Information Sciences 180 (2010) 2142–2155

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Collaborative filtering with ordinal scale-based implicit ratings for mobile music recommendations

Seok Kee Lee [a,1], Yoon Ho Cho [b,*], Soung Hie Kim [a,2]

[a] Graduate School of Management, Korea Advanced Institute of Science and Technology, 207-43 Cheongryangri-Dong, Dongdaemun, Seoul 130-012, Republic of Korea

[b] School of Business Adminstration, Kookmin University, 861-1 Jungnung-dong, Sungbuk-gu, Seoul 136-702, Republic of Korea

ARTICLE INFO

Article history:
Received 17 February 2009
Received in revised form 8 January 2010
Accepted 1 February 2010

Keywords:
Collaborative filtering
Recommender system
Mobile web usage mining
Consensus model

ABSTRACT

Collaborative filtering (CF)-based recommender systems represent a promising solution for the rapidly growing mobile music market. However, in the mobile Web environment, a traditional CF system that uses explicit ratings to collect user preferences has a limitation: mobile customers find it difficult to rate their tastes directly because of poor interfaces and high telecommunication costs. Implicit ratings are more desirable for the mobile Web, but commonly used cardinal (interval, ratio) scales for representing preferences are also unsatisfactory because they may increase estimation errors. In this paper, we propose a CF-based recommendation methodology based on both implicit ratings and less ambitious ordinal scales. A mobile Web usage mining (mWUM) technique is suggested as an implicit rating approach, and a specific consensus model typically used in multi-criteria decision-making (MCDM) is employed to generate an ordinal scale-based customer profile. An experiment with the participation of real mobile Web customers shows that the proposed methodology provides better performance than existing CF algorithms in the mobile Web environment.

1. Introduction

In recent years, an increasing variety of content has been made available in the mobile Web environment, with the mobile music market witnessing a particularly rapid growth. However, customers still experience a great deal of frustration when searching for the music they want on mobile Web devices, owing to the inefficiencies of searching sequentially. When a customer uses a cell phone to log on to a site to download music, the site presents a list of the available best-selling or newest music. The customer pages through the list and selects items to inspect. If the customer likes the item, he/she may buy it. Otherwise, the customer repeats these steps until he/she stumbles over the right item or gives up. With this method, the number of items the customer would be expected to encounter before finding the desired item could far exceed the acceptable level.

These difficulties are partly attributable to typical cell phone features. Compared to PCs, cell phones have smaller screens, fewer input keys, and less sophisticated browsers. Thus, the user interface of the mobile Web application is not as friendly as that of typical Web applications. To make searching more acceptable, a more efficient search aid suggesting only the items meeting the customer's preference is therefore necessary.

* Corresponding author. Tel.: +82 2 910 4950; fax: +82 2 910 5209.
E-mail addresses: seelee@paran.com, pusanlsk@business.kaist.edu (S.K. Lee), www4u@kookmin.ac.kr (Y.H. Cho), seekim@business.kaist.ac.kr (S.H. Kim).
[1] Tel.: +82 2 958 3684; fax: +82 2 958 3604.
[2] Tel.: +82 2 958 3611; fax: +82 2 958 3604.

0020-0255/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2010.02.004

Toward this goal, recommender systems are intended to help customers find music they wish to purchase. To date, a variety of recommendation techniques have been developed. Of these, collaborative filtering (CF) has been the most successful [2,19,21,29,31,57,58,60]. It identifies customers whose tastes are similar to those of the target customer and recommends items that those customers have liked.

CF-based recommender systems require a customer profile to identify preferences and make recommendations. To create a customer profile, two profiling techniques are commonly used: explicit ratings and implicit ratings. Explicit ratings are techniques in which customers examine each item and assign it a value on an agreed rating scale. Implicit ratings are techniques that gather information about a customer's shopping behaviors and represent preferences as ratings without the customer's intervention.

Although explicit ratings are well-understood and fairly precise, some problems still arise when applying them in the mobile domain. As previously mentioned, the user interface of mobile devices is typically poor, and the cost of using the mobile Web through these devices is high. Therefore, explicit ratings may present a burden to customers. Enforcing a rating that may alter a customer's normal pattern of searching [5,44] is another problem of explicit ratings. In this regard, implicit ratings may be considered as better techniques.

As alternatives to rating scales in implicit ratings, there are different scale types such as ordinal and cardinal (interval and ratio) scales. It is generally known that cardinal scales are more accurate because of their high sensitivity. The more values for the parameters "like" or "dislike" a scale has, the more sensitive it is and the more refined information it is able to provide [17]. In preference representation, higher sensitivity enables any system to imply both the order and degree of preference, resulting in higher quality of recommendations. Cardinal scales are sensitive scale types in which, compared to the ordinal scale, it is easy to increase the number of possible values that can be reliably used. For this reason, cardinal scales have been widely used in explicit ratings, making it possible for customers to represent their preferences quite accurately.

Implicit ratings have also used cardinal scales to increase the accuracy of estimation. However, it is uncertain whether cardinal scales are also better in implicit ratings. Previous research has reported that a less accurate scale may be better in some situations wherein sufficient information is not available for estimation [17,59]. A highly sensitive scale in such situations might unfortunately increase the variance of estimated values, thereby increasing the *estimation error* (i.e., the gap between the actual preference and those predicted by estimation). A customer's explicit preference is a comprehensive mixture of several attributes, including demographic characteristics, shopping behaviors, and psychological characteristics, among others. For exact estimation, implicit ratings should use as many of these attributes as possible. However, often in implicit ratings, information about these attributes is not available because of difficulties in its acquisition. For example, it is extremely difficult to obtain psychological information because it is not usually exposed. Therefore, the implicit rating approach has no choice but to use insufficient information. Hence, a less accurate ordinal scale might be better than a cardinal scale in this situation.

In sum, a mechanism which can implicitly analyze a customer's preference and represent it using an ordinal scale would be necessary for an effective CF-based recommendation system in the mobile Web environment. However, this problem has not received much attention from CF-related research communities. In this paper, we propose a new CF recommendation methodology for the mobile music, called collaborative filtering with ordinal scale-based implicit ratings (CoFoSIM).

In order to capture implicit preference information, we develop a mobile Web usage mining technique (mWUM) that applies data mining techniques to discover customer behavior patterns by using mobile Web log data. Through mWUM, all recorded transactions in mobile Web logs are individually analyzed, and a large amount of corresponding ordinal preference information is collected. This collected information is insufficient because it is partial and somewhat conflicting. Specifically, individual preference information from a single transaction cannot provide complete preference information, but it is restricted to some parts of the entire spectrum of choices. Moreover, preferences regarding a specific type of music experienced more than once in several transactions may conflict because a customer's preference is changeable over time. Therefore, a mechanism that makes possible a compromised preference by aggregating this partial and conflicting preference information is required as a supplement to mWUM.

A possible solution is the consensus model, which is one of many significant research topics in the area of multi-criteria decision-making (MCDM). This model addresses the problem of aggregating preferences of a set of individuals into a compromised preference. To date, various consensus models have been investigated [7,9,15,16,34]. Among them, a consensus model suggested by Cook and Kress (called the CK method) is well-known and has been applied in various settings. This model generates a compromised preference based on the view of distance, which minimizes differences (or conflicts) among a set of ordinal preferences. Our proposed CoFoSIM applies this CK method to generate a compromised preference from various pieces of ordinal preference information that are both partial and conflicting.

To evaluate the performance of CoFoSIM, we performed an experiment in the real mobile Web environment. Hundreds of mobile Web customers were involved, and their behavior data were collected and analyzed real-time. Two additional CF-based recommender systems were developed as benchmarks. The first benchmark system adopts mWUM and represents preferences as cardinal scale-based. The second benchmark system adopts mWUM and represents compromised preferences as ordinal scale-based, but it does not apply the CK method. Our experimental results show that the performance of CoFoSIM is superior to the benchmarking systems.

The rest of this paper is organized as follows: We begin by reviewing the related research literature in Section 2. Section 3 provides an explanation of our methodology through a detailed breakdown of the phases involved. Section 4 presents and discusses the experimental results. Section 5 concludes with suggestions for future research.

2. Related works

2.1. Collaborative filtering (CF)

Collaborative filtering (CF) is one of the most successful recommendation techniques, which has been proven by abundant research and actual implementations [10,33,54,58,60–62]. The basic idea is to recommend items to a target user by predicting their utility for that user through previous ratings by other users. For several decades, many CF-based recommender systems have been developed by the academia and the industry. The earliest implementation of CF, the Grundy system [49], involves the use of *stereotypes* as a mechanism for building models of users through a limited amount of information on each individual user. Using stereotypes, the Grundy system builds models of individual users and uses these to book recommendations. The Tapestry system [20] relies on each user to manually identify like-minded users. These earliest CF systems use opinions of people from a close-knit community such as an office workgroup.

However, CF for large communities cannot depend on people who know one another. Several systems, collectively called *automated collaborative filtering*, have emerged to provide personalized recommendations of items by finding a group of other users known as *neighbors*. These statistical approaches typically rely on "ratings" as numerical expressions of user preference. Subsequently, several systems based on automated CF have been developed, including GroupLens [37], Video Recommender [25], Ringo [53], and PHOAKS [56]. GroupLens [37] provides automated neighborhoods for recommendations in Usenet news. Based on a user's ratings on articles, GroupLens automatically recommends other articles to the user. Video Recommender [25] makes recommendations on movies, whereas Ringo [53] uses CF to provide users with recommendations on CDs. The PHOAKS system [56] helps people in finding relevant information on the Web. It differs from the others previously mentioned because it uses implicit ratings to make recommendations.

Because CF-based recommender systems use other users' recommendations (ratings), their advantage lies in the fact that they can deal with any kind of content and recommend any items, even ones that are dissimilar to those already seen in the past. However, the increased use of CF has exposed some limitations such as sparsity, new user problems, and new item problems [4,38]. Fortunately, many approaches have been developed to overcome these limitations. With regard to sparsity, several techniques have been proposed, which are classified into the following categories: the *implicit rating*, *dimensionality reduction*, *item-to-item correlation*, and *hybrid* approaches. The implicit rating approach attempts to increase the number of ratings entered by observing user behavior [1,11,35]. For example, Aggarwal and Yu [1] described the tracking of user behavior on Web sites to implicitly identify user preferences. A formal scheme of WUM for implicitly capturing user behavior has been suggested [11,35]. Singular value decomposition (SVD) and clustering have been used to reduce the dimensionality of sparse rating matrices [6,52]. For example, Sarwar et al. [52] suggested a method of using SVD for matrix factorization, which provides the best lower rank approximations of the original matrix. Instead of identifying a neighborhood of similar users, the item-to-item correlation approach analyzes the user-item matrix to identify relationships between different items and uses these relations to compute the prediction score for a given user-item pair. The hybrid approach uses a combination of content-based filtering and collaborative filtering, which enhances the respective advantages of both techniques [4,13,47]. For example, Claypool et al. [13] and Pazzani [47] implemented separate collaborative and content-based systems and combined the outputs (ratings) obtained from the individual systems into one final recommendation by using either a linear combination of ratings [13] or a voting scheme [47]. Fab [4] is based on traditional CF, but it also maintains content-based profiles for each user. These hybrid recommendation approaches have also been used to effectively overcome both new user and new item problems.

Compared with the large volume of research addressing the above-mentioned common problems of CF, little attention has been paid to the issue of using an ordinal scale in a situation wherein information about a user's preference must be implicitly collected. The use of an ordinal scale or implicitly collecting user preferences has been individually examined. First, a few studies have recently examined the issue of using an ordinal scale for representing user preferences [14,28,30,32]. For example, Kamishima [30] suggested a recommender system called "Nantonac CF," where preferences are represented by an ordinal scale. This study comments on the inappropriateness of cardinal scales for representing an individual user's preference. Cohen et al. [14] and Joachims [28] proposed a method to sort attributed items associated with pairwise compared information. Kazawa et al. [32] studied the learning problem from the perspective of ordered item sets. All of these studies showed that using an ordinal scale is efficient but that they commonly used explicit ratings as a profiling technique. The use of an ordinal scale in implicit ratings-based CF has not been investigated.

Cho and Kim [11] suggested a formal way of using a WUM technique to ascertain implicit ratings in their research on a shopping-mall's product recommendations. Kim et al. [35] suggested a basic method of mWUM in their research on mobile image recommendations. Their study proved that WUM in the mobile environment is an effective technique in alleviating the sparsity problem and in enhancing the quality of the recommendation. However, both of these studies used cardinal scales for ratings and did not investigate the possibility that an ordinal scale might overcome some of the limitations of cardinal scales when used with the implicit rating approach. Our CF-based recommender system suggests a combination of mobile Web usage mining (mWUM) for implicit ratings and an ordinal scale for representing preferences.

2.2. Consensus model

Studies on how to treat the problem of aggregating ordinal preferences into a consensus have been widely conducted in the field of MCDM research. Numerous procedures have been proposed to create a compromise or consensus. These are broadly classified into two categories: *ad-hoc methods* and *distance-based consensus.* The ad-hoc method is a traditional method originating from the parliamentary and committee settings for preference voting needs of the 18th century and has evolved into the *social choice theory* of today [9,15]. In one of the pioneering studies, Condorcet [15] proposed what is now commonly known as the *simple majority rule* method, whereby the alternative x should be declared the winner if for all $y \neq x, x$ is preferred to y by more voters than the number that prefers y to x. Borda [9] suggested the aggregation rule (called Borda's method), which generates consensus by deriving the total of voter-assigned ranks for each alternative. In recommender systems, there are at minimum hundreds of items to be compared, and therefore, using these ad-hoc methods may cause the cyclical ranking problem mentioned in Condorcet [15].

Unlike the ad-hoc methods that commonly use a simple additive rule, distance-based consensus methods examine an aggregation or consensus from the perspective of a distance function [3,7,16,34]. This concept has an intuitive appeal in that a consensus is defined as the set of preferences that are the closest, in a minimum distance sense, to voters' responses [16]. This idea was first advanced by Kemeny and Snell [34] and later adopted by Blin [7], Armstrong et al. [3], and Cook [16]. Kemeny and Snell [34] presented a set of natural axioms that such a measure should satisfy and proved the existence and uniqueness of the measure. Preferences concerning n alternatives are represented in a $n \times n$ pairwise comparison matrix. Blin [7] suggested an alternative to the work of [34] for complete orderings, and Armstrong et al. [3] extended this to include ties (weak orderings). Adding to these distance-based consensus methods, Cook and Kress [16,18] suggested a consensus method to represent the intensity of preference (called the CK method). Unlike previous distance-based methods, the CK method attempts to express the strength of preference by diversifying the preference specifications of Kemeny and Snell's method. Specifically, voters are permitted to express the intensity of their preferences, whereas this is not possible in Kemeny and Snell's method. Incorporating both the preference and the intensity makes the preference information more accurate. For this reason, we adopt the CK method as a consensus method for our CoFoSIM.

3. Methodology

3.1. General behavior patterns in the mobile Web

Before describing the specific phases of CoFoSIM, we first define the general patterns that a customer demonstrates when accessing the mobile Web. Fig. 1 illustrates a common scenario of searching for music on the mobile Web. A customer is first presented with an initial recommendation list from the content provider. The list is commonly generated from best-selling or newest items. The customer browses through the list until he/she finds the preferred item. If the customer finds an interesting title on the list, he/she may then click on that title and access a new page from which new information can be viewed (e.g., title, singer, genre, price, etc.). If the customer does not like the first choice, he/she may go back to the list. Otherwise, the customer may proceed to purchase or listen to the sample to experience it in advance. Subsequently, the customer may decide whether to purchase that content. Finally, the customer may quit the process or go back to the initial list to continue searching.

In comparison, CoFoSIM identifies the customer's preference from information on his/her previous behaviors, and then it creates the customer profile based on these preferences. For the purposes of our research, the following general behavior patterns in the mobile Web environment are defined (based on the works of [11,35]):

Ignore: Not clicking on the title of music on the list page and moving on to the next music title or pages.
Click-through: Clicking on a certain title on the list page and viewing the detailed information.
Pre-listen: Pre-listening a sample of the music that was selected.
Purchase: Buying the music that has been clicked-through or pre-listened.

Based on these patterns, we classify all music items into four groups as purchased music, pre-listened music, clicked-through music, and other music items (ignored). This classification provides an *is-a* relation between the different groups such that purchased music *is-a* music pre-listened and music pre-listened *is-a* music clicked-through [11,35]. From this relation, it is reasonable to obtain a preference order between music items such that {music ignored (never clicked)} $\prec$ {music clicked-through} $\prec$ {music pre-listened} $\prec$ {music purchased}. It also makes sense to assign greater weight to purchased music than to items that were only pre-listened to. Similarly, greater weight is given to pre-listened music than to music that was only clicked-through, and so on.

3.2. Phase 1: mobile Web usage mining (mWUM)

The first phase of our CoFoSIM is mWUM. Specifically, the customer action set is created in this phase. As the customer action set represents the target customer's accumulated shopping behaviors, a customer's transactions remaining in raw mobile logs should be collected first. This data preprocessing step includes the dependent and sequential tasks of data cleaning, user identification, and session identification. The detailed method is described in Step 1. After that, the target customer's transaction file is converted into a specific matrix that represents the customer action set. The detailed method is described in Step 2.

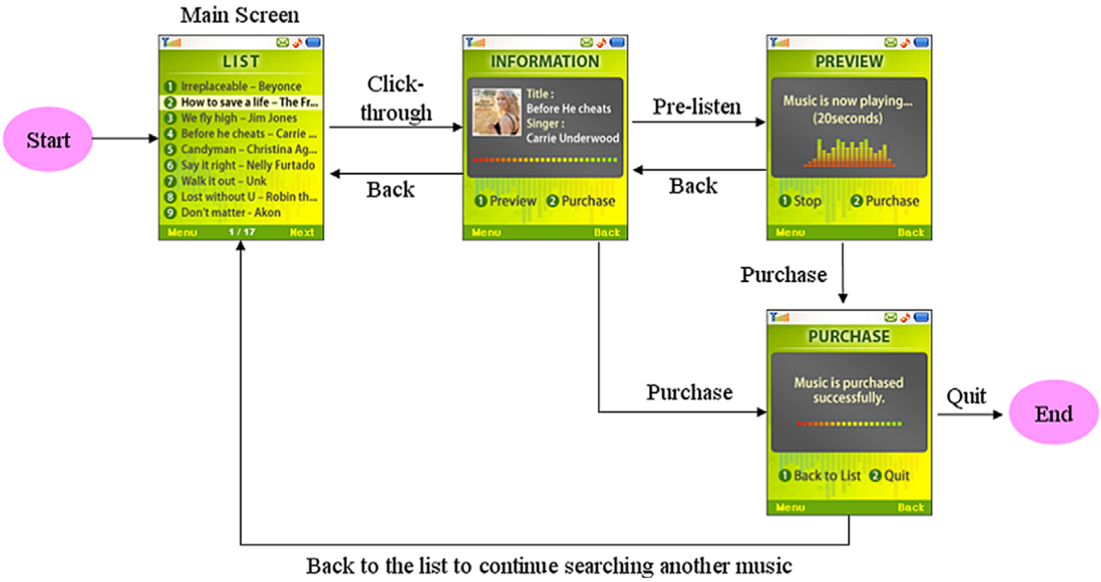

Fig. 1. Music search scenario in the mobile Web (each screenshot is captured from CoFoSIM).

3.2.1. Step 1-1. data preprocessing

Data preprocessing in mWUM consists of a series of dependent tasks including data cleaning, user identification, and session identification. It is similar to WUM in the typical Web environment, but it is simpler and more advantageous because of the unique characteristics of the mobile Web. First, in the typical Web environment, most Web pages contain numerous irrelevant items such as gif, jpg, and swf, and removing them is a time-consuming job. In the mobile Web, however, most pages consist mainly of text, and pages are small with limited components. For example, for mobile Internet protocols such as ME (mobile explorer), the maximum number of graphics per page is restricted to five or less, with each page limited to 23Kbytes maximum. This makes the task of removing irrelevant items a trivial job. Second, in the mobile Web environment, all customers must subscribe to network carriers and receive a unique cell-phone or equipment number (EN). This prerequisite user registration prevents confusion in the user identification task. Furthermore, it provides the opportunity to utilize a customer's demographic and social information for better recommendations. Third, in the mobile Web environment, all pages accessed by a customer at a certain time are recorded with identical values in the specific field (e.g., SESSION_ID) of the log, and this SESSION_ID field can be successfully used to identify each individual session. Using these advantages, CoFoSIM creates session files by sorting all log entries according to the SESSION_ID's values and grouping them together for each customer.

3.2.2. Step 1-2. customer behavior mining

Once the data preprocessing tasks have been completed in Step 1, a customer's session file is created. In this step, a specific matrix called the *customer action set* is created by using the session file. The customer action set $C = (c_{sj})$ is a matrix containing the numerical weights of the target customer's shopping behaviors for music items. It is defined as follows:

Definition 1. Customer Action Set$C$

$$C = (c_{sj}), \text{for which } c_{ij} = \begin{cases} 3, & \text{if } m_j \text{ is } \textit{purchased} \text{ in the } s\text{th session} \\ 2, & \text{if } m_j \text{ is } \textit{pre-listened} \text{ in the } s\text{th session} \\ 1, & \text{if } m_j \text{ is } \textit{clicked-through} \text{ in the } s\text{th session} \\ 0, & \text{if } m_j \text{ is } \textit{ignored} \text{ in the } s\text{th session} \end{cases} \tag{1}$$

where $m_j \in M; s = 1, \ldots, S$ and $j = 1, \ldots, T$

$^*M = \{m_1, \ldots, m_{|M|}\}$: a set of music items.

$^*S$: total number of sessions, $^*T$: total number of music items.

As shown in Definition 1, the cells of a customer action set have four possible degrees of values. Those values from 0 to 3 represent the weight or degree of preference. We place the greatest weight on purchased music items because they reflect the customer's strongest preference. The lowest weight, 0, is assigned to music items that are ignored because the customer passes them by, and thus, they are considered to be least preferred. The intermediate weights, 2 and 1, are assigned to items that are pre-listened and clicked-through, respectively. The rows in the customer action set increase according to the frequency of the customer's visits to the site. We assume that there is a target customer (Customer A) and that a customer action set is created for the customer as follows in Table 1:

Example 1. See Table 1.

3.3. Phase 2: ordinal scale-based customer profile creation

The customer profile describes a customer's product interests or preferences. In this phase, the customer profile, which represents its components on an ordinal scale, is created. As described below, this process requires three sequential steps.

3.3.1. Step 2-1. preference intensity matrix creation

As shown by Example 1, a customer action set for each customer has L rows corresponding to the number of rows in the session file. Usually, customers have limits on the number of music items they are capable of browsing through, owing to limited resources (e.g., time and money) and the inconvenience of visiting mobile Web sites. Therefore, individual rows of customer action sets contain a part of the preference information, and all must be aggregated to make a unified customer profile over all music items. To incorporate preferences, including the degrees, we use the term *preference intensity matrix* in [16,18] and modify it for our research purposes.

Definition 2. The preference intensity matrix $P^l = (p^l_{ij})$ is a matrix for which

$$(p^l_{ij}) = c_{li} - c_{lj}, \tag{2}$$

where $l = 1$ to $L, i = 1$ to T, and $j = 1$ to T. L is the total number of rows in the customer action set, and T is the total number of music items. Here, p^l_{ij} represents the relative strength/weakness of preference that a certain music item m_i has in comparison with other music items m_j in the lth session. The value ranges from -3 to 3, where a more preferred item has a greater positive value as it's p_{ij}. Suppose that a customer purchased music m_1 and pre-listened to music m_2 in the first session. According to Definition 2, the weight of $m_1(c_{11})$ is 3 and $m_2(c_{12})$ is 2. Thus, p^1_{12} is 1 ($3 - 2 = 1$), and p^1_{21} is -1 ($2 - 3 = -1$). It is interpreted that m_1 is preferred to m_2 and that the degree of the preference gap is 1. Through this calculation, each music item emerging from each session results in a p_{ij} score. As an example, the preference intensity matrix P^1 generated from the first row of the customer action set described in Example 1 is specified as follows in Table 2:

Example 2. See Table 2.

As there are L rows in the customer action set, L preference intensity matrices are generated for each customer. These L matrices are compromised to an optimal one in Step 2 and end up as a component of the customer profile described in Step 3.

3.3.2. Step 2-2. optimal preference intensity matrix creation

After Step 1, a certain customer's L individual preference information about music is identified with degrees by L corresponding preference intensity matrices $\{P^l\}^L_{l=1}$. In this step, preference intensity matrices $\{P^l\}^L_{l=1}$ for a specific customer are put together. And then, an *optimal preference intensity matrix* $\hat{X}$, which implies comprehensive preference information on that customer's music preferences, is generated. Then, $\hat{X}$ can be defined as follows:

Definition 3. The optimal preference intensity matrix $\hat{X}$ is the preference intensity matrix for which

$$F(\hat{X}) = \sum_{l=1}^{L} d(P^l, \hat{X}) = \min_{Q \in \beta} \sum_{l=1}^{L} d(P^l, Q), \tag{3}$$

where the distance function $d(P^l, Q) = \frac{1}{2}\sum_{i,j}\left|p^l_{ij} - p^Q_{ij}\right|$ and β is a set of all $n \times n$ preference intensity matrices. Thus, the optimal preference intensity matrix $\hat{X} = (\hat{x}_{ij})$ is a matrix that is the closest, in the l_1 norm sense, to the set of matrices $\{P^l\}^L_{l=1}$. Hence, (3) becomes

$$\min \sum_{i<j} \sum_{l=1}^{L} \left|p^l_{ij} - x_{ij}\right| \; subject\ to: \tag{4}$$

Table 1
Customer action set for Customer A.

Session	Music					
	m_1	m_2	m_3	m_4	m_5	m_6
ff4918007	3	2	2	0	0	1
ff4918010	2	3	2	0	1	1
ff4918012	2	0	3	1	2	2
ff4918013	1	0	0	2	3	1
ff4918014	2	0	1	2	0	2

Table 2
Preference intensity matrix P^1 for the first row of customer action set in Example 1.

	m_1	m_2	m_3	m_4	m_5	m_6
m_1	0	1	1	3	3	2
m_2	−1	0	0	2	2	1
m_3	−1	0	0	2	2	1
m_4	−3	−2	−2	0	0	−1
m_5	−3	−2	−2	0	0	−1
m_6	−2	−1	−1	1	1	0

$$x_{ik} - \sum_{j=i}^{k-1} x_{jj+1} = 0, i = 1, \ldots, T-2 \text{ and } k = i+2, \ldots, T \tag{5}$$
$$-3 \leqslant x_{ij} \leqslant 3, x_{ij} = \text{integer},$$

Consequently, the optimal solution to relaxed Problems (4) and (5) is given by $x^*_{ij} = median\ \{p^l_{ij}, l = 1, \ldots, L\}, i < j$.

Example 3. According to (3)–(5), the target Customer A's optimal preference intensity matrix $\hat{X}$ is as follows in Table 3:

3.3.3. Step 2-3. ordinal scale-based customer profile creation

Finally, the ordinal scale-based customer profile for our recommender systems is created in this step. This requires a series of transformations that begin from the optimal preference intensity matrix. As mentioned above, value of p_{ij} in the optimal preference intensity matrix contains the relative strength/weakness of an individual customer's preference between m_i and m_j. Therefore, by reversely iterating Step 1 to the optimal preference intensity matrix, we arrive at the sorted sequence O_b, which is denoted as $O_b = m_1 \succ m_2 \succ m_3 \succ \cdots \succ m_{|M_b|}$. This indicates that customer b prefers music m_1 to m_2 and prefers m_2 to m_3, and so on. The order that underlies this sorted sequence must be declared in terms of ordinal numbers. For this, we use the term*rank* from Lin et al. [39]. The rank $r(O_b, m_j)$ refers to the cardinal number indicating the position of the music item m_j in O_b. For example, for $O_b = m_1 \succ m_3 \succ m_2, r(O_b, m_2)$ is 3. Using $r(O_b, m_j)$, the sorted sequence O_b is transformed to a sequence of rankings. The sequence of rankings $E_b = \{r(O_b, m_j) | m_j \in M_b.\}$ denotes a set of rankings of music for customer b. In our methodology, E_b is a component of the ordinal scale-based customer profile representing customer$b$'s specific preferences. Therefore, a complete customer profile $U = \{E_1, \ldots, E_Y\}$ is created by iterating these three steps repeatedly for all Y customers (Y is the total number of customers). Given U, the task of CF is performed in Phase 3.

Example 4. From the optimal preference intensity matrix in Example 3, Customer A's $O_A = m_1 \succ m_2, m_3 \succ m_5, m_6 \succ m_4$ and $E_A = \{1, 2, 2, 4, 3, 3\}$ are created. Suppose that there are five more Customers B, C, D, E, and F and that all the steps in Phase 2 are iterated for these customers to make complete customer profiles. The ordinal scale-based customer profile, U, is thus created as follows in Table 4:

3.4. Phase 3: neighborhood formation and recommendation generation

Given the customer profile, U, the CF-based recommendation procedure for a target customer is performed in this phase.

3.4.1. Step 3-1. neighborhood formation

This step computes the similarity between customers and forms a neighborhood between a target customer and a group of like-minded customers. In a previous study, Kamishima [30] suggested an algorithm that treats orders (rankings) as scores by using the GroupLens method [37,48]. For our study, we use this method in a similar manner. The similarities between the target customer a and other customers b are as follows:

Table 3
Optimal preference intensity matrix for Customer A.

	m_1	m_2	m_3	m_4	m_5	m_6
m_1	–	1	1	3	2	2
m_2	–	–	0	2	0	1
m_3	–	–	–	2	1	1
m_4	–	–	–	–	−1	−1
m_5	–	–	–	–	–	0
m_6	–	–	–	–	–	–

Table 4
Sample of the customer profile.

Customer	Music										
	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_{10}	m_{11}
A	1	2	2	4	3	3					
B	2	2	1	3	2	3	1	2		3	1
C	1	3	2		3	1	3	1	2		
D	2	1		4	3	3		1	4		3
E	1	2	3	3	3	2	3	3	4	4	4
F		2		3	2	4		1	4	1	2

$$R_{ab} = \frac{\sum_{m_j \in M_{ab}} (r(O_a, m_j) - \bar{r}_a)(r(O_b, m_j) - \bar{r}_b)}{\sqrt{\sum_{m_j \in M_{ab}} (r(O_a, m_j) - \bar{r}_a)^2} \sqrt{\sum_{m_j \in M_{ab}} (r(O_b, m_j) - \bar{r}_b)^2}}, \tag{6}$$

where $M_{ab} = M_a \cap M_b$ and $\bar{r}_b = \frac{\sum_{m_j \in M_{ab}} (r(O_b, m_j))}{|M_{ab}|}$. Using the similarity measure above, this phase determines which previous customers will be used in the recommendation for the target customer. In our methodology, the best-n-neighbor technique is adopted because of its superior performance over others [24].

Example 5. Suppose that Customer A in Example 4 is a target customer. We can calculate R_{Ab} between Customer A and others by using the customer profile shown in Example 4: $R_{AB} = 0.63, R_{Ac} = 0.30, R_{AD} = 0.81, R_{AE} = 0.70$, and $R_{AF} = 0.43$. Assuming that n = 3 in the best-n-neighbor technique, Customers B, D, and E are then selected to comprise the neighborhood of Customer A.

3.4.2. Step 3-2. recommendation generation

This step derives the *top-N* recommendation from the neighborhood of customers. For each customer, we produce a recommendation list of N music items that the target customer is most likely to purchase. Previously purchased music items are excluded from the recommendation list to broaden each customer's purchase patterns or coverage. The preference for music j by the target customer $a, \hat{r}_{aj}$, is estimated by the following function:

$$\hat{r}_{aj} = \frac{\sum_{b \in \widetilde{M}_j} R_{ab} \times (r(O_b, m_j) - \bar{r}_b)}{\sum_{b \in \widetilde{M}_j} |R_{ab}|}, \tag{7}$$

where $\widetilde{M}_j = \{b | E_b \in U \text{ s.t. } m_j \in M_b\}$. The music items are sorted in an ascending order of these estimated preferences, and the highly ranked N items are recommended.

Example 6. According to the results in Example 5, we can calculate $\hat{r}_{Aj}$ of Customer A for all of the music items: $\hat{r}_{A1} = 1.64, \hat{r}_{A2} = 1.58, \hat{r}_{A3} = 2.07, \hat{r}_{A4} = 3.34, \hat{r}_{A5} = 2.67, \hat{r}_{A6} = 2.63, \hat{r}_{A7} = 2.07, \hat{r}_{A8} = 1.92, \hat{r}_{A9} = 3.74, \hat{r}_{A10} = 3.55$, and $\hat{r}_{A11}$=2.70. Previously purchased items by Customer A, such as $\{m_1, m_2, m_3, m_5\}$, are excluded from the recommendation list to broaden Customer A's purchase patterns. Thus, assuming that N is 3, we can generate a recommendation list $R = \{m_8, m_7, m_6\}$.

4. Experimental evaluation

4.1. Experiments

In this section, we discuss the experimental methods used in this paper. The purposes of the experiments are presented in Section 4.1.1. Experimental designs concerning data collection, benchmark systems, and evaluation metrics are described in Section 4.1.2.

4.1.1. Purposes of experiments

The proposed CoFoSIM is a CF-based recommender system in which customer preference is ordinal scale-based and compromised by the CK method. These characteristics of CoFoSIM could decrease the estimation error and, in turn, enhance the quality of recommendations. To verify our suggestion, the performance of CoFoSIM is evaluated in terms of the *degree of estimation error* and the *quality of recommendation* and is compared with that of other CF-based recommender systems with different rating scales or consensus models. Major questions to be answered through the experiment are as follows:

(1) *Does the use of the ordinal scale result in less estimation error compared with the use of the cardinal scale?*
(2) *Does the estimation error vary depending on which consensus model is selected?*
(3) *Does a decreased possibility of estimation error result in better recommendation quality?*

4.1.2. Experiment design

To answer these questions, we conducted a "live user experiment," measuring the estimation error and the quality of recommendations. Although an alternative off-line analysis has been widely used in previous studies, a live user experiment is much more appropriate for this study for a number of reasons. First, a customer's explicit ratings, required to evaluate the estimation error, can be acquired only from a live user experiment. Second, some experimental attributes that may affect the performance of the system (e.g., system interface, neighborhood size, the number of recommendations, etc.) can be controlled in a live experiment, and thus the experimental results can clearly show whether the performance of the system is affected by the accuracy of preference estimations [53].

4.1.2.1. Benchmark systems. In a live user experiment setting, actual systems are made available to a community of customers, and the performance of those systems was evaluated in real-time. For this, we developed a Web-based CoFoSIM running on a PC with the same interface found in a cell-phone-based CoFoSIM. Similarly, two other recommender systems were developed as a benchmark system. One was a cardinal scale-based recommender system (CS-RS) employing mWUM, wherein user preference was represented on a cardinal scale. The other was an ordinal scale-based recommender system (OS-RS), which also used mWUM. However, in the latter, user preference was represented on an ordinal scale. The mechanism of CS-RS is identical to those suggested in our prior work [11]. Through mWUM, CS-RS first analyzes behaviors of customers for each item and then acquires the cardinal scale-based preference information by summarizing the weights pre-assigned to each behavior type. An ordinal scale-based preference of OS-RS can be acquired by simply converting the cardinal scale-based preference of CS-RS into an ordinal scale-based one. The actual values of ordinal ratings to specific items are determined by comparing the size of the cardinal ratings in CS-RS. The ordinal rankings increase in accordance with the increase in higher cardinal ratings. Coincidentally, the resulting ordinal preference of OS-RS is equivalent to those that could be determined by using the classical consensus model, called Borda's method [9].

As a common factor for these systems, we fixed the neighborhood size of the three systems equally to 10 to prevent this parameter from affecting the quality of CF recommendations [51]. The number of recommendation lists (top-*N*) was fixed to nine, the maximum number displayed on the cell phone screen.

4.1.2.2. Data collection. Using these three systems, the experiment were performed between May 1 and June 18, 2007. A total of 317 real mobile Web customers who had previous experience purchasing music from real mobile Web sites had participated in the experiment. The participants were divided into three individual groups: A, B, and C. Group A participants received only the music list generated by CS-RS, and Group B participants received the music list generated by OS-RS when they accessed the system. Group C participants were for CoFoSIM. To collect the participants' implicit ratings of preference, we asked the participants to log on to each system and freely search and purchase music at anytime during the entire experimental period. All of the behavior histories of the participants were then recorded in the mobile logs after they logged out. The resulting database, which eventually became the source of our preference estimation, consisted of a total of 1619 transactions (as shown in Table 5).

In addition, explicit ratings of customers' preferences were collected after the experimental period. Seventy participants per group were randomly selected to compromise 210 participants for this phase. The participants were provided with five different music items and were asked to represent their actual preference by using a pre-assigned rating scale type particular to that group. For example, participants from Group A rated their preference on a cardinal scale (from 1 to 5), whereas the others used an ordinal scale (from 1 to 5).

4.1.3. Evaluation metrics

In our experiment, two separate metrics were used to evaluate the quality of recommendations and the degree of estimation error. Published metrics for evaluating the quality of recommendations have been developed since the initial work of Resnick et al. [48]. Among these, *recall* and *precision* metrics have been most widely used in recommender system research [11,12,51]. As precision and recall are inversely related, a combination metric called *F1 metric* is used to give equal weight to both precision and recall [12,36,50,51]. We used these three metrics together, and the resulting scores of the three systems were compared.

For evaluating the degree of estimation error, accuracy metrics that measure the gap between the estimated ratings by the system and the actual ratings entered by the user have been widely used [24,51,53]. Commonly used accuracy metrics include mean absolute error (MAE) and correlation [23,51,52]. In our experiment, the accuracies of the systems that use

Table 5
Transaction statistics for the three recommender systems.

System	Participants	Total	Transactions		# of purchases
			# Per system	# Per customer	Per transaction
CS-RS	106		552	5.21	1.32
OS-RS	106	1619	539	5.08	1.55
CoFoSIM	105		528	5.03	1.70

different rating scales (cardinal vs. ordinal) were measured and compared. For such a purpose, MAE is inappropriate because it tends to be lower when preferences are represented on an ordinal scale. For example, suppose that there are four music items and that Customer G rates his explicit preference as (5, 3, 2, and 1). If the estimated preference by the cardinal scale-based recommender system *V* is (5, 4, 4, and 3), the resulting MAE of system *V* will be $(|5-5|+|4-3|+|4-2|+|3-1|)/4 = 1.25$. When converting these cardinal scale-based ratings into ordinal scale-based ones and assuming these as estimated values of an ordinal scale-based system *W*, the resulting estimated preferences will be (1, 2.5, 2.5, and 4). As Customer G's explicit preference can also be represented as (1, 2, 3, and 4), the MAE of system *W* becomes 0.25. At a glance, this result indicates that the estimation of system *W* is five times more accurate, although this could be an inflated interpretation. On the other hand, in the case of Correlation metrics, variants such as the Pearson correlation and Spearman's ranked correlation can be individually used. These two correlation metrics originate from similar basic ideas, and their coefficients range from -1 to 1. Assuming the same situation stated above, the Pearson correlation for system *V* is 0.956 and Spearman's ranked correlation for system *W* is 0.949. The resulting scores are not equal, but compared with the MAE result, the difference between the scores is not large. Therefore, comparing the scores of these two correlation metrics is a more reasonable alternative than using MAE. Despite the slightly lacking mathematical foundation, several studies in various research domains used these two correlation metrics independently for their comparison experiments [8,22,24,41–43,45,55].

4.2. Experimental results

4.2.1. Answer to Question 1: variation of estimation error by rating scales

In order to see whether the ordinal scale could decrease the estimation error better than the cardinal scale, we compared the accuracy of CS-RS and OS-RS by using the *Correlation* metric. This comparison is reasonable because the mechanisms of CS-RS and OS-RS are identical, except that the former represents the preference on a cardinal scale and the latter represents the preference on an ordinal scale. The experimental results are shown in Table 6. As shown in the third row of Table 6, the average correlation metric of OS-RS during the seven weeks is 0.6677, which is approximately 29% higher than that of CS-RS. An independent sample *T*-test was performed to compare the results of OS-RS with those of CS-RS. It shows that the different mean of the correlation metric between the two systems appears to be statistically significant (at $p < 0.01$). It indicates that OS-RS yields a smaller estimation error and is superior to CS-RS in terms of the accuracy of estimation. The superior performance of OS-RS stems from the use of the ordinal scale. As previously discussed, in explicit ratings, a cardinal scale is the more accurate scale type because of its high sensitivity. It enables customers to represent their preferences as closely as possible. However, in our experimental settings, no intervention was made to customer opinions; instead, preferences were indirectly estimated by using insufficient information. Additional customer information such as demographic, psychological, and contextual information was not included in estimating preference. Therefore, the estimated values by using the cardinal scale may stray quite far from the actual preference because the highly sensitive cardinal scale provides a wider range of possible values to be used. This may increase the variance of estimated values, thereby increasing the estimation error. Hence, it can be expected that the ordinal scale, which is a less sensitive scale type, would result in less estimation error in any implicit rating situation.

4.2.2. Answer to Question 2: variation of estimation error by consensus models

To answer Question 2, we compared the correlation metrics of CoFoSIM with those of OS-RS. These two systems used the ordinal scale, although different schemes were applied in formulating the ordinal scale-based preference. In this study, Borda's method was applied to OS-RS, whereas the CK method was applied to CoFoSIM. As shown in the fourth row of Table 6, the average correlation metric of CoFoSIM is 0.7436, which is approximately 11% higher than that of OS-RS. An independent sample *T*-test shows that the different mean of the correlation between CoFoSIM and OS-RS is statistically significant (at $p < 0.01$). It indicates that CoFoSIM yields a smaller estimation error and is superior to OS-RS in terms of accuracy of estimation. As the only difference between these two systems is the consensus model applied, the superior performance of CoFoSIM might stem from the use of the CK method. These results are consistent with those of prior studies in the MCDM field and verify two things: the resulting compromised preference can be slightly different according to the consensus model applied, and the CK method is more advantageous in formulating a more accurate ordinal scale-based preference than

Table 6
Correlation statistics of the three recommender systems.

	Mean	Std. deviation	*t*-value	Sig. (2-tailed)
CS-RS	0.5182*	0.2320		
OS-RS	0.6677**	0.1744	−4.309[a](d.f = 138)	.000
CoFoSIM	0.7436**	0.1423	−2.822[b](d.f = 138)	.005

[a] Compared to CS-RS.
[b] Compared to OS-RS.
* Measured by Pearson correlation.
** Measured by Spearman correlation.

traditional methods such as Borda's method. As mentioned earlier, the CK method reflects the intensity as well as the order of preference in formulating the compromised preference, whereas Borda's method focuses mainly on the order of preference. The intensity of preference implies something beyond the pure order of preference between pairs of items, resulting in a more accurate compromised preference, which in turn results in less estimation error. Therefore, we can conclude that less estimation error would result from the application of the better consensus model.

4.2.3. Answer to Question 3: relationship between the estimation error and recommendation quality

The experimental results in Table 6 show that levels of estimation error in the recommender system may vary depending on the type of scales used and consensus models applied. In order to see whether different levels of estimation error influence the performance of the recommender system, we measured the recommendation quality of the three systems throughout the experimental period. The final results are shown in Table 7.

It is evident in Table 7, that OS-RS shows better performance than CS-RS in all three evaluation metrics. After the seventh period, the precision, recall, and F1-metric of OS-RS are approximately 60%, 15%, and 44% higher than those of CS-RS, respectively. A series of independent sample *T*-tests shows that all of the differences are statistically significant (i.e., the respective *t*-values for the three evaluation metrics are 3.96, 16.25, and 5.43 at $p < 0.01$). Similarly, the precision, recall, and F1-metric of CoFoSIM are approximately 16%, 12%, and 15% higher than those of OS-RS, respectively, at a significance level of 1 percent (i.e., the respective *t*-values are 4.19, 4.33, and 3.54 at $p < 0.01$). The above two results indicates that the performance of CoFoSIM is superior to OS-RS and that the performance of OS-RS exceeds that of CS-RS (i.e., CoFoSIM $\succ$ OS-RS $\succ$ CS-RS with respect to the performance of the systems). This is confirmed by performing additional one-way ANOVA tests for the variations in precision, recall, and F1. The statistics results are ($F = 32.2$, $p < 0.01$), ($F = 9.5$, $p < 0.01$), and ($F = 17.9$, $p < 0.01$), respectively. As previously discussed, the estimation error is the smallest in CoFoSIM and the largest in CS-RS. The sequence is identical with those sorted by the performance of the systems. This indicates that, if the system were to imply less estimation error, the system would yield better performance. Based on these results, we suggest that the level of estimation error could negatively influence the performance of recommender systems.

Therefore, considering the fact that implicit ratings are more required to the mobile Web, we can conclude that CoFoSIM is a viable CF-based recommender system for the mobile Web. CoFoSIM is expected to enhance the quality of recommendations while mitigating customers' burden of explicit ratings. These advantages may eventually lead to increases in the purchase conversion rate of customers.

4.3. Implications, limitations and further studies

Implications of the study findings are now discussed. First, we verified the existence of a relationship between estimation error and the quality of recommendations by using a live user experiment. Numerous CF-related studies have proposed that improved recommendation quality might be guaranteed if preference estimation could be executed more accurately. However, to the best of authors' knowledge, few studies have tried to "verify" these relationships by measuring the actual existence of estimation error in an actual experiment. Second, we also verified the necessity of using an ordinal scale in mobile Web-based CF recommender systems, where gathering a variety of preference information is inconvenient or sometimes impossible. This finding is inconsistent with the findings of most related studies: these studies have assumed the general superiority of more sensitive rating scales such as the cardinal scale. In addition, we also developed a method of applying classical consensus models of MCDM to arrive at an ordinal scale-based preference and to enhance the quality of recommendations.

This study has a few limitations, as well. First, although we verified the existence of a relationship between the estimation error and recommendation quality, the size of the coefficients between the two factors was not calculated. Thus, further study about other factors that might influence the quality of recommendations is required. Second, as a common factor for the recommender systems, the neighborhood size and the number of recommendations were fixed at ten and five because changing them in the middle of a real-time experiment would have been difficult. However, the size of the neighborhood and the number of recommendations have been reported to have a significant impact on recommendation quality [24,52]. Thus, additional iterations might optimize the settings of these factors and improve the performance of CoFoSIM. Third, although only two consensus models were applied and tested, a large number of consensus models with varying characteristics have already been studied. Therefore, applying various consensus models and finding better consensus methods

Table 7
Overall statistics of precision, recall and F1 metric of the three systems.

Measure	Method			*t*-value	
	CS-RS (a)	OS-RS (b)	CoFoSIM (c)	a vs. b	b vs. c
Precision	0.130	0.208	0.241	3.96*	4.19*
Recall	0.339	0.391	0.439	16.25*	4.33*
Fl	0.188	0.272	0.311	5.43*	3.54*

* $p < 0.01$ is the probability that the null hypothesis is true.

for improving the performance would be a worthwhile endeavor. Fourth, although the experiments were performed by simulating the real mobile Web, CoFoSIM's superior performance was based on data sets with relatively few items, customers, and transactions. Applications in the real mobile world typically involve larger numbers of items and customers. Thus, further study is required to test CoFoSIM in various situations where existing data possess different levels of sparsity or scalability. Fifth, in some respects, the precision, recall, and F1scores of our CoFoSIM seem relatively low. However, a tendency toward relatively low levels in these metrics is common to all implicit ratings-based CF recommender systems. Because preference information is not provided by customers but estimated by the system, implicit ratings have an essential weakness in terms of recommendation accuracy; this has been evidenced in many previous CF studies that employed implicit ratings as their profiling technique [10,12,26,27,40,46,54,63]. However, as mentioned in Section 1, implicit ratings are better at overcoming the inconveniences of the mobile Web. In addition, the performance of these systems varies by the datasets used in analyzing customer preference. Our original CoFoSIM used only transaction data, but its performance may increase if other types of implicit data are also used. Thus, to enhance the performance of CoFoSIM, further studies to test the efficient use of diverse datasets would be meaningful. Sixth, our proposed CoFoSIM is a sort of a static approach, which repeats its recommendation process by the addition of one newly provided session record. The quantity of session records used in one update varies by how frequently the system operator updates the customer profile. Different update cycles may diversify the composition of the customer profile, thereby affecting the recommendation quality. Thus, it would be important to find an optimal update cycle for CoFoSIM. Moreover, a trade-off issue between the update frequency and the performance of the system should be considered. It is evident that the update cycle of the customer profile needs to be controlled with the consideration of data sparsity at that time. If the recommender system were to operate with a very sparse customer profile (e.g., near the beginning of the recommendation process), system operators would need to update their customer profile as frequently as possible to rapidly increase the performance of the system. If the customer profile were to become less sparse as time goes by, the system performance would slowly increase while the system overload would rapidly increase. Therefore, it would be better to lessen the frequency of updates to decrease system overload. In this regard, a sensitivity analysis to determine the optimal update cycle according to changes in the sparsity level would be worthwhile.

5. Conclusion

One sign indicating the spread of mobile Web services is the rapid growth of the industry selling music, graphics, games, and other mobile content. However, customers still experience frustration in searching for the content they want because of the peculiar characteristics of the mobile Web, such as inconvenience and high cost. A recommender system resolving these distinct problems is thus needed to provide a pleasant shopping experience to customers. In this paper, we propose a new CF-based recommendation methodology for mobile Web music, called CoFoSIM. Compared to other CF-based recommender systems, CoFoSIM is characterized by three aspects: (1) it implicitly captures preference information by using the mWUM technique. As the user interface of mobile devices is poor and searching activity entails telecommunication costs, implicit ratings that minimize customer intervention may be better for the mobile Web environment; (2) CoFoSIM represents customer preference on an ordinal scale. Admitting that customer preference information on items should be indirectly estimated by implicit ratings, representing such information on an ordinal scale may decrease the estimation error and finally contribute to enhancing the quality of recommendations; and (3) in order to compromise several pieces of preference information, CoFoSIM applies a famous consensus model called the CK method, which has been widely used in the area of MCDM to solve ordinal consensus-making problems. To the best of our knowledge, this method has never been applied to a recommender system. In this regard, this study makes such an application and shows that the method successfully aggregates a great amount of partial preference information into a compromised preference. Unlike other consensus methods such as Borda's method [9], the CK method's ability to incorporate both the order and intensity of preference enables a compromised preference with less estimation error, resulting in better recommendation quality. The experiments to verify the performance of CoFoSIM by using music items and the participation of real mobile Web customers confirm that CoFoSIM produces less estimation error and works better than existing CF-based recommender systems. As a realistic solution for music recommendation problems in the mobile Web environment, CoFoSIM offers the following benefits to both consumers and suppliers of mobile music. (1) Customers will be able to purchase content with much lower connection time on the mobile Web because they will be able to easily find the desired items; and (2) mobile content providers will be able to improve their profitability and revenues because their purchase conversion rate will be improved through increased customer satisfaction.

Acknowledgement

This work was supported by Research Program 2008 of Kookmin University in Korea.

References

[1] C.C. Aggarwal, P.S. Yu, Data mining technique for personalization, Data Engineering 23 (1) (2000).
[2] H.J. Ahn, A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem, Information Sciences 178 (1) (2008) 37–51.
[3] R.D. Armstrong, W.D. Cook, L. Seiford, Priority ranking and consensus formation: the case of ties, Management Science 28 (6) (1982) 638–645.
[4] M. Balabanovic, Y. Shoham, Fab: content-based, collaborative recommendations, Communications of the ACM 40 (3) (1997) 66–72.

[5] C. Basu, H. Hirsh, W.Cohen, Recommendation as classification: using social and content based information in recommendation, in: Proceedings of the 15th National Conference on Artificial Intelligence, 1998, pp. 714–720.
[6] D. Billsus, M.J. Pazzani, Learning collaborative information filters, in: Proceedings of the 15th International Conference on Machine Learning, 1998, pp. 46–54.
[7] J.M. Blin, A linear assignment formulation of the multiattribute decision problem, Review Automatique, Information et Recherche Operationnelle 10 (6) (1976) 21–23.
[8] E. Bofinger, Maximizing the correlation of grouped observations, Journal of the American Statistical Association 65 (December) (1970) 1632–1638.
[9] J.C. Borda, Memoire sur les elections au scrutin, Histoire de l'Academie Royale de Science, Paris, 1981.
[10] L. Chen, F. Hsu, M. Chen, Y. Hsu, Developing recommender systems with the consideration of product profitability for sellers, Information Sciences 178 (4) (2008) 1032–1048.
[11] Y.H. Cho, J.K. Kim, Application of Web usage mining and product taxonomy to collaborative filtering in e-commerce, Expert Systems with Applications 26 (2004) 233–246.
[12] J. Cho, K. Kwon, Y. Park, Collaborative filtering using dual information sources, IEEE Intelligent Systems 22 (3) (2007) 30–38.
[13] M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes, M. Sartin, Combining content-based and collaborative filters in an online newspaper, in: ACM SIGIR'99 Workshop on Recommender Systems, Berkeley, CA, 1999.
[14] W.W. Cohen, R.E. Schapier, Y. Singer, Learning to order things, Journal of Artificial Intelligence Research 10 (1999) 243–270.
[15] M. Condorcet, Essai sur L'Application de L'Analyse a la Probabilite des Decisions Rendues a La Pluralite des Voix, Paris, 1785.
[16] W.D. Cook, Distance-based and ad hoc consensus models in ordinal preference ranking, European Journal of Operation Research 172 (2) (2006) 369–385.
[17] W.D. Cook, Optimal allocation of proposals to reviewers to facilitate effective ranking, Management Science 51 (4) (2005) 655–661.
[18] W.D. Cook, M. Kress, Ordinal ranking with intensity of preference, Management Science 31 (1) (1985) 26–32.
[19] C. Cornelis, J. Lu, X. Guo, G. Zhang, One-and-only item recommendation with fuzzy logic techniques, Information Sciences 177 (22) (2007) 4906–4921.
[20] D. Goldberg, D. Nichols, D.M. Oki, D. Terry, Using collaborative filtering to weave an information tapestry, Communications of the ACM 35 (12) (1992) 61–70.
[21] X. Guo, J. Lu, Intelligent e-government services with personalized recommendation technique, International Journal of Intelligent Systems 22 (5) (2007) 401–417.
[22] J. Hauke, T. Kossowski, Comparison of Values of Pearson's and Spearman's Correlation Coefficient on the Same Sets of Data, Working Paper, 2007.
[23] J.L. Herlocker, Understanding and Improving Automated Collaborative Filtering Systems, Ph.D. Thesis, Minnesota University, 2000.
[24] J. Herlocker, J. Konstan, L. Terveen, J. Riedl, Evaluating collaborative filtering recommender systems, ACM Transactions on Information Systems 22 (1) (2004) 5–53.
[25] W. Hill, L. Stead, M. Rosenstein, G. Furnas, Recommending and evaluating choices in a virtual community of use, in: Proceedings of Conference of Human Factors in Computing Systems, 1995.
[26] Z. Huang, H. Chen, D. Zeng, Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering, ACM Transactions on Information Systems 22 (1) (2004) 116–142.
[27] Z. Huang, D. Zeng, H. Chen, Analyzing consumer-product graphs: empirical findings and applications in recommender systems, Management Science 53 (7) (2007) 1146–1164.
[28] T. Joachims, Optimizing search engine using click through data, in: Proceedings of the Eighth International Conference on Knowledge Discovery and Data Mining, 2002, pp. 133–142.
[29] Y. Kadoya, M. Fuketa, E. Atlam, K. Morita, S. Kashiji, J. Aoe, An efficient e-mail filtering using time priority measurement, Information Sciences 166 (1–4) (2004) 213–229.
[30] T. Kamishima, Nantonac collaborative filtering: recommendation based on order response, in: Proceedings of the Ninth International Conference on Knowledge Discovery and Data Mining, 2003, pp. 583–588.
[31] B.Y. Kang, D.W. Kim, S.J. Lee, Exploiting concept clusters for content-based information retrieval, Information Sciences 170 (2–4) (2005) 443–462.
[32] H. Kazawa, T. Hirao, E. Maeda, SVM and its application to sentence selection, in: Proceedings of 2002 Workshop on Information-Based Induction Sciences, 2002.
[33] P. Kazienko, M. Adamski, AdROSA – adaptive personalization of web advertising, Information Sciences 177 (11) (2007) 2269–2295.
[34] J.G. Kemeny, L.J. Snell, Preference ranking: an axiomatic approach, Mathematical Models in the Social Sciences (1962) 9–32.
[35] C.Y. Kim, J.K. Lee, Y.H. Cho, D.H. Kim, VISCORS: a visual-content recommender for the mobile Web, IEEE Intelligent Systems 19 (3) (2004) 32–38.
[36] J.K. Kim, Y.H. Cho, W.J. Kim, J.R. Kim, J.Y. Suh, A personalized recommendation procedure for Internet shopping support, Electronic Commerce Research and Applications 3 (2002) 301–313.
[37] J. Konstan, B.N. Miller, D. Maltz, J. Herlocker, L.R. Gordon, J. Riedl, GroupLens: applying collaborative filtering to Usenet news, Communications of the ACM 40 (3) (1997) 77–87.
[38] W.S. Lee, Collaborative learning for recommender systems, in: Proceedings of International Conference of Machine Learning, 2001.
[39] W. Lin, S. Alvarez, C. Ruiz, Efficient adaptive-support association rule mining for recommender systems, Data Mining and Knowledge Discovery 6 (2002) 83–105.
[40] D. Liu, C. Lai, W. Lee, A hybrid of sequential rules and collaborative filtering for product recommendation, Information Sciences 179 (20) (2009) 3505–3519.
[41] W.S. Martin, Effects of scaling on the correlation coefficient: additional considerations, Journal of Marketing Research 15 (May) (1978) 304–308.
[42] E.C. May, S.J.P. Spottiswoode, L.V. Faith, A methodological issue in the study of correlation between psycho physiological variables, The Journal of Parapsychology 1 (2001). December.
[43] D.G. Morrison, N.E. Toy, The effect of grouping continuous variables on correlation coefficients, Marketing Science 1 (4) (1982) 379–389.
[44] D.M. Nichols, Implicit rating and filtering, in: Proceedings of the Fifth DELOS Workshop on Filtering and Collaborative Filtering, 1997, pp. 31–36.
[45] J.I. Odiase, S.M. Ogbonmwan, Correlation analysis: exact permutation paradigm, Matematiqki Vesnick 59 (2007) 161–170.
[46] Y. Park, N. Chang, Individual and group behavior-based customer profile model for personalized product recommendation, Expert Systems with Applications 36 (2009) 1932–1939.
[47] M. Pazzani, A Framework for Collaborative, Content-based, and Demographic Filtering, Artificial Intelligence Rev, 1999.
[48] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, J. Riedl, GroupLens: an open architecture for collaborative filtering of Netnews, in: Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work, 1994, pp. 175–186.
[49] E. Rich, User modeling via stereotypes, Cognitive Science 3 (4) (1979) 329–354.
[50] C.J. Rijsbergen, Information Retrieval, second ed., Butterworths, London.
[51] B. Sarwar, Sparsity, Scalability, and Distribution in Recommender Systems, Ph.D. Thesis, Minnesota University, 2001.
[52] B. Sarwar, G. Karypis, J.A. Konstan, J. Riedl, Application of dimensionality reduction in recommender system – a case study, in: Proceedings of the ACM WebKDD Workshop, 2000.
[53] U. Shardanand, P. Maes, Social information filtering: algorithms for automating 'word of mouth', in: Proceedings of the CSCW '98, 1998, pp. 345–354.
[54] Y. Shih, D. Liu, Product recommendation approaches: collaborative filtering via customer lifetime value and customer demands, Expert Systems with Applications 35 (12) (2008) 350–360.
[55] G.A. Simon, Efficacies of measures of association for ordinal contingency table, Journal of the American Statistical Association 73 (363) (1978) 545–551.
[56] L. Terveen, W. Hill, B. Amento, D. McDonald, J. Creter, PHOAKS: a system for sharing recommendations, Communications of the ACM 40 (3) (1997) 59–62.

[57] R. Vezian, D. Militaru, Collaborative filtering: theoretical positions and a research agenda in marketing, Internationl Journal of Technology Management 28 (2004) 31–45.
[58] M. Vozalis, K. Margaritis, Using SVD and demographic data for the enhancement of generalized collaborative filtering, Information Sciences 177 (2007) 3017–3037.
[59] C. Wang, W.A. Wulf, Towards a framework for security measurement, in: Proceedings of National Information Systems Security Conference, 1997, pp. 522–533.
[60] B. Xie, P. Han, F. Yang, R. Shen, H.J. Zeng, Z. Chen, DCFLA: a distributed collaborative-filtering neighbor-locating algorithm, Information Sciences 177 (2007) 1349–1363.
[61] C.H. Yun, M.S. Chen, Mining mobile sequential patterns in a mobile commerce environment, IEEE Intelligent Systems, Man and Cybernetics, Part C: Applications and Reviews 37 (2) (2007) 278–295.
[62] Y. Zhang, J. Jiao, An associative classification-based recommendation system for personalization in B2C e-commerce applications, Expert Systems with Applications 33 (2) (2007) 357–367.
[63] C. Ziegler, S. McNee, J. Konstan, G. Lausen, Improving recommendation lists through topic diversification, in: International World Wide Web Conference (WWW), 2005, pp. 22–32.