Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2014, 6(4):30-40

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

A new personalized three-dimensional recommendation approach for C2C E-commerce context

Ai Danxiang#, Zuo Hui@ and Yang Jun*

#School of Management, Guangdong University of Technology, Guangdong, China
@School of Economics and Commerce, Guangdong University of Technology, Guangdong, China
**School of Management, Guangdong University of Technology, Guangdong, China*

Abstract

Recommender systems have been viewed as powerful tools to filter overloaded information in the e-commerce environment. But traditional two-dimensional recommendation methods, which only explore the relevance between customers and products, are not applicable for the recommendation space in C2C (Customer to Customer) e-commerce context that involves three types of entities: buyers, sellers and products. In this paper, we propose a three-dimensional approach to explore the relevance among buyers, sellers and products, and provide personalized "seller and product" recommendations for buyers. Firstly, similarities between sellers are calculated based on seller features. Then the spare data in the three-dimensional historical rating set are supplemented and based on which buyer similarities are calculated to find neighbors who have similar product preferences with the target buyer. Finally, a three-dimensional rating prediction model is used to predict the unknown ratings that the buyer may give to candidate "seller and product" combinations. A real data experiment is conducted and the results prove the effectiveness of the proposed approach.

Keywords: Three-dimensional Recommendation; Recommender System; Collaborative Filtering; Content-based Filtering; Customer to Customer

---

INTRODUCTION

The prevalence of e-commerce has brought great convenience to people's daily life, but at the same time, the swift growth of data size has led to the growing problem of information overload [1]. E-commerce recommendation applications arise out of this background that analyze the customer buying behavior to find customer interest or purchase patterns, and recommend proper products or services [2-3]. Recommender systems have been widely used in B2C (Business-to-Customer) e-commerce websites, but they have limitations when applied in C2C (Consumer-to-Consumer) e-commerce context. Since the B2C shopping site acts as the sole vendor in the trades, customers need only select products but consider little about sellers. Thus most recommendation methods used in B2C are two-dimensional, for they explore the relevance of two types of entities: customers and products [4]. However, in the context of C2C e-commerce, the website is not involved in the transactions, but only plays as a third party providing an open online platform for buyers and sellers. There are often multiple sellers in one C2C shopping site. The same product is usually sold by many sellers simultaneously in the site. Buyers have to select both sellers and products. When the number of sellers is increasing, buyers are as difficult to filter the information of sellers as that of products. So the recommendation task in C2C changes from traditional "recommend the top N products to a customer" to "recommend the top N seller and product combinations to a buyer".

We believe the recommendation problem in C2C described above is three-dimensional, for it involves three types of entities: buyers, sellers and products. Traditional two-dimensional recommendation methods cannot perfectly solve this kind of problem. Additional seller selection paradigms have to be used to supplement two-dimensional recommender systems after they generate product recommendations. The simplest way is selecting sellers randomly, which means all sellers who sell the corresponding product are possible to be recommended. More commonly used way is ranking the sellers who sell the recommended product by some criterion that most buyers may concern, e.g., sales volume, credits, price etc. However, such two-dimensional recommendation combined with simple seller selection is not enough to model the complicated buyer behavior filtering both products and sellers in C2C context. For one thing, different buyers have different demands for sellers, and they usually select sellers referring to various features, not just one. So filtering sellers by a single pre-set criterion cannot reflect the personalized preferences of buyers. For another, a buyer's interest in sellers and that in products may affect each other. For example, when a buyer cannot find a satisfied seller, he may give up buying the original product and choose some substitute. So the buyer's preferences for products and that for sellers should be treated equally in the recommendation process. Currently used methods obviously give more weights to products than to sellers.

To solve the personalized recommendation problem in C2C e-commerce context, we propose a new three-dimensional approach. Our approach has two aspects of utilities: (1) synthetically considering the buyer's preferences for both sellers and products and recommending "seller and product" combinations to the target buyer. We learn the buyer's preferences for products using collaborative filtering method [5], namely, find neighbors by measuring buyer-buyer similarity based on their co-ratings of products, and recommend the products that the neighbors have rated highly in the past. The buyer's preferences for sellers is learned using context-based filtering method [6], namely, measure seller-seller similarity based on their features, and recommend the sellers that have a high degree of similarity to those the target buyer or his neighbors have rated highly in the past; (2) alleviating the influence of spare data. Data sparsity is a well-known problem in recommender system [4], and it becomes more severe in the three-dimensional recommendation space. Since there are large number of "seller and product" combinations in a C2C shopping site, the number of ratings already obtained is usually very small compared to the number of ratings that need to predict. So we design a rating inference mechanism to supplement the data set and ensure the validity of the recommendation calculations in our approach.

1. TRADITIONAL TWO-DIMENSIONAL RECOMMENDATION METHODS

Traditional two-dimensional recommendation methods can be divided into two main groups: collaborative filtering (CF) and content-based filtering (CB) [7].

CF has thus far been the most successful recommendation technique and used in a number of different applications [8]. It works by recommending items to a target user through a process of identifying neighbors (people who share similar preferences for items) and looking for those items that the target user is most likely to select [9]. Before the recommendation process, the user profile should be constructed, which is usually represented as a user-item matrix and formed from historical ratings. Table 1shows an example of the matrix. The value scope of the ratings is usually 0 to1 or -1 to 1. "NULL" means the user has not rated the item yet.

Table 1: Example of user-item rating matrix

	i_1	i_2	i_3	…	i_m
u_1	0.42	NULL	0.71	…	NULL
u_2	NULL	0.25	0.57	…	0.80
…	…	…	…	…	…
u_n	0.31	0.45	NULL	…	0.66

Once the user profile is built, CF systems try to predict the utilities of items for a particular user based on the items previously rated by other users. Algorithms for collaborative recommendations can be grouped into two general classes: memory-based and model-based [9]. Memory-based algorithms [9-12] are heuristics that make rating predictions based on the entire collection of items previously rated by the users. That is, the value of the unknown rating $r(u,i)$ for user u and item i is computed as an aggregate of the ratings of some other users(neighbors) for the same item i [13], i.e.,

$$r(u,i) = \underset{u' \in \hat{U}}{aggr}\, r(u',i) \quad (1)$$

where aggr denotes the aggregation function, like weighted sum or adjusted weighted sum. $\hat{U}$ denotes the set of K neighbors that have rated item i and are similar to user u (K can range anywhere from 1 to the number of all users). Neighborhood of u is decided by the similarity measure between users u and u', $sim(u,u')$. Various approaches have been used to compute $sim(u,u')$. In most of these approaches, $sim(u,u')$ is based on the ratings of items that both users u and u'have rated. The two most popular approaches are the correlation-based approach [10-11] and the cosine-based approach [8][14]. In contrast to memory-based methods, model-based algorithms use the collection of ratings to learn a model, which is then used to make rating predictions [15-17].

CB methods [18-21] learn a profile of the user's interests based on the features presented in items that the user has rated. Schafer et al. [2] call this "item-to-item correlation". The type of user profile derived by a content-based recommender depends on the learning method employed. Decision trees, neural nets, and vector-based representations have all been used.

2. THREE-DIMENSIONAL RECOMMENDATION APPROACH FOR C2C E-COMMERCE CONTEXT
2.1PROBLEM DEFINITION
The recommendation problem inC2C context can be formally described as follows.

● Let $B=\{b_1, b_2, b_3,\ldots,b_n\}$ be the set of buyers, $S=\{s_1, s_2, s_3,\ldots,s_k\}$ be the set of sellers, and $P=\{p_1, p_2, p_3,\ldots,p_m\}$ be the set of products.

● Given
(1) the set of historical ratings $R = \{r(b,s,p)|b \in B, s \in S, p \in P\}$, where $r(b,s,p)$ indicates the rating that buyer b gives to product p and its seller s. R forms a three-dimensional recommendation space, where the three dimensions correspond to buyer, seller and product respectively, as illustrated in Figure 1.If buyer b has not given any rating to seller s and product p, then $r(b,s,p) = \text{NULL}$;

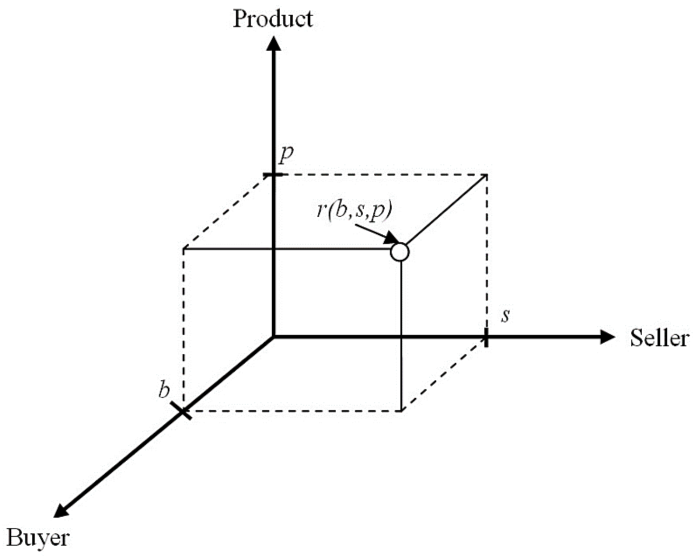

Fig 1:Three-dimensional recommendation space in C2C context

(2) the seller feature matrix SFM. SFM is built with the seller features provided by the reputation and trust system in theC2C shopping site. The vertical axis corresponds to sellers, the horizontal axis corresponds to features, and each cell $sfm(s,f)$ indicates the value of feature f belongs to seller s. Since the measurement scales vary greatly for different features, the data of features should be generalized and ordered in advance. Table 2 illustrates an example of SFM;

Table 2: Example of seller features matrix SFM

	f_1	f_2	f_3	…	f_m
s_1	5	1	2	…	2
s_2	4	4	7	…	6
…	…	…	…	…	…
s_n	2	3	4	…	9

(3) the sale relation matrix SRM. In SRM, the vertical axis corresponds to sellers, the horizontal axis corresponds to products, and each cell $srm(s,p)$ indicates whether seller s sells product p or not. ("1" means selling, "0" means not selling). Table 3 illustrates an example of SRM.

Table 3: Example of sale relation matrix SRM

	p_1	p_2	p_3	…	p_m
s_1	0	1	1	…	0
s_2	0	0	1	…	1
…	…	…	…	…	…
s_n	1	1	0	…	0

● The task of the recommendation is: according to R, SFM and SRM, predicting the unknown rating of $r(b,s,p)$ and recommending the top N "seller and product" combinations with the highest predicted ratings.

2.2 MODEL OF THE PROPOSED APPROACH

As mentioned above, we adopt the ideas of both CF and CB techniques, and compound them in our three-dimensional approach. Figure 2 shows the model of the approach.

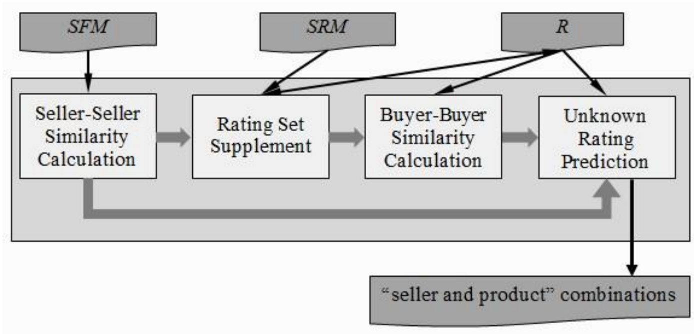

Fig 2:Model of three-dimensional recommendation approach

The recommendation process in the model comprises four steps:
(1)Calculate seller-seller similarities based on seller features;
(2)Employ rating inference mechanism based on seller-seller similarities and sale relations, and supplement R with the inferred ratings;
(3)Calculate buyer-buyer similarities based on the co-ratings of products;
(4)Predict the unknown ratings for the target user and the candidate "seller and product" combinations based on seller-seller similarities and buyer-buyer similarities.

2.2.1SELLER-SELLER SIMILARITY CALCULATION

To learn the buyer's preferences for sellers, we measure seller similarities using the CB method. Let $sim(s,s')$ be the similarity between seller *s* and *s'*, we calculate it using cosine measure based on sellers' features as follows:

$$sim(s,s') = \frac{\sum_{i=1}^{K} sfm(s,f_i)sfm(s',f_i)}{\sqrt{\sum_{i=1}^{K} sfm(s,f_i)^2 \sum_{k=1}^{n} f_k(s',f_i)^2}} \quad (2)$$

where $sfm(s,f_i)$ and $sfm(s',f_i)$ denote the values of feature f_i that seller s and s' have in SFM. K denotes the total number of the features in SFM.

2.2.2 RATING SET SUPPLEMENT

To solve the sparse data problem caused by more dimensions, we design a mechanism to supplement the historical ratings set R. The basic idea is that buyers may give similar ratings if the products are the same and the sellers are similar. Suppose buyer b has never rated product p and its seller s, but b has rated p and its other sellers, then the possible value of $r(b,s,p)$ can be inferred according to the existing ratings and seller similarities. The calculation of the inferred rating $r_{inference}(b,s,p)$ is as follows:

$$r_{inference}(b,s,p) = \frac{\sum_{s' \in S(b,p)} r(b,s',p) \times sim(s,s')}{M} \quad (3)$$

where $S(b,p) = \{s | srm(s,p) = 1, r(b,s,p) \neq NULL\}$ contains all sellers who sells p and buyer b has rated them and p. M denotes the number of sellers in $S(b,p)$. $sim(s,s')$ denotes the similarity between sellers s and s'. All inferred ratings will be filled into R to replace the corresponding NULLs.

2.2.3 BUYER-BUYER SIMILARITY CALCULATION
To learn the buyer's preferences for products, we measure buyer similarities using the CF method. That is, the similarity is computed based on the ratings of all products that the buyers have co-rated. However, the rating in R is a comprehensive evaluation given to both seller and product, so the independent rating about the product should be derived before similarity measure.

The rating $r(b,s,p)$ that buyer b gives to seller s and product p can be considered as the weighted sum of the two independent ratings that b gives to s and that b gives to p, i.e.,

$$r(b,s,p) = w_s r(b,s) + w_p r(b,p) \quad (4)$$

where w_s and w_p are weights indicates the extent of importance that the buyer attaches to the seller or the product. $r(b,s)$ can be computed as the average of all existing ratings involving b and s in R, i.e.,

$$r(b,s) = \frac{\sum_{p' \in P(b,s)} r(b,s,p')}{L_p} \quad (5)$$

where $P(b,s) = \{p | srm(s,p) = 1, r(b,s,p) \neq NULL\}$ contains all products sold by seller s and buyer b has rated s and them. L_p denotes the number of products in $P(b,s)$. Thus, if we suppose that buyers pay equal attention to sellers and products, namely $w_s = w_p = 0.5$, then according to Equation (4) and Equation (5), we can derive $r(b,p)$ as follows:

$$r(b,p) = 2r(b,s,p) - \frac{\sum_{p' \in P(b,s)} r(b,s,p')}{L_p} \quad (6)$$

Since it is possible that buyer b has bought and rated product p from several different sellers, the derived value of $r(b,p)$ may not be unique. In this situation, we take the average of all $r(b,p)$ s as the final rating.

Once the independent ratings that buyers give to products are derived, the buyer-buyer similarities can be measured as the traditional CF method does. Let $sim(b,b')$ be the similarity between buyer b and b', we calculate $sim(b,b')$ using Pearson correlation coefficient as follows:

$$sim(b,b') = \frac{n\sum_{p \in \bar{P}} r(b,p) r(b',p) - \sum_{p \in \bar{P}} r(b,p) \sum_{p \in \bar{P}} r(b',p)}{\sqrt{n\sum_{p \in \bar{P}} r(b,p)^2 - (\sum_{p \in \bar{P}} r(b,p))^2}\sqrt{n\sum_{p \in \bar{P}} r(b',p)^2 - (\sum_{p \in \bar{P}} r(b',p))^2}} \quad (7)$$

where $\bar{P}$ denotes the set of all products co-rated by both b and b'.

2.2.4UNKNOWN RATING PREDICTION

To predict the unknown rating that a buyer gives to a candidate "seller and product" combination, we extends the predict model of the traditional two-dimensional CF to involve three types of entities. Similar with Equation 1, the unknown rating $r(b,s,p)$ is computed as an aggregate of the ratings for the same product p given by K neighbors, i.e.,

$$r(b,s,p) = \underset{\substack{b' \in \hat{B} \\ s' \in S(p)}}{\text{aggr}}\, r(b',s',p) \tag{8}$$

where $\hat{B}$ denotes the set of neighbors (K buyers who have rated product p and are similar to buyer b). $S(p)$ denotes the set of sellers who sell product p.

We use the adjusted weighted sum as the aggregation function, then the three-dimensional prediction model of $r(b,s,p)$ is as follows:

$$r(b,s,p) = \frac{\sum_{\substack{b' \in \hat{B} \\ s' \in S(p)}} sim(b,b') \times sim(s,s') \times r(b',s',p)}{\sum_{\substack{b' \in \hat{B} \\ s' \in S(p)}} \left| sim(b,b') \times sim(s,s') \right|} \tag{9}$$

Buyer-buyer similarity $sim(b,b')$ and seller-seller similarity $sim(b,b')$ are used as weights for rating $r(b',s',p)$ in the model.That is, the more similar b with b', and s with s', the more weights $r(b',s',p)$ will carry in the prediction of $r(b,s,p)$.

3. EXPERIMENTAL EVALUATION

3.1 DATA SET

We use real-world sales data to examine the performance of the proposed approach. The data come from Taobao, the largest C2C website in China. We collect the sale records of Canon's digital cameras and accessories during a three months period from August to October2011, and get 18573 transactions, involving 10261 buyers,842 sellers and 139 products. To make the data suitable for evaluation, we filter the records by restricting the buyers to those who have participated at least two different transactions. Finally 5218 transactions, involving 1633 buyers, 327 sellers, 139 products, are preserved. Buyers in Taobao are required to give a rating after each valid transaction. All the transaction records and the ratings constitute the historical rating set R. The original ratings from Taobao have three possible grades: good, medium and bad. For our analysis, we replace the three grades with the scores of 1, 0.5 and 0 respectively.

We also collect sale relations and seller features in the same period of time. 1386 sale relations related to the sellers and products in R are obtained to form SRM. And 18 seller features represented by the reputation system in Taobao are used to form SFM, including sales volume, favorite popularity, browsed times, price, inventory, if describe the truth, if exchangeable within 7 days, if certified products, foul play, refund dispute, penalty, is mall, credits, good rating ratio, virtual commodity transactions, physical commodity transactions, service attitude, delivery speed.

3.2EXPERIMENT DESIGN

The experiment mainly involves four steps:

(1)We employ 5-fold cross-validation approach in the experiment. R is split using 20%-80% ratio and such split is done 5 times. For each of the 5 splits, we designate the 80% part as training dataset R_T, R_T contains 4174 ratings, used to predict unknown ratings and produce recommendation results. The remaining 20% part (contains1044 ratings) is designated as evaluation dataset R_E, used to evaluate the recommendation results. Based on the definition of $R_T$and R_E, we have that $R_E \cap R_T = \varphi$ and $R_E \cup R_T = R$;

(2)We implement the three-dimensional recommendation approach described in Section3.2 on R_T and test it on R_E. The approach produces top N tuples of <seller, product> as recommendation results. Since the size of the neighborhood can substantially affect the recommendation quality, we conduct a pilot test in advance and discover that 20 is the optimal value. To evaluate the sensitivity of varied recommendation number N, we perform the experiment with different N of 5, 10, 15, 20 and 25;

(3) To compare the proposed approach with the existing approaches, we also evaluate the two-dimensional CF approach using 5-fold cross-validation method on our data set. After the two-dimensional CF approach has produced the product recommendations, we use four common ways to select a seller for each recommended product: a) select the seller randomly; b) select the seller with the largest sales volume; c) select the seller with the lowest price; and d) select the seller with the highest credits. Thus all these four approaches also produce N tuples of <seller, product> as recommendation results;

(4) We use precision, recall and F1 metric [22], three widely-used measures to evaluate the predictive performance the recommendations. Precision is the fraction of recommended "seller and product" combinations that the buyer really selects. Recall is the fraction of buyer-really-selected "seller and product" combinations that are F1 metric combines recall and precision and gives equal weight to both measures:

$$\text{recall} = \frac{\text{correctly - recommende d-} < \text{seller, product} > \text{s}}{\text{total-} < \text{seller, product} > \text{s - selected - by - buyers}} \quad (10)$$

$$\text{precision} = \frac{\text{correctly - recommende d-} < \text{seller, product} > \text{s}}{\text{total - recommende d-} < \text{seller, product} > \text{s}} \quad (11)$$

$$\text{F1} = \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \quad (12)$$

3.3 RESULT AND ANALYSIS

The detailed results of recall, precision and F1 in the experiment are reported in Table 4, 5 and 6. The Test No. in tables denotes the number of different R_E/R_T split. The results show that the recall improves and the precision drops when the recommendation number increases.

Table 4: Recall metric for various approaches, recommendation numbers and test numbers

recall	**Recommendation number**	**3D Approach**	**2D Approach (random)**	**2D Approach (largest sales volume)**	**2D Approach (lowest price)**	**2D Approach (highest credits)**
Test No.1	**5**	0.112	0.057	0.035	0.028	0.037
	10	0.157	0.074	0.043	0.039	0.045
	15	0.236	0.113	0.072	0.067	0.080
	20	0.284	0.168	0.124	0.135	0.128
	25	0.322	0.204	0.158	0.163	0.152
Test No.2	**5**	0.109	0.051	0.023	0.021	0.027
	10	0.154	0.075	0.040	0.053	0.047
	15	0.238	0.103	0.078	0.067	0.095
	20	0.309	0.180	0.114	0.135	0.132
	25	0.301	0.208	0.137	0.165	0.143
Test No.3	**5**	0.104	0.040	0.041	0.040	0.039
	10	0.140	0.077	0.049	0.027	0.042
	15	0.233	0.113	0.086	0.067	0.083
	20	0.306	0.179	0.133	0.136	0.142
	25	0.318	0.204	0.166	0.153	0.151
Test No.4	**5**	0.112	0.042	0.017	0.028	0.029
	10	0.127	0.081	0.049	0.039	0.045
	15	0.251	0.110	0.071	0.069	0.059
	20	0.287	0.173	0.127	0.142	0.128
	25	0.297	0.191	0.139	0.172	0.150
Test No.5	**5**	0.122	0.070	0.037	0.036	0.049
	10	0.144	0.074	0.043	0.031	0.040
	15	0.247	0.106	0.067	0.073	0.080
	20	0.296	0.143	0.132	0.135	0.116
	25	0.305	0.188	0.166	0.154	0.149

Figure 3 to Figure 7 show the comparisons of F1 metric among the three-dimensional (3D) approach and the four two-dimensional (2D) approaches. In all five tests, the F1 metric reaches an optimal value when the recommendation number increases to 20.

Table 5: Precision metric for various approaches, recommendation numbers and test numbers

precision	Recommendation number	3D Approach	2D Approach (random)	2D Approach (largest sales volume)	2D Approach (lowest price)	2D Approach (highest credits)
Test No.1	5	0.464	0.362	0.301	0.274	0.286
	10	0.391	0.328	0.259	0.261	0.270
	15	0.308	0.270	0.238	0.250	0.247
	20	0.269	0.235	0.178	0.153	0.164
	25	0.228	0.169	0.115	0.104	0.128
Test No.2	5	0.470	0.362	0.280	0.283	0.266
	10	0.372	0.359	0.266	0.261	0.292
	15	0.314	0.261	0.242	0.247	0.240
	20	0.268	0.224	0.180	0.147	0.172
	25	0.228	0.169	0.120	0.11	0.143
Test No.3	5	0.472	0.374	0.289	0.279	0.276
	10	0.400	0.339	0.261	0.262	0.259
	15	0.284	0.257	0.235	0.238	0.240
	20	0.256	0.225	0.180	0.163	0.164
	25	0.243	0.164	0.113	0.117	0.129
Test No.4	5	0.459	0.371	0.296	0.291	0.282
	10	0.400	0.340	0.261	0.268	0.263
	15	0.314	0.262	0.238	0.244	0.240
	20	0.283	0.241	0.178	0.157	0.158
	25	0.241	0.156	0.123	0.109	0.128
Test No.5	5	0.469	0.352	0.305	0.280	0.294
	10	0.377	0.344	0.244	0.255	0.277
	15	0.282	0.281	0.238	0.265	0.254
	20	0.275	0.242	0.167	0.153	0.161
	25	0.233	0.163	0.115	0.106	0.126

Table 6: F1 metric for various approaches, recommendation numbers and test numbers

F1	Recommendation number	3D Approach	2D Approach (random)	2D Approach (largest sales volume)	2D Approach (lowest price)	2D Approach (highest credits)
Test No.1	5	0.180	0.098	0.063	0.051	0.066
	10	0.224	0.121	0.074	0.068	0.077
	15	0.267	0.159	0.111	0.106	0.121
	20	0.276	0.196	0.146	0.143	0.144
	25	0.267	0.185	0.133	0.127	0.139
Test No.2	5	0.177	0.089	0.043	0.040	0.049
	10	0.218	0.124	0.070	0.088	0.081
	15	0.271	0.148	0.118	0.105	0.136
	20	0.287	0.199	0.140	0.141	0.149
	25	0.259	0.186	0.128	0.132	0.143
Test No.3	5	0.171	0.082	0.070	0.054	0.060
	10	0.208	0.125	0.082	0.048	0.072
	15	0.256	0.157	0.126	0.105	0.123
	20	0.279	0.199	0.153	0.148	0.152
	25	0.276	0.182	0.134	0.132	0.139
Test No.4	5	0.180	0.076	0.032	0.051	0.052
	10	0.193	0.131	0.082	0.068	0.076
	15	0.279	0.155	0.109	0.107	0.095
	20	0.285	0.201	0.148	0.149	0.141
	25	0.266	0.171	0.131	0.134	0.138
Test No.5	5	0.193	0.117	0.066	0.064	0.083
	10	0.208	0.122	0.073	0.055	0.070
	15	0.263	0.154	0.104	0.115	0.122
	20	0.285	0.180	0.147	0.143	0.135
	25	0.264	0.175	0.136	0.126	0.137

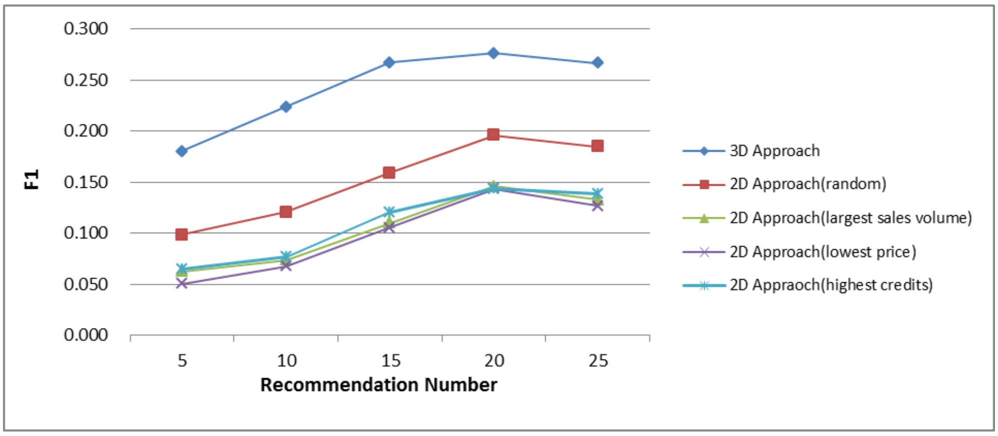

Fig 3: F1 metric for various approaches vs. recommendation number (Test No:1)

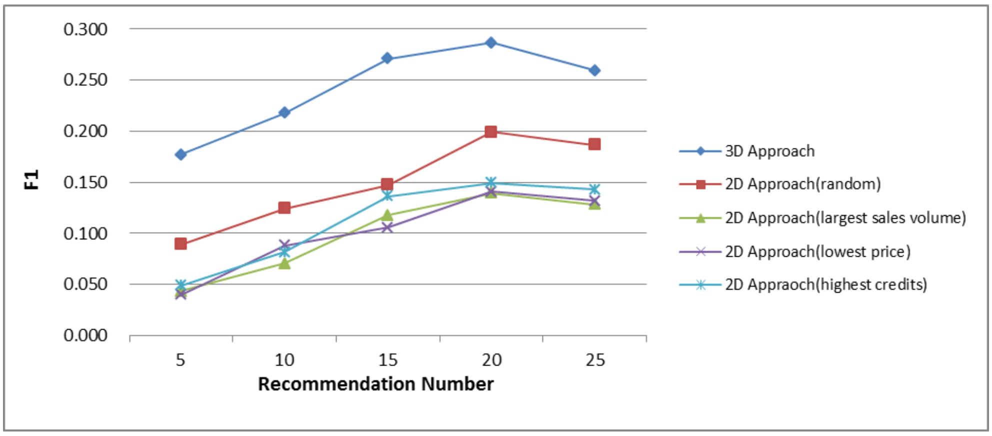

Fig 4: F1 metric for various approaches vs. recommendation number (Test No:2)

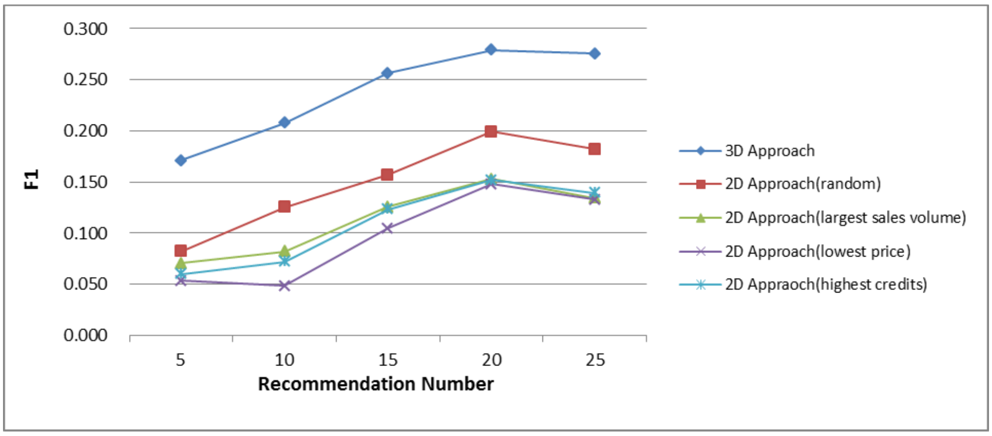

Fig 5: F1 metric for various approaches vs. recommendation number (Test No:3)

All the results show that the three-dimensional approach performs better than the other four two-dimensional approaches at all three metrics. In addition, among the four two-dimensional approaches, selecting seller randomly performs relatively better, and the other three perform with little difference. This can be explained that buyers seldom select sellers by particular sole criterion, but consider various seller features synthetically. When the seller is selected randomly, all features are possible to affect the result, so it outperforms the approaches considering only one feature. This confirms our former viewpoint that the buyer's selection of sellers in C2C context is complicated and personalized behavior, and need to be modeled by more complex means. The experiment results prove that the three-dimensional predictive model is relatively good at reflecting buyer's preferences for both seller and product.

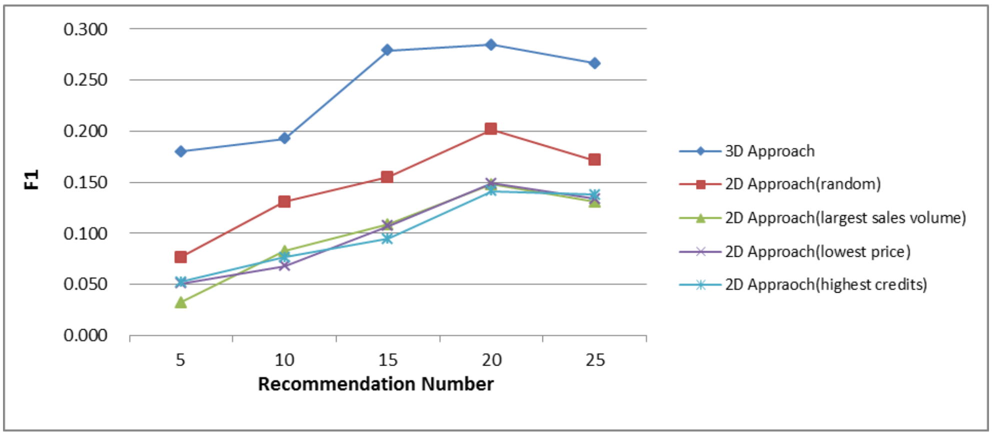

Fig 6: F1 metric for various approaches vs. recommendation number (Test No:4)

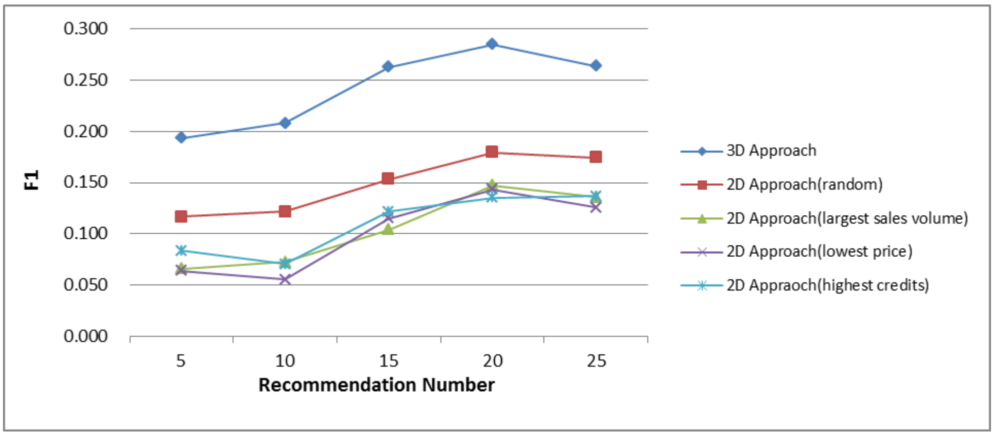

Fig 7: F1 metric for various approaches vs. recommendation number (Test No:5)

CONCLUSION

In this paper, we have proposed a new three-dimensional recommendation approach suitable for C2C e-commerce context, which explores the relevance among buyers, sellers and products, and recommends personalized "seller and product" combinations for buyers. Based on historical ratings, seller features and sale relations, the proposed approach combines the ideas of both the content-based filtering and the collaborative filtering to calculate seller-seller similarities and buyer-buyer similarities, and predict unknown ratings using a three-dimensional prediction model. We have also implemented a rating inference mechanism in our approach to supplement the spare data and improve the effectiveness of the calculations. To evaluate the recommendation quality, we conduct a true data experiment. The results have shown that our approach outperforms the existing approaches in the real C2C e-commerce context.

Acknowledgments
This research was financially supported by the National Social Science Foundation of China (11CTQ020), the National Natural Science Foundation of China (71203037) and the Natural Science Foundation of Guangdong Province(S2012040007883).

REFERENCES

[1] JA Konstan. *ACM Transactions on Information Systems (TOIS)*, **2004**, 22(1),1-4.
[2] JB Schafer; JA Konstan; J Riedl. *Applications of Data Mining to Electronic Commerce*. **2001**, 5-10.
[3] K Wei; J Huang; S Fu. *Proc. of International Conference on Services Systems and Services Management, IEEE*, **2007**, 1-5.
[4] G Adomavicius; R Sankaranarayanan; S Sen, A Tuzhilin.*ACM Transactions on Information Systems (TOIS)*, **2005**, 23(1), 103-145.
[5] X Su;TM Khoshgoftaar.*Advances in Artificial Intelligence*, **2009**(4),1-20.
[6] C Basu;H Hirsh;W Cohen. *Proc.of the 15th National Conference on Artificial Intelligence*, AAAI, **1998**, 714-720.

[7] DH Park; HK Kim; IY Choi;JK Kim. *Expert Systems with Applications*, **2012**, 39(11): 10059-10072.
[8] JA Konstan; BN Miller; D Maltz; JL Herlocker, LRGordon. *Communications of the ACM*, **1997**, 40(3): 77-87.
[9] JS Breese;D Heckerman;C Kadie. *Proc. of the 14th Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann, **1998**, 43-52.
[10] P Resnick;N Iacovou;M Suchak;P Bergstrom;J Riedl. *Proc.of the 1994 ACM Conference on Computer Supported Cooperative Work*, ACM, **1994**, 175-186.
[11] U Shardanand;P Maes. *Proc.of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, **1995**, 210-217.
[12] J Delgado;N Ishii. *Proc. of the ACM SIGIR'99 Workshop Recommender Systems: Algorithms and Evaluation*, ACM, **1999**.
[13] G Adomavicius;A Tuzhilin. *IEEE Transactions on Knowledge and Data Engineering*, **2005**, 17(6), 734-749.
[14] B Sarwar;G Karypis;J Konstan;J Riedl. *Proc of the 10th international conference on World Wide Web*,Elsevier, **2001**, 285-295.
[15] D Billsus; M Pazzani. *User-Modeling and User-Adapted Interaction*, **2001**, 10(2-3), 147-180.
[16] LH Ungar; DP Foster. *Proc.of the Workshop on Recommendation Systems*, AAAI, **1998**, 114-129.
[17] K Goldberg; T Roeder;D Gupta;C Perkins. *Information Retrieval*,**2001**, 4(2), 133-151.
[18] M Balabanovic;Y Shoham. *Communications of the ACM*, **1997**, 40(3),66-72.
[19] K Lang.*Proc.of the 12th International Conference on Machine Learning*, Morgan Kaufmann, Springer, **1995**, 331-339.
[20] J Li;ORZaïane. *Proc.of the 5th International Conference on Electronic Commerce and Web Technologies*, **2004**, 305-315.
[21] M Pazzani;D Billsus. *Machine learning*, **1997**, 27(3), 313-331.
[22] CJ Van Rijsbergen. *Information Retrieval*, second edition, Butterworths, London,**1979**.