Location-based Advertisement Recommendation Model for Customer Relationship Management under the Mobile Communication Environment

Hyunchul Ahn, Kyoung-Jae Kim, Ingoo Han

Location-based advertising or application has been one of the drivers of third-generation mobile operators’ marketing efforts in the past few years. As a result, many studies on location-based marketing or advertising have been proposed for recent several years. However, these approaches have two common shortcomings. First, most of them just suggested the theoretical architectures, which were too abstract to apply it to the real-world cases. Second, many of these approaches only consider service provider (seller) rather than customers (buyers). Thus, the prior approaches fit to the automated sales or advertising rather than the implementation of CRM. To mitigate these limitations, this study presents a novel advertisement recommendation model for mobile users. We call our model MAR-CF (Mobile Advertisement Recommender using Collaborative Filtering). Our proposed model is based on traditional CF algorithm, but we adopt the multi-dimensional personalization model to conventional CF for enabling location-based advertising for mobile users. Thus, MAR-CF is designed to make recommendation results for mobile users by considering location, time, and needs type. To validate the usefulness of our recommendation model, we collect the real-world data for mobile advertisements, and perform an empirical validation. Experimental results show that MAR-CF generates more accurate prediction results than other comparative models.

Keywords: Mobile Recommender System, Location-based Advertising, Collaborative Filtering, Needs Type
I. Introduction

Mobile devices such as phones, PDAs, GPS receivers, and other small devices are everywhere, and their use has begun to impact how individuals manage their increasingly fast-paced lives. For mobile service operators, this trend can be great market opportunities. As a result, many companies are being interested in marketing and sales using mobile devices. In particular, location-based advertising or application has been one of the drivers of third-generation (3G) mobile operators’ marketing efforts in the past few years [UMTS Forum, 2000].

However, the prior approaches for location-based marketing or advertising in both academic and practical areas are quite disappointing [Brunato and Battiti, 2002; Tewari et al., 2002; Yuan and Tsao, 2003; Pousman et al., 2004; Kim et al., 2005; Kwon et al., 2005]. The most critical limitations of these approaches can be summarized into two facts. One is that most of their approaches were not practically useful methodology, but just suggestion of the theoretical architecture. And, the other is that many of their approaches only consider service provider (seller) rather than customers (buyers). Thus, the prior approaches seem to be automated sales or advertising rather than implementations of CRM.

To overcome these limitations, this study presents a novel recommendation model for mobile users. Our model is based on collaborative filtering (CF) - the most frequently used recommendation scheme, but it produces recommendation results using the location information of users. However, we modify the concept of the conventional CF in order to apply it to location-based advertising for mobile users. To do this, we adopt the multi-dimensional personalization model, which is proposed by Schilke et al. [2004]. This model proposes three dimensions to be considered for personalization of mobile users - location, time, and interest - as shown in <Figure 1>.

![Figure 1](image-url)
Among these dimensions, location and time are easily defined operationally because these dimensions have commonly applied basis such as administrative territory, hours, minutes, and so on. In addition, the location as well as the time of usage of the mobile phone users can be technically traced. The mobile phones in service always communicate the base stations of mobile service providers, so the service providers can acknowledge when and where their users are using their mobile phones. As a result, in the case of ‘location’ and ‘time’, it is easy to use the information for generating recommendation result. However, it is never easy to inject the notion of ‘interest’ into a recommendation model. Generally, interests of users are different from each other, so it is very hard to define or classify ‘interest’ in the concrete form. Moreover, it is difficult to get the information on the user’s interests. Because the types of interests are very diverse, it is hard to make inquiries about interests of users. Thus, we simplify the types of interests by adopting the concept of user’s needs type – hedonic, utilitarian, or neutral – from marketing literatures [Hirschman and Holbrook, 1982; Babin, et al., 2000; Chandon et al., 2000]. As a result, our study proposes a novel CF algorithm that is designed to make recommendation results for mobile users by considering location, time, and needs type. To validate the usefulness of our recommendation model, we collect the real-world data for mobile advertisements, and perform an empirical validation.

The rest of the paper is organized as follows. Section 2 briefly reviews the basic concepts of collaborative filtering and user’s needs type, and the next section proposes our research model - Mobile Advertisement Recommender using Collaborative Filtering (MARCF). In section 4, the explanation for the experimental design and results are presented. In the final section, the conclusions of the study are presented.

II. Collaborative Filtering and User’s Needs Type

As mentioned in the previous section, our study presents CF recommendation model that considers mobile users’ location, time, and needs type. Thus, in this section, we briefly review the theoretical backgrounds of two core components in our recommendation model - collaborative filtering and user’s needs type.

2.1 Collaborative filtering as a recommendation method

A recommendation system is one of the ways that such knowledge can be represented. It can be defined as the automated and sophisticated decision support systems that are needed to provide personalized solution in a brief form without going through a complicated search process. It helps users select from available products or services based on their requirements and preferences. Although many different approaches have been applied to the problem of making more accurate and efficient recommender systems, there are two dominant types: content-based and collaborative filtering.

In the content-based (CB) approach, the system analyzes contents of items, and creates a profile that is a representation of a user’s interest in terms of items. Then, the system analyzes contents of items unknown to the user, and
comparing it with his/her profile. Finally, the system constructs recommendation results with the new items that are likely to satisfy the user [Balabanovic and Shoham, 1997]. To put it simply, the key of CB is how to characterize items. However, it is never easy to characterize the items, since it usually requires a manual process using human intelligence. As a result, the application of the CB approaches has been restricted to a few domains such as recommendation of news articles or web pages. Due to this critical limitation, collaborative filtering (CF) is preferred to CB approaches as a recommendation method in practical applications [Billsus and Pazzani, 1998].

The CF method sidesteps the problem of CB by recommending items that other people who seem to have similar preference patterns have liked [Lawrence et al., 2001]. A CF system collects explicit user ratings of items in question (e.g., movie, music, book, and so on). Users are then compared based on how similar their ratings are, and the system constructs recommendation results with the items favored by other people who have similar interests. In summary, CB approaches are based on the similarity between items, however, CF approaches use the similarity between users.

The task of collaborative filtering is to predict the preference of an active user for a target item based on user preference. There are two general classes of collaborative filtering algorithms: the memory-based and the model-based approach. The memory-based CF approach repeatedly scans the preference database (user-item matrix) to locate the peer groups for an active user. A prediction is then computed by weighting the votes of users in the peer groups. The people in the peer groups are identified based on their similarity or nearness in tastes to the active user. Consequently, this method is equivalently called the correlation-based or nearest-neighbor collaborative filtering method.

The model-based CF approach infers a user model from the database of rating histories. The user model is then consulted for predictions. This approach produces predictions in a shorter time in comparison to the memory-based approach. However, it generally requires more time to train the dataset. Moreover, it is not suitable for environments in which user preference models must be updated rapidly or frequently [Schafer et al., 2001]. Thus, in this paper, we will focus on the memory-based CF algorithm.

In general, the memory-based CF algorithm recommends items to an active user according to the following steps:

**Step 1. Similarity calculation**

Similarity between an active user and his/her neighbor is computed using Pearson correlation coefficient, which is defined in the Equation (1).

\[
S_{a,u} = \frac{\sum_i (r_{a,i} - \overline{r_a}) \cdot (r_{u,i} - \overline{r_u})}{\sqrt{\sum_i (r_{a,i} - \overline{r_a})^2} \cdot \sqrt{\sum_i (r_{u,i} - \overline{r_u})^2}}
\]

(where \(-1 \leq S_{a,u} \leq 1\))

where \(S_{a,u}\) is the similarity between the active user \(a\) and each of the other users (\(u\)) who have the co-rated items with the active user \(a\). \(i\) is the index of each item that both user \(a\) and user \(u\) have rated, \(r_{a,i}\) is the rating of user \(a\)
for item $i$, $r_{u,i}$ is the rating of user $u$ for item $i$, $\overline{r_a}$ is the average rating of user $a$, and $\overline{r_u}$ is the average rating of user $u$.

**Step 2. Neighbor selection**

In this step, $n$ neighbors who have the highest similarity with the active user are selected. The similarities between the active user and other users calculated in step 1 are used as a criterion for selecting nearest neighbors.

**Step 3. Prediction**

In step 3, the system computes a prediction for the active user’s unanswered rating from a combination of the selected neighbors’ ratings. The predicted numerical rating of the active user $a$ for a target item $x$ ($\hat{r}_{a,x}$) can be calculated by Equation (2).

$$\hat{r}_{a,x} = \overline{r_a} + \sum_{\nu \in N} (r_{\nu,x} - \overline{r_{\nu}}) \cdot \frac{S_{a,\nu}}{\sum_{\nu \in N} |S_{a,\nu}|} \quad (2)$$

where $\overline{r_a}$ is the average rating of user $a$, $\nu$ is the index of each nearest neighbor, $N$ is the set of the nearest neighbors for user $a$, $r_{\nu,x}$ is the rating of the user $\nu$ for item $x$, $\overline{r_{\nu}}$ is the average rating of user $\nu$, and $S_{a,\nu}$ is the similarity between the active user $a$ and user $\nu$.

The most important advantage of memory-based CF is that it doesn’t require any effort for characterizing items, and it can provide serendipitous recommendations because it is not based on contents of items. In addition, it is suitable for the environments that require up-to-date recommendation results because its database is continually updated.

### 2.2 User’s Needs Type

Needs are defined as requirements for something essential or desirable that is lacking. That is, needs are the most fundamental factors and the starting point of the process generating behavioral outcomes. Thus, understanding a user’s needs at the point of usage is so important to facilitate his/her satisfaction. Some prior studies in marketing literature have identified numerous kinds of needs which influence the process to stimulate people’s behavior, however, it is usual to classify needs into two types: utilitarian and hedonic [MacInnis and Jaworski, 1989].

Utilitarian needs are defined as requirements for products that remove or avoid problems, while hedonic needs are requirements for products that provide social or aesthetic utility. For example, a user who uses a virtual community for obtaining useful information has utilitarian needs, but he/she has hedonic needs when he uses it for social relationship or amusement. The needs type is known to be stimulated by the advertised message (cue), thus advertisers may use utilitarian or hedonic appeals to stimulate consumers’ utilitarian or hedonic.

### III. Research Model

Our recommendation model is basically based on collaborative filtering (CF) algorithm. In general, CF works by building user-item matrix, which contains all the information about the satisfaction levels of users for target items. However, CF in our study should consider other additional information—location, time, and user’s needs type. Consequently, we define a
(a) Conventional scheme of user-item matrix

(b) New scheme of user-item matrix in our study

Figure 2> User–item matrix schemes of conventional CF and the proposed model.
user, item, and corresponding satisfaction level. However, our new scheme should contain three additional information - location, time and needs type. Thus, we modify the traditional user-item matrix as presented in Figure 2 (b). For the information on location, we adopt it as a higher dimension of items because items of our study would be commercial spots located in a specific area (location). Thus, when applying this model to actual recommendation, it is easy to filter items on the basis of location. In the case of time, we specify it into two factors - visiting day and visiting time. In general, the spots we visit become different according to the day of visit (e.g. weekday or weekend) and the time frame of visit (e.g. morning, lunch time, afternoon, or dinner time). As a result, we design our recommendation model to consider both the day and the time of visit independently. Also, we include needs type as one of the additional information that affects user's satisfaction level.

Figure 3 represents the overall process of our proposed recommendation model. For convenience, we call our recommendation model MAR-CF (Mobile Advertisement Recommender model using Collaborative Filtering) hereafter. As shown in Figure 3, the procedure of MAR-CF consists of three steps. The detail explanation for each step is presented as follows:

**Step 1. Filtering**

In step 1, the information about the target user is inputted into MAR-CF system. This information includes the identification of the target user, his or her location (visiting area), current time, and user's current needs type. The information on the user's current needs type is collected by directly inquiring the user, however, other information can be obtained automatically.

After collecting required information on the target user, MAR-CF searches for candidate items for recommendation. At this time, it fil-
ters items in the areas that are located in the areas far from the target user’s current position. By doing this, MAR-CF system can reduce the search space dramatically, and it can improve efficiency of recommendation process.

**Step 2. CF**

After step 1, CF is performed for finding similar neighbors to the target user, and calculating the expected satisfaction level for the items that are inexperienced by the target user. In the conventional CF, the similarity between the users is calculated by using Pearson correlation of the satisfaction levels (ratings) between the users as presented in Equation (1). And then, the expected satisfaction levels for the items of the target user are calculated by applying Equation (2). However, in our MAR-CF model, other additional information like time, needs type should also be considered when calculating similarity and the expected satisfaction level. Thus, MAR-CF system uses ‘adjusted Pearson correlation’ when calculating similarity between users. It can be expressed as an equation form like following Equation (3).

$$S_{a,u} = S_{a,u} \times w_{a,u}$$

(3)

where $S'_{a,u}$ is the adjusted Pearson correlation between the active user $a$ and each of the other users ($u$), $S_{a,u} (1 \leq S_{a,u} \leq 1)$ is Pearson correlation of users’ ratings presented in Equation (1), and $w_{a,u} (0 \leq w_{a,u} \leq 1)$ is the similarity of time(visiting day and visiting time) and corresponding needs type between user $a$’s current status and the user $u$’s inputted status.

By using this equation, the similarity between users can be more refined by considering multi-dimensional factors including time, needs type, and the pattern of satisfaction levels. As a result, more sophisticated recommendation results may be generated for mobile users.

**Step 3. User Interface**

In the last step, the optimal recommendation results produced by MAR-CF are provided to the target user. The results can be transferred by using SMS or MMS. If the system uses MMS, it may include more detail information for the recommended spot such as simple map, photos, and long messages, although SMS can only send 80 alphabetical characters.

**IV. Experiments and Results**

**4.1 Experimental design**

In order to validate the usefulness of our MAR-CF model, we apply ‘empirical validation’ that is based on real-world data. Although our proposed model is based on CF, we cannot use public datasets for CF such as MovieLens, or EachMovie because our model is sophisticated on mobile advertising using additional information on location, time, and user’s needs type. Consequently, we build a Web-based system for collecting appropriate data from mobile users. This data collection system contains the places for shopping, eating, drinking, enjoying, and learning in five major commercial zones of Seoul, Korea. The system totally contains the information on 275 places in Chongro (종로), Daehakro(대학로), Shinchon/Ewha Univ. (신촌/이대), Kangnam Station(강남역), and My
ungdong(영동) areas. And, it is designed to collect data for these spots from actual mobile phone users. Figure 4 shows the input screen for the data collection system. As shown in the figure, our system asks the latest visiting day, visiting time, the objective of visiting (user’s needs type at the point of visit), and satisfaction level for each place.

To simplify the input process, we discretize the candidate values of input variables as presented in Table 1. As shown in Table 1, we assign the numeric code in an interval scale to each candidate value of the most input variables (visiting time, needs type, and ratings). Consequently, it is possible to apply simple numeric operations for the inputted values.

In order to collect the experimental dataset, we operate the data collection system from April to May in 2006. As a result, we collect 9980 ratings from 265 respondents in three universities in Korea. Among them, we eliminate some cases that are seemed to be distorted, and finally select 208 respondents and their ratings for 175 items as an experimental dataset.

In order to apply our model to this real-world dataset, we should first determine how to measure in Equation (3). As explained, our collected dataset has the numeric values for visiting day, visiting time and needs types, which are measured in interval scale. Using this property, we define $w_{kn}$ as following equations.
<table>
<thead>
<tr>
<th>Dimension</th>
<th>Variable</th>
<th>Assigned numeric code</th>
<th>Candidate values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Visiting day</td>
<td>1</td>
<td>Weekday (Mon.-Fri.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Weekend (Sat./Sun.)</td>
</tr>
<tr>
<td></td>
<td>Visiting time</td>
<td>1</td>
<td>Morning / AM 08:00 ~ AM 11:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Lunch / AM 11:00 ~ PM 02:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Afternoon / PM 02:00 ~ PM 05:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>Dinner / PM 05:00 ~ PM 08:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>Night / PM 08:00 ~ PM 11:00</td>
</tr>
<tr>
<td>Needs type</td>
<td>Needs type</td>
<td>1</td>
<td>Hedonic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Utilitarian</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>Ratings level</td>
<td>1</td>
<td>Very dissatisfied</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Dissatisfied</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Somewhat dissatisfied</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>Somewhat satisfied</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Satisfied</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>Very satisfied</td>
</tr>
</tbody>
</table>

\[
 w_{a,u}^X = \frac{w_{a,u}^{\text{day}} + w_{a,u}^{\text{time}} + w_{a,u}^{\text{need}}}{3} \quad (4)
\]

where \( w_{a,u}^X \) means the similarity between the active user \( a \) and each of the other users \( (u) \) from the viewpoint of ‘visiting day’, \( w_{a,u}^{\text{time}} \) means the similarity between the active user \( a \) and each of the other users \( (u) \) from the viewpoint of ‘visiting time’, and \( w_{a,u}^{\text{need}} \) means the similarity between the active user \( a \) and each of the other users \( (u) \) from the viewpoint of ‘needs type’.

In above equation, \( w_{a,u}^{\text{day}} \), \( w_{a,u}^{\text{time}} \) and \( w_{a,u}^{\text{need}} \) can be calculated by using Equation (5).

\[
 w_{a,u}^X = \frac{MAX^X - MIN^X - \text{diff}_{a,u}^X}{MAX^X - MIN^X} \quad (5)
\]

where \( w_{a,u}^X \) means the similarity between the active user \( a \) and each of the other users \( (u) \) from the viewpoint of variable \( X \), \( MAX^X \) means the maximum possible value of the numeric code for variable \( X \), \( MIN^X \) means the minimum possible value of the numeric code for variable \( X \), and \( \text{diff}_{a,u}^X \) means the difference of the numeric code for variable \( X \) between the active user \( a \) and each of the other users \( (u) \).

In Equation (5), the value of \( W_{a,u}^X \) becomes the maximum value (i.e. 1) when \( \text{diff}_{a,u}^X \) becomes 0. In contrast, it becomes the minimum value (i.e. 0) when \( \text{diff}_{a,u}^X \) becomes the maximum (i.e. \( MAX^X - MIN^X \)). In order to combine these functions into conventional CF algorithm, we develop our own private experimental software. This software is implemented using
Microsoft Excel 2003 and its VBA (Visual Basic for Applications). In addition, in order to overcome the scarcity of the items, we use ‘All but 1’, which means that the test set for each test user contains a single randomly selected rating and the observed set (i.e. training set) contains the rest of the ratings.

In addition, to test the effectiveness of MAR-CF model, we also apply three different types of CF models to the same dataset. The first type is the conventional approach of CF. This model is designed to ignore all the information except for on user, item and corresponding ratings. Thus, this model generates recommendation results only considering the pattern of ratings.

The second type is the filtering approach for each variable—visiting day, visiting time, and needs type. In this approach, the recommender model only considers neighbors who have same value for each situational factor (i.e. visiting day, visiting time, or needs type) when calculating similarities between users. Because it does not consider other users that have different values for situational factors at all, the efficiency of the model can be improved. However, this type of recommender models may suffer from the problem of information loss.

The third type is the combination approach of two variables. As presented in Equation (4), our proposed model compromises the values of all the situational factors with equal importance. However, in this type of comparative models only compromises two factors of total situational factors with equal importance (i.e.0.5). As a result, there are three sub-comparative models in this type: (1) combination of ‘visiting day’ and ‘visiting time’, (2) combination of ‘visiting day’ and ‘needs type’, and (3) combination of ‘visiting time’ and ‘needs type’.

4.2 Experimental results

In this study, we set the average MAE (mean absolute error) as the criterion for evaluating performances of the comparative models. The MAE is frequently used in CF literature, and represents the difference between the predicted and actual rating of users [Breese et al., 1998; Sarwar et al., 1998; Goldberg et al., 2001]. Average MAE can be defined as Equation (6).

\[
\text{Avg.MAE} = \left( \frac{1}{N} \sum_{k=1}^{N} \left( \left( \sum_{i=1}^{n} |p_{k,i} - a_{k,i}| / n \right) \right) / N \right)
\]

where \(N\) is the number of users in test dataset \(T\), \(n\) is the number of items in test dataset \(T\), \(p_{k,i}\) is the predicted ratings of user \(k\) for the item \(i\), and \(a_{k,i}\) is the actual ratings of user \(k\) for the item \(i\).

<Table 2> presents overall results of the comparative models and our proposed model. As shown in the table, our proposed model, MAR-CF, shows the minimal average MAE among the comparative models. Thus, we may conclude that our model generates the most accurate prediction results in the recommendation for mobile users. In addition, we can find that the comparative models that apply filtering approach shows very unsatisfactory prediction accuracy. The reason of this phenomenon seems to be information loss due to filtering. We suspect that this may hinder the models from finding generalized patterns of various neighbors.

To examine whether the differences of pre


<table>
<thead>
<tr>
<th>Type</th>
<th>Model name</th>
<th>Variables to be considered</th>
<th>Mean of MAE</th>
<th>S.D. of MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional</td>
<td>PureCF</td>
<td>None</td>
<td>0.9282</td>
<td>0.2811</td>
</tr>
<tr>
<td>Filtering</td>
<td>FD-CF</td>
<td>Visiting day</td>
<td>0.9394</td>
<td>0.2832</td>
</tr>
<tr>
<td></td>
<td>FT-CF</td>
<td>Visiting time</td>
<td>0.9748</td>
<td>0.2719</td>
</tr>
<tr>
<td></td>
<td>FN-CF</td>
<td>Needs type</td>
<td>0.9404</td>
<td>0.2824</td>
</tr>
<tr>
<td>Combination</td>
<td>CDT-CF</td>
<td>Visiting day</td>
<td>0.9271</td>
<td>0.2824</td>
</tr>
<tr>
<td></td>
<td>CDN-CF</td>
<td>Visiting day</td>
<td>0.9270</td>
<td>0.2823</td>
</tr>
<tr>
<td></td>
<td>CTN-CF</td>
<td>Visiting time</td>
<td>0.9275</td>
<td>0.2808</td>
</tr>
<tr>
<td>Proposed model</td>
<td>MAR-CF</td>
<td>Visiting day</td>
<td>Visiting time</td>
<td>Needs type</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FD-CF</th>
<th>FT-CF</th>
<th>FN-CF</th>
<th>CDT-CF</th>
<th>CDN-CF</th>
<th>CTN-CF</th>
<th>MAR-CF</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.693***</td>
<td>8.153**</td>
<td>3.451***</td>
<td>1.232</td>
<td>0.963</td>
<td>0.976</td>
<td>1.682***</td>
</tr>
<tr>
<td>5.312**</td>
<td>0.233</td>
<td>4.074***</td>
<td>3.997***</td>
<td>3.817***</td>
<td>4.107***</td>
<td></td>
</tr>
<tr>
<td>5.766***</td>
<td>8.278***</td>
<td>8.290***</td>
<td>8.187***</td>
<td>8.331***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.886***</td>
<td>4.050***</td>
<td>3.759***</td>
<td>4.037***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.204</td>
<td>0.425</td>
<td>0.638</td>
<td>0.596</td>
<td>0.315</td>
<td>1.2801</td>
<td></td>
</tr>
</tbody>
</table>

* statistically significant at the 10% level, ** statistically significant at the 5% level, *** statistically significant at the 1% level.

Predictive accuracy between the proposed model and other comparative algorithms are statistically significant or not, we apply the t-test for paired samples. This test is usually applied when the two sets of values are from the same sample, such as in a pre-test/post-test situation. It is sometimes called the t-test for correlated samples or dependent samples [Green et al., 2000]. <Table 3> shows the result for the paired-samples t-test. As shown in <Table 3>, MAR-CF outperforms PureCF at the 5% statistical significance level, and all of the filtering-based comparative models at the 1% statistical significance level. We can also find that MAR-CF outperform CTN-CF at the 10% statistical significance level. However, it does not outperform CDT-CF and CDN-CF with statistical significance.

V. Conclusion

In this study, we propose a novel CF algorithm for mobile advertisement recommendation. MAR-CF, our proposed model, is designed to perform CF with consideration of the important dimensions for the personalization of
mobile users' location, time and user's needs type (interest). To validate the usefulness of MAR-CF, we collect primary data from actual users in real-world, and conduct an experiment for empirical validation. Experimental results show that MAR-CF outperforms conventional CF algorithm as well as other comparative models.

In reality, MAR-CF can be implemented under the system architecture presented in <Figure 5>. In addition, <Figure 6> presents the sample scenario of the real-world application of MAR-CF. As presented in these figures, our proposed model can be applied by mobile service operators as a new business model, which is based on 'mobile advertising' and 'permission marketing'.

However, our study also has limitations. First of all, the problem of data scarcity should be resolved. Although we proceed with the data collection on the Web for about a month, the collected dataset still have insufficient ratings. This may cause so-called 'sparsity problem', which means the problem of low-quality recommendations when the system has a few ratings of users, since the users' patterns for measuring the similarity between users become unclear. Consequently, the efforts to mitigate the sparsity problem should be made in the future research.

Second, the relative importance (i.e. importance weights) of each situational variable (visiting day, visiting time, and needs type) should be refined. Currently, our proposed model just uses equal weight for these variables assuming that these components are equally important. However, this basic assumption is very unrealistic. Thus, the methods to differentiate the importance weight of each situational variable should be researched.

Third, our proposed system is designed to directly ask user's current needs type, however
it may cause inconvenience since users should answer the question to get recommendation results. Thus, future research should focus on efforts to predict the current needs of users from implicit data like mobile logs. The recent studies on the context-awareness in an online environment are expected to contribute to this future research direction.

Finally, the usefulness of MAR-CF should be validated in practice. The validation process in our study is quite restricted because our model is not validated in the real-world mobile situation, although the experimental validation is performed using the data collected from real-world users. Thus, in the future, the applicability of MAR-CF should be validated practically by a real-world mobile service provider.

References


[5] Brunato, M., and Battiti, R., "PILGRIM: A Location Broker and Mobility-Aware Re-


이동통신 환경하에서의 고객관계관리를 위한 지역화고 추천 모형

◆ 저자소개 ◆

안현철 (Ahn, Hyunchul)
현재 KAIST 테크노경영대학원 박사후 연구원으로 일하고 있다. KAIST에서 산업경영학을, 그리고 KAIST 테크노경영대학원에서 경영정보시스템 전공으로 경영공학석사와 경영공학박사를 취득하였다. 주요 관심분야는 인공지능 및 데이터마이닝, 고객관계관리, m-CRM, 재무정보시스템 등이다.

김경재 (Kim, Kyoung-Jae)

한인구 (Han, Ingoo)
현재 KAIST 테크노경영대학원 교수로 재직 중이다. 서울대학교 국제경제학사, KAIST 경영과학석사를 취득하였고, University of Illinois at Urbana-Champaign에서 회계정보시스템을 전공하여 경영학박사를 취득하였다. 주요 관심분야는 지능형 신용평가시스템, 인공지능을 이용한 재무예측, 지식기반 가치평가, 정보시스템 감사 및 보안 등이다.

◆ 본 논문은 지난 2006 한국경영정보학회 추계학술대회에서 최우수논문상을 수상했던 논문을 경영정보학연구 학술지에 적합하도록 내용을 편집, 보강하여 경영정보학연구 학술위원회에 의해 12월 7일 게재확정된 논문임을 밝힙니다.