product that has been purchased. After that, the system will update both the customer profile DB and the ratings products file.

Step_7: when the customer previously registered in the system using his name and password, which is used to register in the system in the first time sign up into the system. The system uses this information (name, password) to get a customer ID from the customer profile.

Step_8: The system uses CF to generate a new list of the recommendations to the customer Based on customer's purchasing history.

Step_9: The CF connects to the DB to extract information from customer profile DB and products rating file, in order to calculate the correlation degree between the product chosen by the customer previously and the products existing in the DB.

Step_10: The CF generates a recommendations list, which gets a high percentage of correlation between the products that has already been selected by a registered customer.

Step_11: The system provides a recommendations list to the registered customer.

Step_12: The customer chooses one product from a recommended list according to his/her preferences, then the system imposes on the customer to rate the product that has been purchased. After that, the system will update both the customer profile DB and the ratings products file, Similar to Step 6.

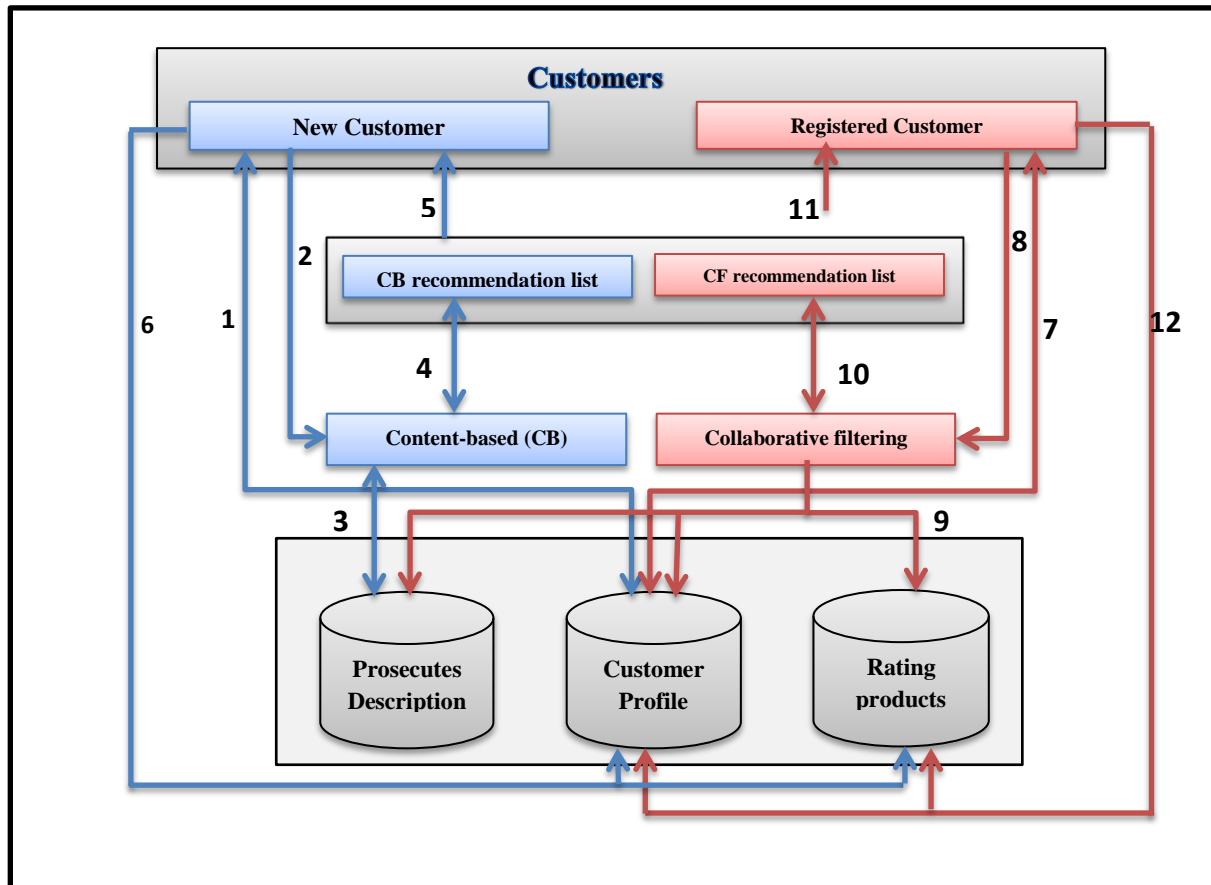


Fig.(2): General steps of the proposed system.

Experimental Result and Evaluation

A-The Dataset:

The data set contains metadata of 320 Office Products, In addition to 500 customers buy different products from these products and have rated it. [12]

B- CBF Stage

For example, when Customer selects a particular product such as that (Classroom Decoration) from the list of product the following points are used to calculate the similarities among products:

1. The proposed system extracts ID of the selected product from the product profile in the DB in order to use it to extract the main category, which is for the selected product (Office Products, *Office & School Supplies*, *Education & Crafts*, *Arts & Crafts Supplies*, *Classroom Decorations*).
2. Convert main category to the array of token (x) = (*Office*, *Products*, *Office*, *&*, *School*, *Supplies*, *Education*, *&*, *Crafts*, *Arts*, *&*, *Crafts*, *Supplies*, *Classroom*, *Decorations*).
3. For each sub_categories extract the main categories for it and convert it to the array of token (y_i).

$i = 1 \dots n$, where n is the number of sub_categories.

For example: The other *sub_category* = (*paper*).

The *main category is* = (*Office Products*, *Office & School Supplies*, *Paper*), convert to array of token

(y_i) = (*Office Products*, *Office*, *&*, *School*, *Supplies*, *Paper*).

4. Integrate array (x) with array (y_i) in a unique array (k) without repeating the type of words.
 $K =$ (*Office*, *Products*, *&*, *School*, *Supplies*, *Education*, *Crafts*, *Arts*, *Classroom*, *Decorations*, *Paper*).
5. Generate vector (A) by searching for each value of the array (x) within the array (k) if the value exists put one in vector (A) else put zero in vector (A).
 Vector $A = x$ (1,1,1,1,1,1,1,1,1,1,0).
6. Repeat step5 for each type of product to generate vector (B).
 Vector $B = y$ (1,1,1,1,1,0,0,0,0,0,1).
7. Cosine similarity is used to measure the similarity between vector (A) and vector (B) as follows:

$$\vec{A} \cdot \vec{B} = 1 \times 1 + 1 \times 1 + 1 \times 1 \dots \dots \dots = 5$$

$$\|\vec{A}\| = \sqrt{1^2 + 1^2 + 1^2 \dots \dots \dots} = 3.162$$

$$\|\vec{B}\| = \sqrt{1^2 + 1^2 + 1^2 \dots \dots \dots} = 2.449$$

$$S(\vec{A}, \vec{B}) = \frac{5}{3.162 \times 2.449} = 0.645$$

Consequently the similarity between the selected product and all other categories is displayed in Table (1).

Table (1)
The results of using the cosine similarity function.

Product name	Similarity
Appointment Books & Planners	0.47
Class Records & Lesson Books	0.66
Wide-format Paper	0.55
Copy & Multipurpose Paper	0.55
Inkjet Printer Paper	0.25
Photo Paper	0.59
Dry Erase Boards	0.52
Impact & Dot Matrix Printer Ribbons	0.28
Pins & Tacks	0.50
Rubber Bands	0.50
Correction Tape	0.52
Erasable Markers	0.50
Mechanical Pencil Leads	0.47
Wooden Colored Pencils	0.50
Ballpoint Pens	0.50
Compasses	0.50
Protractors	0.50
Printers	0.38
Conference Room Tables	0.33
Calculators Basic	0.59
Laser Printers	0.35
Lateral File Cabinets	0.30
Mobile File Cabinets	0.30
Vertical File Cabinets	0.30
Racks & Displays	0.31
Shelving & Storage	0.30
Storage Cabinets	0.30
Lecterns & Podiums	0.31
Mail Carts	0.33
Desk Chairs	0.33
Stools	0.33
Task Chairs	0.33
Computer Workstations	0.33
Office Desks	0.35
Bulletin Boards	0.55
Inkjet Printer Ink	0.33
Pencil Sharpeners	0.52
Mechanical Pencils	0.48
Desktop Staplers	0.52
Liquid White Glues	0.45
Store Sign Holders	0.50

8. If the result of similarity ≥ 0.48 and ≤ 0.98 , the ID of the sub_categories adds to the array.

9. Then the system uses the ID of sub_category to recommend one product for each sub_category.

C- The CF Stage

This stage begins when the customer who already registered in the system requested sign in to the system.

The system nominates similar products by rating to the old purchased products using item_based CF technique. Also, the system displays the products that the customer has to purchase in the past in the same page.

In the proposed system the item-based technique uses Pearson's Correlation in order to compute correlation between similar products, as it is shown in equation (2).

The prediction on a certain item (i) for a customer (a) can be calculated by using the equation (3).

For example

If the customer who has the ID number (51), bought the product which has the ID number (140), then gave the product rating is given (5).

To nominate suitable products to the customer, the proposed system performs the following calculations:

Calculate the correlation between the products

1. Calculate the average rating for each customer who bought the product (140), with the exception of active customer (51), as it is shown in the following table.

Table (2): shows ID for the customers who bought the product (140), in addition to the rating of their selected product. The table below starts from customer ID 7 and ends with customer ID 3739.

Table (2)
Customer's Average Rating.

No	Customer ID	Rating	No	Customer ID	Rating
1	7	1	15	2398	4
2	17	1	16	2562	4
3	184	4	17	2853	3
4	375	4	18	2879	1
5	615	5	19	2889	5
6	623	5	20	2957	5
7	726	5	21	3205	5
8	1110	5	22	3301	4
9	1240	3	23	3485	4
10	1264	4	24	3537	4
11	1963	5	25	3578	4
12	2040	3	26	3739	1
13	2363	4	Total		98
14	2379	5	Rating average		3.76923

2. Calculate rates averages for all products except that are on the list of previous purchases for active customer (51), for example, the product which has the ID number (10), as it is shown in the following table.

Table (3): shows ID for the customers who bought the product (10), in addition to the rating their choice for this product.

Table (3)
Customer Rating.

No	Customer ID	Rating
1	10	4
2	96	4
3	357	4
4	1192	4
5	1277	3
6	1641	4
7	2173	3
8	2562	5
9	2889	4
10	3205	4
11	3427	4
12	3485	3
13	3739	3
Total		49
Rating average		3.76923

3. Correlation between the producer (140) and the product (10) is calculated by the following equation:

$$s(140,10) = \frac{\sum_{c \in C} (r_{c,140} - \bar{r}_{140})(r_{c,10} - \bar{r}_{10})}{\sqrt{\sum_{c \in C} (r_{c,140} - \bar{r}_{140})^2} \sqrt{\sum_{i=1}^n (r_{c,10} - \bar{r}_{10})^2}}$$

Where

$$C = (2562, 2889, 3205, 3485, 3739)$$

$$\bar{r}_{140} = 3.76$$

$$\bar{r}_{10} = 3.76$$

$$s(140,10) = \frac{(4-3.76)(5-3.76)+(5-3.76)(4-3.76)+(5-3.76)(4-3.76)+\dots}{\sqrt{(4-3.76)^2+(5-3.76)^2+\dots} \sqrt{(5-3.76)^2+(4-3.76)^2+\dots}}$$

$$s(140,10) = 0.498$$

Notice from the above equation, the similarity between the products (140.10) is calculated only to customers who bought producers together, who have the following ID numbers (2562, 2889, 3205, 3485, 3739). Thus, the process is repeated between the producer (140) and the remaining the products in the DB, as shown in the following table.

Table (4)
Illustrates the results of calculating correlation between the product (140) and all products in the DB.

ID (j)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
S (140,j)	0	0	0	-1	0	0	1	0	0	0.49	0	0	0	0	0	0
ID (j)	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
S (140,j)	1	0	0	0	0	0	-1	0	1	0	0	0	0	0	0	0
ID (j)	33	34	35	36	37	38	39	40	41	42	43	45	46	47	48	49
S (140,j)	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
ID (j)	50	51	52	53	54	55	56	57	59	60	61	62	63	64	65	66
S (140,j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82
S (140,j)	0	0	0.99	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98
S (140,j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
ID (j)	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114
S (140,j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130
S (140,j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146
S (140,j)	0	0	0	0	0	0	0	0	-0.97	0	0	0	0	0	0	0
ID (j)	147	148	149	150	151	152	153	154	155	156	158	159	160	161	162	163
S (140,j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	165	166	167	168	169	170	172	173	174	175	177	178	179	180	181	182
S (140,j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1
ID (j)	183	184	185	186	188	189	190	191	192	193	194	195	196	197	198	199
S (140,j)	0	-1	0	0	0	0	0	0	0	0	0	0	1	0	0	0
ID (j)	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215
S (140,j)	0	0	0	0	0	0	0	1	0	-1	0	0	0	0	0	0
ID (j)	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231
S (140,j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247
S (140,j)	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.28	0
ID (j)	248	249	250	251	252	253	255	256	257	258	259	260	261	262	263	264
S (140,j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	265	266	267	268	269	270	271	272	273	274	275	276	278	279	280	281
S (140,j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	282	283	284	285	286	287	288	289	290	292	293	294	295	296	297	298
S (140,j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314
S (140,j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0
ID (j)	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	
S (140,j)	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	

Calculating item_based prediction

The proposed system uses the average weight to predict the rating of the customer (51) on the product (140). Prediction equation will apply to all products, but in this example will be applied to the product (10) only.

$$P_{51,10} = \frac{\sum_{140 \in I} r_{51,140} \times S(140,10)}{\sum_{140 \in I} |S(140,10)|}$$

where I represents all the products in the customer's purchasing history, in this example the product (140) is the only product in customer's purchasing history. And

$r_{51,140}$ represents the rating of costumer (51) on product (140)

$$P_{51,10} = \frac{5 \times 0.498}{|0.498|} = 5$$

The following table shows the results of prediction for all products in the DB.

Table (5)
The results of prediction for all products.

ID (j)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
P _{51,j}	0	0	0	-5	0	0	5	0	0	5	0	0	0	0	0	0
ID (j)	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
P _{51,j}	5	0	0	0	0	0	-5	0	5	0	0	0	0	0	0	0
ID (j)	33	34	35	36	37	38	39	40	41	42	43	45	46	47	48	49
P _{51,j}	0	5	0	0	0	0	0	0	0	5	0	0	0	0	0	0
ID (j)	50	51	52	53	54	55	56	57	59	60	61	62	63	64	65	66
P _{51,j}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82
P _{51,j}	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98
P _{51,j}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
ID (j)	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114
P _{51,j}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130
P _{51,j}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146
P _{51,j}	0	0	0	0	0	0	0	0	-5	0	0	0	0	0	0	0
ID (j)	147	148	149	150	151	152	153	154	155	156	158	159	160	161	162	163
P _{51,j}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	165	166	167	168	169	170	172	173	174	175	177	178	179	180	181	182
P _{51,j}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-5
ID (j)	183	184	185	186	188	189	190	191	192	193	194	195	196	197	198	199
P _{51,j}	0	-5	0	0	0	0	0	0	0	0	0	0	5	0	0	0
ID (j)	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215
P _{51,j}	0	0	0	0	0	0	0	5	0	-5	0	0	0	0	0	0
ID (j)	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231
P _{51,j}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247
P _{51,j}	5	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0
ID (j)	248	249	250	251	252	253	255	256	257	258	259	260	261	262	263	264
P _{51,j}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	265	266	267	268	269	270	271	272	273	274	275	276	278	279	280	281
P _{51,j}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	282	283	284	285	286	287	288	289	290	292	293	294	295	296	297	298
P _{51,j}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ID (j)	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314
P _{51,j}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-5	0
ID (j)	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	
P _{51,j}	0	0	0	-5	0	0	0	0	0	0	0	0	0	0	0	

Thus, the system will nominate products that get a rating close to the customer evaluation of products in the purchasing history. However, according to the results obtained in this example, the products that will be nominated are (7, 10, 17, 25, 34, 42, 69, 98, 196, 207, 232,246).

Switching between Item_based CF and CBF

There are situations in which the proposed system switches or combines techniques of

item_based CF and CBF, as shown in the following points:

1. If the customer has previously registered in the system, but did not buy any product, then the system will use CBF in the recommended products to the customer according to his current choices.
2. If the customer is buying products, but these products do not have any similarity rating with products in the DB, thus the system will use the CBF technique to

- nominate products to the customer, these products are similar to the existing products in the purchasing history.
3. If the customer bought products, but these products are new products and have not been rated by the other customers, thus, the system will use the CBF technique to nominate products to the customer, these products are similar to the existing products in the purchasing history.
 4. The customer bought products, but some purchased products are new products which have not been rated by other customers, while the remainder products have been rated by other customers, thus the proposed system will merger item_based CF in order to nominate products similar to products that have been previously rated with the CBF in order to nominate products similar to new products that do not have any rating.

Conclusions

After achieving the building and implementations of the proposed system, the following conclusions are drawn:

1. The using of the CBF in the proposed system contributes to overcoming the drawbacks such as a new customer, new product, synonymy, and scalability.
2. Selecting the upper limit of the similarity degree between the products maximums to (0.98), contributes to solving the problem of nomination of the same products which have been selected by the customer previously.
3. The using of the cosine function in order to find the similarity between products at the CBF stage contributes to the nomination products in a simple and efficient manner.
4. The using of Item_based CF, contributes to solving the problem of diversity and snap in nomination products to the customers, which the CBF suffers from.
5. The using Item_based CF increases speed of the process of nomination products to registered customers because it depends on finding similarities between the products, and not between the customers, where the number of products is much less than the number of customers.
6. Merger between the CBF and item_based CF contributes to addressing the problems that exist in both types, therefore building an integrated system is characterized by the ability to nominate products in an easy and efficient manner.

References

- [1] Przemyslaw K., Pawel K., "Personalized Integration of Recommendation Methods for E_commerce", International Journal of Computer Science & Applications, Vol. 3, Issue 3, pp12-26, 2006.
- [2] Nicolas K., Maria M., "Consumer s' Interest in Personalized Recommendations: The Role of Product-Involvement and Opinion Seeking", 1530-1605/07, (IEEE), 2007.
- [3] Francesco R., Lior R., Bracha S, "Introduction to Recommender Systems Handbook", (Springer), pp 1-300, 2011.
- [4] P. N. Vijaya Kumar and V. Raghunatha Reddy, "A Survey on Recommender Systems (RSS) and Its Applications", (IJIRCCE), Vol. 2, Issue 8, August 2014.
- [5] Tejal A., R.S. Sonar, N. J. Uke, "A Survey on Recommendation System", (IJIRAE), ISSN, pp 2349-2163, V. 2, Issue 1, 2015.
- [6] Bobadilla J., Ortega F., Hernando A., Gutiérrez A., "Recommender systems survey", Volume 46, PP 109-132 (ELSEVIER), 2013.
- [7] Clare M. K., Jan O., John K., "Designing Personalized User Experiences in e Commerce", KLUWER ACADEMIC PUBLISHERS, (Volume : 5), 2004.
- [8] Vaishnavi S., Karthik S., "Design Recommender System based on customer reviews", (IJRET), Volume: 02, Issue: 11 | Nov- 2013.
- [9] Isinkaye F.O., Folajimi Y.O. Ojokoh B.A., "Recommendation systems: Principles, methods and evaluation", Egyptian Informatics V 16, pp216-273, Journal, 2015,.
- [10] Mohd A. H., Omar J., Ramachandram S., "Collaborative Filtering Based Recommendation System: A survey", (IJCSE), Vol. 4 No. 05 May 2012.
- [11] Saša B., Mirjana M. Zita B., "Choosing a Collaborative Filtering Algorithm for e-Commerce", Management Information Systems, Vol. 3, 2008.

- [12] Available at (<http://jmcauley.ucsd.edu/data/amazon>).
- [13] Prajyoti L., Bidisha R., “Dynamic Recommendation system using web usage mining for E-commerce users”, (ELSEVIER), Ramachandram Vol.45, pp 60-69, 2015.
- [14] Shano S., Shalini B., “Recommender System using Collaborative Filtering and Demographic Characteristics of Users”, (IJRITCC), ISSN: 2321-8169 Vol.3 Issue: 7, July 2015.
- [15] Gilbert B., Hazem H., Wassim H., Lama N., “A Hybrid Approach with Collaborative Filtering for Recommender Systems”, IEEE, 2013.
- [16] Anil P., Neev P., Tanvi B., Rekha S., “Non-Personalized Recommender Systems and User-based Collaborative Recommender Systems”, New York, USA, (IJ AIS), Vol. 6– No. 9, March 2014.
- [17] Purna Dwivedi and Nikita Chheda, “A Hybrid Restaurant Recommender”, International Journal of Computer Applications, Vol. 55, NO 16, October 2012.
- [18] RVVSV P., VALLI K., “A Categorical Review Of Recommender Systems”, (IJDPS), Vol.3, No.5, September 2012.
- [19] Megharani M. and Uma N., “Recommender System Methods and Feedback Mechanisms: A Survey”, (IJRCCE), Vol. 3, Issue 11, November 2015.
- [20] Xiaoyuan S. Taghi M. K., “A Survey of Collaborative Filtering Techniques”, Hindawi Publishing Corporation Advances in Artificial Intelligence, Article ID 421425, pp19, 2009.
- [21] Kumar A., Saurabh K., Vipin K., Kumar P., “A Review on Personalized Information Recommendation System Using Collaborative Filtering”, (IJCSIT), Vol. 2 (3), pp 1272-1278, 2011.