Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2013, Article ID 957184, 10 pages
http://dx.doi.org/10.1155/2013/957184

Research Article

A Novel Nonadditive Collaborative-Filtering Approach Using Multicriteria Ratings

Yi-Chung Hu

Department of Business Administration, Chung Yuan Christian University, Chung-Li 32023, Taiwan

Correspondence should be addressed to Yi-Chung Hu; ychu@cycu.edu.tw

Received 13 June 2013; Accepted 23 September 2013

Academic Editor: Jung-Fa Tsai

Although single-criterion recommender systems have been successfully used in several applications, multicriteria rating systems which allow users to specify ratings for various content attributes of individual items are gaining importance in recommendation context. An overall rating of an unrated item is often obtained by the weighted average method (WAM) when criterion weights are available. However, the assumption of additivity for the WAM is not always reasonable. For this reason, this paper presents a new collaborative-filtering approach using multicriteria ratings, in which a nonadditive technique in Multicriteria decision making (MCDM), namely, the Choquet integral, is used to aggregate multicriteria ratings for unrated items. Subsequently, the system can recommend items with higher overall ratings for each user. The degrees of importance of the respective criteria are determined by a genetic algorithm. In contrast to the additive weighted average aggregation, the Choquet integral does not ignore the interaction among criteria. The applicability of the proposed approach to the recommendation of the initiators on a group-buying website is examined. Experimental results demonstrate that the generalization ability of the proposed approach performs well compared with other similarity-based collaborative-filtering approaches using multicriteria ratings.

1. Introduction

It is well known that personalized recommender systems can avoid information overload by providing items which are more relevant to consumers [1–3]. These recommended items with higher predicted overall ratings enable greater cross-selling to be achieved. In particular, single-criterion recommender systems have been successful in a number of personalization applications. The key property of such systems is that users are required to offer only a single-criterion or overall rating for each consumed item. In other words, users cannot express their preference for each criterion of these items. However, practical problems are often characterized by several criteria [4]. Recommender systems should benefit from leveraging multicriteria information because it can potentially improve recommendation accuracy [5]. Several online sites, such as Zagat's Guide for restaurants, Buy.com, and Yahoo! Movies, provide a recommendation service that uses multicriteria ratings for each object. In contrast to Yahoo! Movies, the rating on each criterion for a restaurant in Zagat's Guide is the same for all users, which means that multicriteria ratings in Zagat's Guide are not personalized. The uniqueness of Yahoo! Movies indicates that personalization multicriteria recommender systems may become an important component in personalization applications.

Multicriteria rating problems could be an important issue for the next generation of recommender systems [5, 6]. Multicriteria recommender systems have been addressed in several approaches. For instance, Schafer [7] presented a metarecommendation system with DynamicLens which lets users express their preference and relative importance for each criterion. As a result, the system will filter the preferred items on the basis of their requirements. Lee et al. [8] proposed intelligent agent-based systems for personalized recommendations. Because they assume that the value or rank of each criterion is the same for all users, such systems do not take personalization into account. Ricci and Werthner [9] employed case-based querying to recommend travel planning with multiple criteria (e.g., location and activity).

These recommendation systems do not allow users to specify their subjective perception of the various criteria of individual items.

From the viewpoint of multicriteria decision making (MCDM), the overall rating has a certain relationship with the multicriteria ratings for an item. For this, in contrast to the multicriteria recommendation systems previously described, Adomavicius and Kwon [5] presented a framework for an aggregation-function-based approach to leverage multicriteria rating information, in which the rating for each criterion can be estimated by the traditional similarity-based approach using single-criterion ratings. Then the output of an aggregation function is regarded as a predicted value of the overall rating of an unrated item. The weighted average method (WAM) is often used as an aggregation function and can aggregate ratings on different criteria when the criterion weights are available [5, 10]. Note that the criterion weights in WAM are interpreted as the relative importance of the criteria. The WAM is a simple decomposed method and assumes that criteria do not interact [11]. However, because the criteria are not always independent, the assumption of additivity is not always reasonable [12] and may affect recommendation performance. Thus, this motivates us to use a nonadditive technique, the Choquet integral [13–16], as an aggregation function. Actually, the Choquet integral is a generalization of the weighted average [17, 18]. To address this, this paper further proposes a novel nonadditive multicriteria recommendation method on the basis of the Choquet integral. Furthermore, because the goal of the proposed approach is to recommend correctly a set of a few relevant items to each user, a variation of the popular F-measure is presented to evaluate recommendation performance. To achieve high accuracy, a genetic algorithm (GA) is implemented to determine the parameter specifications.

The remainder of the paper is organized as follows. Section 2 briefly introduces several collaborative-filtering approaches using multicriteria ratings, including the similarity-based and the aggregation-function-based approaches. Section 3 describes the proposed aggregation-function-based collaborative-filtering approach using the Choquet integral and presents an accuracy metric for recommendation performance evaluation. A GA-based method for constructing a nonadditive recommendation model is demonstrated in Section 4. Section 5 applies the proposed nonadditive approach to initiator recommendation on a group-buying website in Taiwan. Section 6 contains the discussion and conclusions.

2. Collaborative-Filtering Approaches Using Multicriteria Ratings

Collaborative-filtering approaches rely on the ratings of a user as well as those of other users in the system [19]. The key idea is to recommend items that users with similar preferences have liked in the past. Traditional collaborative-filtering approaches using single-criterion rating can be categorized into two classes: neighborhood-based and model-based approaches. Adomavicius and Kwon [5] presented the similarity-based and aggregation-function-based methods within neighborhood-based approaches. The latter is the focus of this paper.

2.1. Similarity-Based Approach. Assume that a system asks each user to offer feedback on n criteria with respect to a consumed item or a person with whom he or she has a connection. Let R_0 denote the set of possible overall ratings, and let R_i denote the set of possible ratings for each individual criterion ($1 \le i \le n$). For the (user, item) pairs, the rating function R in a multicriteria recommender system is defined as follows:

$$R\,(\text{user}, \text{item}) \longrightarrow R_0 \times R_1 \times \cdots \times R_n. \tag{1}$$

For instance, for simplicity only the reputation (criterion 1) and the response (criterion 2) are used to evaluate an initiator on a group-buying website (i.e., $n = 2$) for the initiator recommendation. User Randy might assign ratings of 5, 7, and 6 to the reputation, the response, and the overall rating, respectively, for initiator Ryan. Of course, it is necessary that Randy has already joined the group confirmed by Ryan. Therefore, R(Randy, Ryan) can be denoted by ($r_0^{\text{Randy, Ryan}}$, $r_1^{\text{Randy, Ryan}}$, $r_2^{\text{Randy, Ryan}}$) = (6, 5, 7). If user Frances has not yet joined the group confirmed by Ryan, the recommender system would directly estimate the overall rating that Frances would give to Ryan (i.e., $r_0^{\text{Frances,Ryan}}$) by estimating R.

In this example, without losing generality, the estimate of the overall rating that Frances would give to Ryan is based on the similarity between Frances and user u, denoted by sim(Frances, u), who rated Ryan; meanwhile, the similarity is calculated according to the initiators that Frances and user u have both rated previously. The more similar Frances and u are, the more $r_0^{u,\text{Ryan}}$ would contribute to $r_0^{\text{Frances, Ryan}}$.

The cosine-based similarity measure is most commonly used to derive sim_i(Frances, u) on criterion i. sim_i(Frances, u) is defined as

$$\begin{aligned}\text{sim}_i\,(\text{Frances}, u) = &\left(\sum_{j\in I(\text{Frances},u)} r_i^{\text{Frances},j} r_i^{u,j}\right)\\ &\times\left(\sqrt{\sum_{j\in I(\text{Frances},u)}\left(r_i^{\text{Frances},j}\right)^2}\right.\\ &\left.\times\sqrt{\sum_{j\in I(\text{Frances},u)}\left(r_i^{u,j}\right)^2}\right)^{-1},\\ &= 1,\ldots,n,\end{aligned} \tag{2}$$

where I(Frances, u) represents the sets of initiators rated by both Frances and user u. sim(Frances, u) is a used-based average obtained by aggregating the individual similarities in several ways as follows:

(1) average similarity:

$$\text{sim}\,(\text{Frances}, u) = \frac{1}{n+1}\sum_{i=0}^{n}\text{sim}_i\,(\text{Frances}, u)\,; \tag{3}$$

(2) worst-case similarity:

$$\text{sim}(\text{Frances}, u) = \min_{i=0,\ldots,n} \text{sim}_i(\text{Frances}, u). \quad (4)$$

In addition to the cosine-based similarity measure, a distance-based similarity can be formulated as follows:

$$\text{sim}(\text{Frances}, u) = 1 \times \left(1 + \frac{1}{|I(\text{Frances}, u)|} \times \sum_{j \in I(\text{Frances}, u)} d(R(\text{Frances}, j), R(u, j))\right)^{-1}, \quad (5)$$

where $d(R(\text{Frances}, j), R(u, j))$ can be derived by various distance metrics, for example,

(i) Manhattan distance:

$$d(R(\text{Frances}, j), R(u, j)) = \sum_{i=0}^{n+1} \left| r_i^{\text{Frances}, j} - r_i^{u, j} \right|; \quad (6)$$

(ii) Euclidean distance:

$$d(R(\text{Frances}, j), R(u, j)) = \sqrt{\sum_{i=0}^{n+1} \left(r_i^{\text{Frances}, j} - r_i^{u, j} \right)^2}; \quad (7)$$

(iii) Chebyshev distance:

$$d(R(\text{Frances}, j), R(u, j)) = \max_{i=0,\ldots,n} \left| r_i^{\text{Frances}, j} - r_i^{u, j} \right|. \quad (8)$$

The predicted overall rating $r_0^{\text{Frances, Ryan}}$ can then be defined by a weighted average of $r_0^{u, \text{Ryan}}$:

$$r_0^{\text{Frances, Ryan}} = \overline{e(\text{Frances})} + \frac{\sum_u \text{sim}(\text{Frances}, u) \left(r_0^{u, \text{Ryan}} - \overline{e(u)} \right)}{\sum_u \text{sim}(\text{Frances}, u)}, \quad (9)$$

where $\overline{e(u)}$ represents the average overall rating of user u. This formulation has been used by the well-known GroupLens [20] to provide personalized predictions for Usenet news [21].

As for the single-criterion rating systems, the rating function R for the (user, item) pairs is defined as follows:

$$R(\text{user}, \text{item}) \longrightarrow R_0, \quad (10)$$

where $\text{sim}(\text{Frances}, u)$ is simply specified as $\text{sim}_0(\text{Frances}, u)$ to obtain $r_0^{\text{Frances, Ryan}}$.

2.2. Aggregation-Function-Based Approach. In contrast to the similarity-based approach, the aggregation-function-based approach assumes that a certain relationship exists between the overall rating and the multicriteria ratings of items. Undoubtedly, the aggregation function plays an important role for an aggregation-function-based approach. The rating function R is defined as follows:

$$R(\text{user}, \text{item}) \longrightarrow R_i, \quad i = 1, \ldots, n. \quad (11)$$

Following the example in the previous subsection, instead of computing the individual similarity weights, it is necessary to estimate that ratings of the reputation (i.e., $r_1^{\text{Frances, Ryan}}$) and the response (i.e., $r_2^{\text{Frances, Ryan}}$) that Frances would give to Ryan can be estimated by a user-based deviation-from-mean method as follows:

$$r_i^{\text{Frances, Ryan}} = \overline{e_i(\text{Frances})} + \frac{\sum_u \text{sim}_i(\text{Frances}, u) \left(r_i^{u, \text{Ryan}} - \overline{e_i(u)} \right)}{\sum_u \text{sim}_i(\text{Frances}, u)}, \quad i = 1, 2, \quad (12)$$

where $\overline{e_i(u)}$ represents the average overall rating of user u for criterion i. $r_i^{\text{Frances, Ryan}}$ can be obtained by considering the cosine-based similarity measure. Then $r_0^{\text{Frances, Ryan}}$ is further predicted by aggregating $r_1^{\text{Frances, Ryan}}$ and $r_2^{\text{Frances, Ryan}}$. Therefore, the focus of the similarity-based and the aggregation-function-based approaches is quite different. The WAM is often used to aggregate the partial ratings (i.e., $r_1^{\text{Frances, Ryan}}$ and $r_2^{\text{Frances, Ryan}}$) when w_1 and w_2 have been assigned:

$$r_0^{\text{Frances, Ryan}} = w_1 r_1^{\text{Frances, Ryan}} + w_2 r_2^{\text{Frances, Ryan}}, \quad (13)$$

where $w_1 + w_2 = 1$. This means that a classical set function μ can be defined on $\{r_1^{\text{Frances, Ryan}}, r_2^{\text{Frances, Ryan}}\}$ with $\mu(r_1^{\text{Frances, Ryan}}) = w_1$ and $\mu(r_2^{\text{Frances, Ryan}}) = w_2$ such that $\mu(\{r_1^{\text{Frances, Ryan}}, r_2^{\text{Frances, Ryan}}\}) = \mu(r_1^{\text{Frances, Ryan}}) + \mu(r_2^{\text{Frances, Ryan}})$. The additivity of μ indicates that there exists no interactions among $r_1^{\text{Frances, Ryan}}$ and $r_2^{\text{Frances, Ryan}}$. Unfortunately, as mentioned above, this assumption is not warranted in many applications [4]. Because the fuzzy integral does not assume the independence of elements [17, 18], it is reasonable to obtain $r_0^{\text{Frances, Ryan}}$ by using a nonadditive Choquet integral to aggregate $r_1^{\text{Frances, Ryan}}$ and $r_2^{\text{Frances, Ryan}}$.

3. Non-Additive Aggregation-Function-Based Collaborative-Filtering Approach

In this section, the fuzzy measure used for describing the interaction among attributes in a set is first described in Section 3.1. Section 3.2 presents the proposed approach using the Choquet integral and an accuracy metric for recommendation performance evaluation.

3.1. Description of the Interaction Using a Fuzzy Measure. Let $P(X)$ denote the power set of $X = \{x_1, x_2, \ldots, x_n\}$, where X is called the feature space. Then $(X, P(X))$ is a measurable space. A nonadditive and nonnegative set function $\mu : P(X) \to [0, 1]$ is a fuzzy measure that satisfies the following conditions [22–24]:

(1) $\mu(\emptyset) = 0$;

(2) for all $R, S \in P(X)$, if $R \subset S$, then $\mu(R) \leq \mu(S)$ (monotonicity);

(3) for every sequence of subsets of X, if either $R_1 \subseteq R_2 \subseteq \cdots$ or $R_1 \supseteq R_2 \supseteq \cdots$, then $\lim_{i\to\infty}\mu(R_i) = \mu(\lim_{i\to\infty}R_i)$ (continuity).

When $\mu(X) = 1$, μ is said to be regular. The fuzzy measure is developed to consider the interaction among attributes towards the objective attribute [17] by replacing the usual additive property with the monotonic property.

Let μ_k denote $\mu(\{x_k\})$, which is called a fuzzy density, and $E_k = \{x_k, x_{k+1}, \ldots, x_n\}$ $(1 \leq k \leq n)$, where $0 \leq \mu_k \leq 1$. Interaction among the attributes of E_k can be described using $\mu(E_k)$, which expresses the relative importance or discriminatory power of E_k. This means that μ can be regarded as an importance measure and then μ_k can be interpreted as the degree of importance of x_k. $\mu(E_k)$ may be less than or greater than $\mu_k + \mu_{k+1} + \cdots + \mu_n$, thereby expressing an interaction among the elements $x_k, x_{k+1}, \ldots, x_n$ [12]. For instance, if $\mu_1 = 0.3$, $\mu_2 = 0.5$, and $\mu(\{x_1, x_2\}) = 0.9$, then $\mu(\{x_1, x_2\}) > \mu_1 + \mu_2$ indicates that the joint contribution of x_1 and x_2 to the decision or the objective attribute is greater than the sum of their individual contributions. This indicates that they would enhance each other.

Among the various options for μ, a λ-fuzzy measure is a convenient means of computing the fuzzy integral [12, 23, 25]. For all $R, S \in P(X)$ with $R \cap S = \emptyset$, μ is a λ-fuzzy measure if it satisfies the following property:

(1)

$$\mu(\emptyset) = 0, \qquad \mu(X) = 1; \tag{14}$$

(2)

$$\mu(R \cup S) = \mu(R) + \mu(S) + \lambda\mu(R)\mu(S), \quad \lambda \in (-1, \infty). \tag{15}$$

The advantage of using the λ-fuzzy measure is that, after determining the fuzzy densities $\mu_1, \mu_2, \ldots, \mu_n$, λ can be uniquely determined from the condition $\mu(X) = 1$. $\mu(E_k)$ can be further determined by λ and μ_j as follows:

$$\mu(Ek) = \frac{1}{\lambda}\left[\prod_{i=k\cdots n}(1 + \lambda\mu_i) - 1\right]. \tag{16}$$

As mentioned above, $\mu(R \cup S)$ expresses the importance of $R \cup S$. The value of λ determines the nature of the interaction between R and S. If $\lambda > 0$, there is a multiplicative effect between R and S (i.e., R and S are superadditive); if $\lambda < 0$ there is a substitutive effect between R and S (i.e., R and S are subadditive). If $\lambda = 0$, then R and S are not interactive: $\mu(R \cup S) = \mu(R) + \mu(S)$. Actually, the sign of λ can be identified by $\sum_{i=1}^{n}\mu_i$. In other words, if $\sum_{i=1}^{n}\mu_i > 1$, then $-1 < \lambda < 0$; if $\sum_{i=1}^{n}\mu_i < 1$, then $\lambda > 0$; and if $\sum_{i=1}^{n}\mu_i = 1$, then $\lambda = 0$.

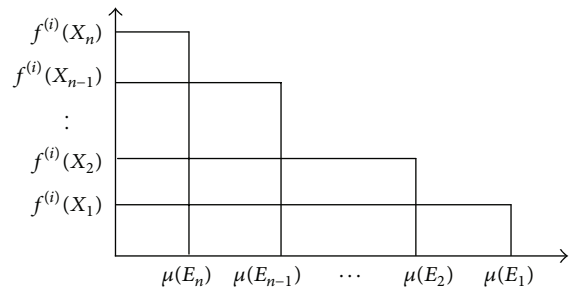

Figure 1: Graphical representation of Choquet fuzzy integral.

3.2. The Proposed Non-Additive Approach

3.2.1. Aggregating Multicriteria Ratings Using the Choquet Integral. To consider interactions among criteria, a nonadditive collaborative-filtering approach is proposed to estimate, for instance, $r_0^{u,v}$, using the Choquet integral, where v is an initiator but has not been rated by u. The Choquet integral, which is a generalization of the linear Lebesgue integral (e.g., the weighted average method) [17], can be represented in terms of fuzzy measures [12, 18]. Let $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{in}) = (r_1^{u,v}, r_2^{u,v}, \ldots, r_n^{u,v})$. For the proposed nonadditive aggregation-function-based approach, the synthetic evaluation of $\mathbf{x}_i$ can be further obtained by the Choquet integral. Let $f^{(i)}$ with respect to $\mathbf{x}_i$ be a non-negative, real-valued measurable function defined on X such that $f^{(i)}(x_k) = x_{ik}$ $(k = 1, 2, \ldots, n)$ falls into a certain range. The element in X with $\min\{f^{(i)}(x_k) \mid k = 1, 2, \ldots, n\}$ is renumbered as one, where $f^{(i)}(x_k)$ denotes the performance or observation value of x_k with respect to $\mathbf{x}_i$. In other words, all elements x_k are rearranged in order of descending $f^{(i)}(x_k)$, so that $f^{(i)}(x_1) \leq f^{(i)}(x_2) \leq \cdots \leq f^{(i)}(x_n)$. As illustrated in Figure 1, the Choquet integral $(c)\int f^{(i)}\mathrm{d}\mu$ over X of f with respect to μ is defined as follows:

$$(c)\int f^{(i)}\mathrm{d}\mu = \sum_{k=1}^{n} f^{(i)}(x_k)\left(\mu(E_k) - \mu(E_{k+1})\right), \tag{17}$$

where $\mu(E_{n+1})$ is specified as zero and $R_0(u, v) = (c)\int f^{(i)}\mathrm{d}\mu$. If $\sum_j \mu_j$ is equal to one, the Choquet integral coincides with the WAM.

3.2.2. Evaluating Recommendation Performance. The performance of a recommendation approach can be evaluated by decision-support accuracy metrics which determine how well the recommendation approach can predict items the user would rate highly. Commonly used metrics are precision, recall, and the F-measure. Precision is the number of truly high overall ratings expressed as a fraction of the total number of ratings that the system predicted they would be high; recall is the number of correctly predicted high ratings expressed as a fraction of all the ratings known to be high, while the F-measure is the harmonic mean of precision and recall. Often, there is an inverse relationship between precision and recall because it is possible to increase one at

the cost of reducing the other. Usually, precision and recall scores are not discussed in isolation. Therefore, it pays to take into account the F-measure [1] as follows:

$$F = \frac{2\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}, \tag{18}$$

where recall and precision are evenly weighted. However, the actual weights on precision and recall should be dependent on the goal of a recommendation approach. van Rijsbergen [26] further proposed the F_β measure as follows:

$$F_\beta = \left(1 + \beta^2\right) \frac{\text{precision} \cdot \text{recall}}{\beta^2 \text{precision} + \text{recall}}, \tag{19}$$

where $\beta \geq 0$. F_β weights recall higher than precision when $\beta > 1$ and weights precision higher than recall when $\beta < 1$. Clearly, the original F-measure is simply the F_1 measure.

From the viewpoint of practicality, many users of recommendation applications are typically interested in only the few highest-ranked item recommendations [5]. Therefore, the popular metric related to precision, namely, precision-in-top-N ($N = 1, 2, 3, \ldots$), should be taken into account for the proposed method. This metric is defined as the fraction of truly high overall ratings among those the system predicted would be the N relevant items for each user. Furthermore, it is reasonable to place more emphasis on precision than on recall, to highlight the significance of precision for users. For this, a variation of the F_β measure that places emphasis on precision-in-top-N is presented to evaluate the performance of the proposed nonadditive recommendation approach:

$$vF_\beta = \left(1 + \beta^2\right) \frac{\text{precision-in-top-}N \cdot \text{recall}}{\beta^2 \text{precision-in-top-}N + \text{recall}}, \tag{20}$$

where $0 \leq \beta \leq 1$.

4. A GA-Based Method for Constructing a Non-Additive Recommendation Model

4.1. Constructing a Non-Additive Recommendation Model Using Multicriteria Ratings. The proposed model does not involve any complicated mechanisms for selecting the free parameters. Because decision-makers cannot easily pre-specify the criterion weights, a real-valued GA-based method is used here, involving the basic operations of selection, crossover, and mutation [27, 28] to determine the optimal values of the criterion weights (i.e., $\mu_1, \mu_2, \ldots, \mu_n$). Let n_{size} and $n_{\max}$ denote the population size and the total number of generations, respectively. The following steps are used to construct a recommendation model using the proposed nonadditive collaborative-filtering approach.

Algorithm 1. Construct a nonadditive recommendation model using the collaborative-filtering approach.

Input. Population size (N_{pop}); stopping condition (N_{con}, i.e., total number of generations); number of elite chromosomes (N_{del}); crossover probability (Pr_c); mutation probability (Pr_m); the value of β for vF_β measure; the value of N for the precision-in-top-N measure; a set of training patterns.

Output. A nonadditive recommendation model using the collaborative-filtering approach with a higher vF_β measure.

Method

Step 1. Population Initialization. Generate an initial population of n_{size} chromosomes, each consisting of n real-valued parameters. Randomly assign a real value chosen from the interval $[0, 1]$ to each parameter in a chromosome.

Step 2. Chromosome Evaluation. Compute the fitness value for each chromosome. Because the objective of the algorithm is to construct a nonadditive recommendation model using the collaborative-filtering approach with a higher vF_β measure, the vF_β measure is used as the fitness function for evaluating a chromosome.

Step 3. Generation of New Chromosomes. Generate a new generation of n_{size} chromosomes by selection, crossover, and mutation.

Step 4. Elitist Strategy. Randomly remove n_{del} ($0 \leq n_{\text{del}} \leq n_{\text{size}}$) of the newly generated chromosomes. Insert n_{del} copies of the chromosome with maximum fitness in the previous generation.

Step 5. Termination Test. Terminate the algorithm if $n_{\max}$ generations have been generated; otherwise, return to Step 2.

When the stopping condition is satisfied, the algorithm is terminated and whichever chromosome has the maximum fitness among all generations serves as the desired solution. It is noted that the above algorithm can also be applied to construct an additive recommendation model using the WAM.

4.2. Genetic Operations. Let P_k denote the population generated in generation k ($1 \leq k \leq n_{\max}$). Chromosome i ($1 \leq k \leq n_{\text{size}}$) generated in P_k is represented by $\mu_{i1}^k \mu_{i2}^k \cdots \mu_{in}^k$. After evaluating the fitness value for each chromosome in P_k, selection, crossover, and mutation are applied until n_{size} new chromosomes have been generated for P_{k+1}. These genetic operations are described in more detail below. In contrast to a nonadditive recommendation model, for an additive recommendation model using the WAM, μ_{ij}^k can be specifically set as follows before evaluating the fitness value:

$$\mu_{ij}^k = \frac{\mu_{ij}^k}{\zeta}, \tag{21}$$

where $\zeta = \mu_{i1}^k + \mu_{i2}^k + \cdots + \mu_{in}^k$.

4.2.1. Selection. Using the binary tournament selection, two chromosomes from the current population are randomly selected, and the one with the higher fitness is placed in a mating pool. This process is repeated until there are n_{size} chromosomes in the mating pool. Next, n_{size} pairs of chromosomes from the pool are randomly selected for mating.

The crossover and mutation operations are applied to the parents to generate children.

4.2.2. Crossover. For each mated pair of chromosomes i and j, $\mu_{i1}^k \mu_{i2}^k \cdots \mu_{in}^k$ and $\mu_{j1}^k \mu_{j2}^k \cdots \mu_{jn}^k$, each pair of genes has a probability Pr_c of undergoing the crossover operation. The operations are performed as $\mu_{iw}^{k\,\prime} = a_w \mu_{iw}^k + (1 - a_w)\mu_{jw}^k$, $\mu_{jw}^{k\,\prime} = (1 - a_w)\mu_{iw}^k + a_w \mu_{jw}^k$ $(1 \le w \le n)$, where a_w is a random number in the interval $[0, 1]$. Two new chromosomes are thereby generated, which will replace their parents in generation P_{k+1}.

4.2.3. Mutation. There is a probability Pr_m that the mutation operation will be performed on each real-valued parameter in new chromosomes generated by the crossover operation. To avoid excessive perturbation of the gene pool, a low mutation rate is used. If a mutation occurs for a gene, it will be changed by adding a number randomly selected from a specified interval. After crossover and mutation, n_{del} $(0 \le n_{\mathrm{del}} \le n_{\mathrm{size}})$ chromosomes in P_{k+1} are randomly removed from the set of new chromosomes (those formed by genetic operations) to make room for additional copies of the chromosome with maximum fitness value in P_k.

5. Application to Initiator Recommendation on a Group-Buying Website

Group-buying websites play the role of a transaction platform between businesses and consumers. The websites call a group of consumers with the same needs for the purchase of items and then negotiate with vendors to obtain the best price or to get a special discount. Groupon, which has the high market share, is a representative group-buying website. With group-buying activities increasing and their associated websites expanding rapidly, a market research institute in Virginia, BIA/Kelsey, predicted that the group-buying market in the United States will reach US$ 39.3 billion in 2015 [29]. Undoubtedly, group-buying has become an important transaction model for online shopping.

In Taiwan, the Institute for Information Industry (MIC) of Taiwan reported that the group-buying market reached approximately US$ 2.39 billion in 2010 and is expected to reach US$ 3 billion by 2011 [30]. In the application, the proposed recommendation approach has been applied to initiator recommendation on one popular group-buying website in Taiwan. Its sales volume was over US$ 0.17 billion in 2010. On this website, users often search for appropriate initiators who can bargain with vendors over the price for certain items. However, due to the large number of search results, users have suffered from the problem of information overload, especially for hot items. Furthermore, application domains in previous research do not involve the initiator recommendations for the group-buying. This motivates us to apply the recommendation techniques to this website.

The computer simulations were programmed in Delphi 7.0 and executed on a 2 GHz dual-CPU Pentium computer. Section 5.1 describes the data collected. Section 5.2 describes the parameter specification of the GA for the computer simulations. In Section 5.3, the performance of the proposed nonadditive recommendation approach is compared with several recommendation approaches using multicriteria ratings.

5.1. Data Description. A total of 211 undergraduates with business administration as their major subject and who were familiar with group-buying participated in the experiment. Each subject was asked to rate twenty experienced initiators selected from the above-mentioned website. Each subject assigned an overall rating and five criterion ratings, namely, ability, reputation, responsiveness, trust, and interaction, to each initiator. These criteria are described below.

Ability. This represents knowledge and techniques that can be used to solve problems for customers [31]. Initiators are expected not only to have the experience in confirming groups, but also to solve any problems that arise in group-buying.

Reputation. This indicates whether the initiators have the ability and intention to fulfill their promises [32].

Responsiveness. In a virtual community, members always expect to receive responses to their posted information from other members [33]. In the group-buying context, this can promote the establishment of trust among initiators and group members. Responsiveness indicates whether the initiators tried to respond to any problems reported by the group members.

Trust. Based on the viewpoint of trust in [34], it is considered that the initiators are expected to have a positive expectation of the intention.

Interaction. Regular communication between sellers and buyers is helpful in building up the trust of buyers in sellers [35]. Therefore, initiators are expected to discuss progress and problems actively with group members during a group-buying session.

All ratings range from zero, representing "very unsatisfactory," to ten, representing "very satisfactory." The overall rating indicates how much a user appreciates the initiator. Because the decision-support measures are used to estimate accuracy, it is necessary to define every overall rating on a binary scale [21] as "high-ranked" or "not high-ranked". It should be reasonable to translate these overall ratings into a binary scale by treating the ratings greater than or equal to seven as high-ranked and those less than seven as not high-ranked. In other words, the initiators whose ratings are greater than or equal to seven are relevant to the users.

The leave-ten-out technique is used to examine the generalization ability of different recommendation. In each of the iterations, ten evaluations given by a subject are randomly selected to serve as test data, and the remaining evaluations were chosen as the training data. This means that test data are produced for each subject in each of the iterations. Then the overall rating was predicted for each evaluation in the test data based on the information in the training data. The precision-in-top-N for the test data can also be obtained.

Table 1: vF_β of different methods.

Classification method	β							
	1	1/2	1/3	1/4	1/5	1/6	1/7	1/8
Single-criterion rating	0.545	0.676	0.735	0.762	0.776	0.784	0.789	0.793
Average similarity	0.546	0.681	0.742	0.770	0.785	0.794	0.799	0.803
Worst-case similarity	0.536	0.671	0.733	0.762	0.777	0.785	0.791	0.795
Manhattan distance	0.563	0.707	0.773	0.803	0.820	0.829	0.835	0.839
Euclidean distance	0.553	0.691	0.753	0.782	0.797	0.806	0.812	0.815
Chebyshev distance	0.553	0.688	0.749	0.777	0.792	0.801	0.806	0.810
WAM	0.569	0.682	0.735	0.756	0.769	0.793	0.789	0.792
Non-additive approach	0.657	0.719	0.751	0.779	0.797	0.822	0.837	0.851

This procedure is iterated until the evaluations of each of the subjects have been used as the test data. Because the results of a random sampling procedure may be dependent on the selection of evaluations, the above procedure is repeated five times.

5.2. GA Parameter Specifications. A number of factors can influence GA performance, including the size of the population and the probabilities of applying the crossover and mutation operators. Unfortunately, there is no standard procedure for choosing optimal GA specifications. Based on the principles suggested by Osyczka [36] and Ishibuchi et al. [37], the parameters are specified as follows:

(i) $n_{\text{size}} = 50$: GA populations commonly range from 50 to 500 individuals. Hence, 50 individuals is an acceptable size;

(ii) $n_{\max} = 500$: the stopping condition is specified according to the available computing time;

(iii) $n_{\text{del}} = 2$: only a small number of elite chromosomes are used;

(iv) $\text{Pr}_c = 1.0$, $\text{Pr}_m = 0.01$: since Pr_c controls the range of exploration in the solution space, most sources recommend choosing a large value. To avoid generating an excessive perturbation Pr_m should be set to a small value;

(v) $N = 5$: it is assumed that users required the system to recommend the five most highly ranked initiators. In other words, the precision-in-top-5 is incorporated into the vF_β measure;

(vi) β is specified as $1, 1/2, 1/3, \ldots, 1/8$. A recommender system is constructed for each β value.

Although the above specifications are somewhat subjective, the experimental results show that they are acceptable. Therefore, customized parameter tuning is not considered for the proposed approach.

5.3. Performance Evaluation. The proposed nonadditive approach is compared with several collaborative-filtering approaches introduced in the previous section: the traditional single-criterion rating approach, similarity-based approaches using multicriteria ratings with average similarity, worst-case similarity, distance-based similarity as described in [5], and an aggregation-function-based approach using the WAM. By a random sampling procedure with 50% training data and 50% test data, the average results of these approaches are obtained from five trials. For each evaluation in the test data, the evaluated initiator is treated as the target item. Each approach considered is used to predict the overall rating that the target item would have by using the training data. Then, "high-ranked" or "not high-ranked" for each target item could be determined by its predicted overall rating.

By maximizing vF_β with the precision-in-top-5 measure for a certain β, it is interesting to examine the precision-in-top-1 and precision-in-top-3 for the proposed approach. The idea is to identify β for which the proposed method can perform well compared with the similarity-based methods considered for vF_β and various precision-in-top-N measures ($N = 1, 3, 5$). Table 1 shows that the vF_β value of all the methods improved from $\beta = 1$ to $\beta = 1/8$. This is reasonable because vF_β approaches precision-in-top-N when β is sufficiently small. The notable results in Tables 1, 2, and 3 can be summarized as follows.

(1) An approach with a higher precision-in-top-5 measure is not guaranteed to have a higher vF_β. For instance, for $\beta = 1$ and $1/2$; although the vF_β value of the proposed nonadditive approach exceeds the similarity-based approaches considered, it can be seen in Table 2 that the precision-in-top-5 measures obtained by the proposed approach are inferior to those obtained by the similarity-based approaches and the WAM.

(2) When $\beta \leq 1/5$, the proposed nonadditive approach performs well compared with the similarity-based methods using the average similarity and the worst-case similarity for vF_β and different precision-in-top-N measures ($N = 1, 3, 5$).

(3) When $\beta \leq 1/7$, the proposed nonadditive approach performs well compared with similarity-based methods using distance-based similarities for vF_β, the precision-in-top-3 measure, and the precision-in-top-5 measure.

(4) The proposed nonadditive approach is inferior to the approaches using distance-based similarities for

Table 2: Various precisions (%) of different methods.

Method	Precision-in-top-1	Precision-in-top-3	Precision-in-top-5
Single-criterion rating	82.02	82.90	80.45
Average similarity	83.02	83.35	81.48
Worst-case similarity	83.52	83.57	80.81
Manhattan distance	93.71	88.55	85.22
Euclidean distance	91.51	86.44	82.77
Chebyshev distance	89.28	85.36	82.17

the precision-in-top-1 measure (i.e., Manhattan and Euclidean distances) regardless of the value of β.

(5) When $\beta \leq 1/5$, the proposed nonadditive approach performs well compared with the single-criterion rating approach for vF_β and various precision-in-top-N measures ($N = 1, 3, 5$).

(6) The proposed nonadditive approach outperforms the WAM for vF_β and various precision-in-top-N measures ($N = 1, 3, 5$) when $\beta \leq 1/4$.

(7) Among the similarity-based approaches, the single-criterion rating approach seems to perform worst for vF_β and various precision-in-top-N measures ($N = 1, 3, 5$) regardless of the value of β.

(8) The WAM outperforms the single-criterion rating approach for the precision-in-top-5 measure when $\beta \leq 1/3$. The precision-in-top-1 and precision-in-top-3 measures of the WAM are inferior to those of the single-criterion rating approach.

Therefore, by incorporating the precision-in-top-5 measure into the fitness function (i.e., vF_β) with appropriate β values, the proposed nonadditive approach is found to outperform the traditional single-criterion rating approach and the similarity-based approaches using multicriteria ratings for vF_β, the precision-in-top-3 and the precision-in-top-5 measures. Moreover, the proposed nonadditive approach outperforms the additive WAM when using appropriate β values for vF_β and different precision-in-top-N measures. The experimental results indicate that leveraging the multicriteria ratings of initiators could be helpful in improving the prediction performance of the traditional single-criterion rating approach. In addition, it is found that λ is less than -0.9 for the best solution. In other words, there exists a substitutive interaction effect among the attributes. Therefore, it seems quite reasonable to use the fuzzy integral with a λ-fuzzy measure as an aggregation function instead of the WAM.

6. Discussion and Conclusions

It is known that the assumption of independence among criteria in the WAM is not always reasonable. The main issue that this paper addresses is the additivity of the WAM for the multicriteria aggregation-function-based approach. In view of the nonadditivity of the fuzzy integral, this paper contributes to present a nonadditive aggregation-function-based approach using multicriteria ratings by combining the similarity-based approach using single-criterion ratings with the Choquet fuzzy integral. The former is used to predict multicriteria ratings and the latter to generate an overall rating for an item. Because many users of recommendation applications are typically interested in only the few highest-ranked item recommendations, a variant of the F_β measure has been designed by incorporating the precision-in-top-N metric into the F_β measure to assess prediction performance. Both the proposed nonadditive approach and the similarity-based approaches with average/worst-case similarity use the cosine-based similarity measure to obtain the similarity on each criterion between two users. Whereas the former (i.e., the proposed nonadditive approach) derives the rating for each criterion from individual similarities and then estimates the overall rating for an item, the latter (i.e., the similarity-based approaches with average/worst-case similarity) aggregates the individual similarities to compute the overall similarity between two users.

The proposed nonadditive approach is applied to initiator recommendation on a group-buying website in Taiwan in order to examine its prediction performance. The criterion weights for the fuzzy integral are determined automatically by a GA. Compared with the traditional single-criterion rating approach, several collaborative-filtering approaches using multicriteria ratings, and the aggregation-function-based approach using the WAM, it can be seen that the proposed nonadditive collaborative-filtering approach yields encouraging vF_β, precision-in-top-3, and precision-in-top-5 measures with appropriate β values. Therefore, the proposed approach is able to improve recommendation performance by leveraging multicriteria information. This indicates that the proposed nonadditive approach has applicability to other application domains. Note that the choice of β is eventually dependent on proprietors of websites. Besides, identification of the best collaborative-filtering approach is not possible because there is no such thing as the "best" approach [25].

It is notable that, when a traditional single-criterion rating approach is treated as a baseline collaborative-filtering approach, the experimental results indicate that the similarity-based approaches using multicriteria ratings of initiators perform better than the traditional single-criterion rating approach. Moreover, the proposed nonadditive approach and the WAM outperform the traditional single-criterion rating approach with appropriate β values. However, it is not possible to conclude that recommendation approaches using multicriteria ratings outperform the traditional single-criterion rating approach in all domains where multicriteria information exists.

Table 3: Various precisions (%) of WAM and non-additive approach.

Method	Precision	β							
		1	1/2	1/3	1/4	1/5	1/6	1/7	1/8
WAM	Precision-in-top-1	77.01	78.62	79.81	80.72	81.03	81.04	81.66	81.98
	Precision-in-top-3	80.09	81.15	81.61	80.96	80.29	81.64	81.37	81.36
	Precision-in-top-5	78.37	80.43	81.34	81.25	81.19	82.20	82.46	82.07
Non-additive approach	Precision-in-top-1	76.16	78.84	80.21	83.91	86.16	92.45	90.15	88.44
	Precision-in-top-3	76.03	78.19	80.33	82.89	85.32	88.68	88.67	88.58
	Precision-in-top-5	74.88	77.68	79.42	83.08	85.50	88.67	88.63	89.14

As mentioned above, the precision-in-top-1 and precision-in-top-3 measures for the proposed nonadditive approach are generated by a system that is trained by incorporating the precision-in-top-5 measure into the fitness function, whereas the precision-in-top-1 measures obtained by the proposed nonadditive approach are inferior to those obtained by approaches using distance-based similarities. It would be interesting to examine whether the precision-in-top-1 measure obtained by the proposed approach can be improved when the system is trained by incorporating the precision-in-top-1 measure into the fitness function.

Acknowledgments

The author would like to thank the anonymous referees for their valuable comments. This research is partially supported by the National Science Council of Taiwan under grant NSC 102-2410-H-033-039-MY2.

References

[1] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich, *Recommender Systems: An Introduction*, Cambridge University Press, New York, NY, USA, 2010.

[2] M. Gao, Z. Wu, and F. Jiang, "Userrank for item-based collaborative filtering recommendation," *Information Processing Letters*, vol. 111, no. 9, pp. 440–446, 2011.

[3] S.-L. Huang, "Designing utility-based recommender systems for e-commerce: evaluation of preference-elicitation methods," *Electronic Commerce Research and Applications*, vol. 10, no. 4, pp. 398–407, 2011.

[4] S. Opricovic and G.-H. Tzeng, "Compromise solution by MCDM methods: a comparative analysis of VIKOR and TOPSIS," *European Journal of Operational Research*, vol. 156, no. 2, pp. 445–455, 2004.

[5] G. Adomavicius and Y. Kwon, "New recommendation techniques for multicriteria rating systems," *IEEE Intelligent Systems*, vol. 22, no. 3, pp. 48–55, 2007.

[6] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734–749, 2005.

[7] J. B. Schafer, "DynamicLens: a dynamic user-interface for a meta-recommendation system," in *Proceedings of Beyond Personalization 2005: A Workshop on the Next Stage of Recommender Systems Research*, pp. 72–76.

[8] W.-P. Lee, C.-H. Liu, and C.-C. Lu, "Intelligent agent-based systems for personalized recommendations in Internet commerce," *Expert Systems with Applications*, vol. 22, no. 4, pp. 275–284, 2002.

[9] F. Ricci and H. Werthner, "Case-based querying for travel planning recommendation," *Information Technology and Tourism*, vol. 4, no. 3-4, pp. 215–226, 2002.

[10] N. Manouselis and C. Costopoulou, "Experimental analysis of design choices in multi-attribute utility collaborative filtering," in *Personalization Techniques and Recommender*, G. Uchyigit and M. Y. Ma, Eds., pp. 111–134, World Scientific, River Edge, NJ, USA, 2008.

[11] T. Onisawa, M. Sugeno, Y. Nishiwaki, H. Kawai, and Y. Harima, "Fuzzy measure analysis of public attitude towards the use of nuclear energy," *Fuzzy Sets and Systems*, vol. 20, no. 3, pp. 259–289, 1986.

[12] W. Wang, "Genetic algorithms for determining fuzzy measures from data," *Journal of Intelligent and Fuzzy Systems*, vol. 6, no. 2, pp. 171–183, 1998.

[13] T. Murofushi and M. Sugeno, "An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure," *Fuzzy Sets and Systems*, vol. 29, no. 2, pp. 201–227, 1989.

[14] T. Murofushi and M. Sugeno, "A theory of fuzzy measures: representations, the Choquet integral, and null sets," *Journal of Mathematical Analysis and Applications*, vol. 159, no. 2, pp. 532–549, 1991.

[15] T. Murofushi and M. Sugeno, "Some quantities represented by the Choquet integral," *Fuzzy Sets and Systems*, vol. 56, no. 2, pp. 229–235, 1993.

[16] M. Sugeno, Y. Narukawa, and T. Murofushi, "Choquet integral and fuzzy measures on locally compact space," *Fuzzy Sets and Systems*, vol. 99, no. 2, pp. 205–211, 1998.

[17] Z. Wang, K.-S. Leung, and G. J. Klir, "Applying fuzzy measures and nonlinear integrals in data mining," *Fuzzy Sets and Systems*, vol. 156, no. 3, pp. 371–380, 2005.

[18] Z. Wang, K.-S. Leung, and J. Wang, "A genetic algorithm for determining nonadditive set functions in information fusion," *Fuzzy Sets and Systems*, vol. 102, no. 3, pp. 463–469, 1999.

[19] C. Desrosiers and G. Karypis, "A comprehensive survey of neighborhood-based recommendation methods," in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds., pp. 107–144, Springer, New York, NY, USA, 2011.

[20] J. A. Konstan, B. N. Miller, D. Maltz, J. L. Herlocker, L. R. Gordon, and J. Riedl, "Applying collaborative filtering to usenet news," *Communications of the ACM*, vol. 40, no. 3, pp. 77–87, 1997.

[21] J. Herlocker, J. Konstan, A. Borchers, and J. Riedl, "An algorithmic framework for performing collaborative filtering," in *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 230–237, 1999.

[22] M. Sugeno, *Theory of fuzzy integrals and its applications [Ph.D. thesis]*, Tokyo Institute of Technology, Tokyo, Japan, 1974.

[23] M. Sugeno, "Fuzzy measures and fuzzy integrals—a survey," in *Fuzzy Automata and Decision Processes*, pp. 89–102, North-Holland, New York, NY, USA, 1977.

[24] Z. Y. Wang and G. J. Klir, *Fuzzy Measure Theory*, Plenum Press, New York, NY, USA, 1992.

[25] L. I. Kuncheva, *Fuzzy Classifier Design*, Physica, Heidelberg, Germany, 2000.

[26] C. J. van Rijsbergen, *Information Retrieval*, Butterworth, London, UK, 1979.

[27] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, Mass, USA, 1989.

[28] K. F. Man, K. S. Tang, and S. Kwong, *Genetic Algorithms: Concepts and Designs*, Springer, London, UK, 1999.

[29] B. I. A. Kelsey, 2011, http://www.biakelsey.com.

[30] "Foreseeing Innovative New Digiservices in Institute for Information Industry," 2011, http://www.find.org.tw/find/home.aspx?page=many&id=290.

[31] J. Singh and D. Sirdeshmukh, "Agency and trust mechanisms in consumer satisfaction and loyalty judgments," *Journal of the Academy of Marketing Science*, vol. 28, no. 1, pp. 150–167, 2000.

[32] S. Ganesan, "Determinants of long-term orientation in buyer-seller relationships," *Journal of Marketing*, vol. 58, pp. 1–19, 1994.

[33] C. M. Ridings, D. Gefen, and B. Arinze, "Some antecedents and effects of trust in virtual communities," *Journal of Strategic Information Systems*, vol. 11, no. 3-4, pp. 271–295, 2002.

[34] D. M. Rousseau, S. Sitkin, R. S. Burt, and C. Camerer, "Not so different after all: a cross-discipline view of trust," *Academy of Management Review*, vol. 23, pp. 393–404, 1998.

[35] P. M. Doney and J. P. Cannon, "An examination of the nature of trust in buyer-seller relationships," *Journal of Marketing*, vol. 61, no. 2, pp. 35–51, 1997.

[36] A. Osyczka, *Evolutionary Algorithms for Single and Multicriteria Design Optimization*, Physica, New York, NY, USA, 2002.

[37] H. Ishibuchi, T. Nakashima, and M. Nii, *Classification and Modeling with Linguistic Information Granules: Advanced Approaches to Linguistic Data Mining*, Springer, 2004.