A Web-based personalized recommendation system for mobile phone selection: Design, implementation, and evaluation

Deng-Neng Chen [a], Paul Jen-Hwa Hu [b], Ya-Ru Kuo [c], Ting-Peng Liang [c,d,*]

[a] Department of Management Information Systems, National Pingtung University of Science and Technology, Taiwan, ROC
[b] Accounting and Information Systems, David Eccles School of Business, University of Utah, Taiwan, ROC
[c] Department of Information Management, National Sun Yat-sen University, Taiwan, ROC
[d] Department of Information Systems, City University of Hong Kong, Hong Kong

ARTICLE INFO

ABSTRACT

Recommendation systems that provide appropriate solutions to users to reduce their decision complexity have become popular in the Internet world. Designing and evaluating such systems remain essential challenges to researchers and practitioners. Toward that end, a critical task is how to obtain user preferences. Mobile phones have become indispensable in everyday life, yet fierce market competition, characterized by rapid introductions of different models with novel designs and advanced features, have made consumers' purchase decision making increasingly complex. As a well-established, multiple criteria decision technique, analytic hierarchy processing (AHP) provides an intuitive model of a hierarchical structure capable of supporting complex product comparisons and evaluations by consumers. In this paper, we illustrate the application of an AHP-based mechanism to develop a Web-based recommendation system and empirically evaluate the prototype by conducting a controlled experiment with 244 mobile phone users, focusing on both content and system satisfaction. Our evaluation includes benchmark systems built on rank-based analysis and an equal weight-based system as comparative baselines. Overall, the results suggest the viability and value of using AHP to construct effective recommendation systems. Subjects appear satisfied with the recommendations by the AHP-based system, though its relatively demanding input requirements may need mitigation and adequate interface designs. This study contributes to research and practice in recommender systems in general and helps develop mobile phone recommendation systems for online stores and consumers in particular.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

With the growing penetration of the Internet and e-commerce, personalized recommendations that identify appropriate products or services for customers to reduce their information load and search costs become increasingly critical. Many online vendors, including Amazon and Netflix, have implemented recommendation systems to assist their consumers. However, central to the development of an effective recommendation system is identifying customer preferences, which can be analyzed by previous shopping history or eliciting customers' purchasing criteria. For example, a system might use previous shopping history to suggest complementary products (e.g., recommending cereal to customers who buy milk). Yet when the consumer already has decided to buy a product, and just needs to select the brand or model, such simple data mining approaches are no longer adequate. Rather, recommendations need to rely on product attributes and user decision criteria.

For example, multiple criteria decision methods (MCDM), such as analytic hierarchy process (AHP) methods, help offer recommendations when decisions involve trade-offs among different decision criteria. In prior research into AHP, studies target problem analysis and formulation or system design and implementation, without addressing the user experiences of these techniques online, even though such issues are critical for the application of AHP in electronic commerce and Web-based decision support. We attempt to demonstrate the use of the AHP technique in a multi-criteria product recommendation in the context of a Web-based mobile phone recommendation system. Our application demonstrates the technical feasibility of the system, as we demonstrate empirically through a performance comparison with two benchmark systems based on users' satisfaction.

We choose the mobile phone selection process as our study context for several reasons. First, as mobile phones become increasingly indispensable in everyday life, the number of users worldwide has grown from 170 million in 1996 to 2.5 billion in

* Corresponding author at: Department of Information Management, National Sun Yat-sen University, Taiwan, ROC.
E-mail address: tpliang@mail.nsysu.edu.tw (T.-P. Liang).

2006, an average growth rate of 230 million users per year (Pyramid Research, 2006). In some countries or regions, such as Taiwan and Hong Kong, mobile phone penetration has exceeded 100% (NCC, 2006). A growing number of brands and models compete in this fierce market on the basis of innovative design and advanced functionality. As a result, product comparisons by consumers are becoming more and more difficult, thus favoring the use of a computer-based decision support system to assist consumers in finding what they need or want.

Second, mobile phone selection involves a set of variables, which means we can formulate it as a MCDM problem (Ahn, 2006; MacCrimmon, 1973). Analyses of consumer behavior can identify salient preferences for or expectations about design, functionality, features, appearance, and price. The resulting product search space, described by combinations of these variables, can be cognitively overwhelming for consumers whose limited processing capacity may not support consistent, systematic comparisons. A system-based approach to support consumers' mobile phone selections can mitigate the information overload problem and increase their satisfaction, particularly when it features easy-to-use, Web-based recommendation systems (Liang, Lai, & Ku, 2006). The use of an established MCDM technique to design a recommendation system that can perform efficient and consistent product analyses and comparisons based on customer-provided preferences and constraints is desirable because of its well-formulated analytical basis and systematic analysis (Ahn, 2006; MacCrimmon, 1973; Ryu, 1999). This system-based approach also can reduce the stringent search costs consumers suffer in the form of time or cognitive processing requirements, which should increase their satisfaction. In particular, we adopt the AHP method, which uses pairwise comparisons, explicitly specifies the analysis, and provides robust built-in consistency assessment, validated measurement scales, intuitiveness, and ease of use (Saaty, 1980, 1988; Saaty & Kearns, 1991).

A good recommendation system should be able to improve user satisfaction, a key attribute for customer loyalty and continued use (Taylor & Todd, 1995), which is indispensable to information systems success (DeLone & McLean, 1992, 2003). Measures of user satisfaction with a recommendation system rely of the recommendation content and the system itself. For example, Doll and Torkzadeh (1988) examine user satisfaction from both information and system perspectives and consider them two closely related but distinct constructs. Liang et al. (2006) measure user satisfaction with personalized services by focusing on both the system and its contents. We similarly concentrate on user satisfaction with a system and its recommendations, hereafter referred to as *system satisfaction* and *content satisfaction*, respectively.

Specifically, we empirically evaluate the superiority of the AHP-based system in a controlled experiment with 244 mobile phone users, compared with benchmark systems that use rank-based and equal-weight methods to make recommendations. Our experimental results show that the AHP-based system induces greater user satisfaction than do the benchmark methods. In turn, this study contributes to extant literature by demonstrating the viability and desirability of Web-based recommendation systems for consumers' mobile phone selections, and it further underscores the value of established MCDM techniques to support complex product selections. The results we obtain also generalize to similar decision-making scenarios in other application domains.

The remainder of this paper is organized as follows: Section 2 describes the task of mobile phone selection and provides an overview of the AHP technique. Section 3 explains our AHP model for mobile phone selection, constructed in accordance with the general AHP process, and discusses the associated attribute measurement normalization. In Section 4, we describe our system architecture design and implementation, followed by discussions of our hypotheses and evaluation study design in Section 5. Section 6 highlights the key evaluation results and their implications. We conclude in Section 7 with a summary and future research directions.

2. Background overview and motivation

In this section, we analyze the mobile phone selection problem, summarize the AHP technique, and highlight our motivation.

2.1. Mobile phone selection decision and personalized recommendation support

The rapid expansion of product variety and continually compressed product cycle times have made mobile phone selection increasingly challenging for consumers. The resulting product search space is enormous, making a manual approach to consistent product selection time-consuming and difficult, if not impossible. A system-based approach, in contrast, can support systematic evaluations and consistent comparisons while mitigating the cognitive processing and time required of consumers.

Recommendation systems are (online) computer-based software capable of automatically identifying appropriate choices from a large number of alternative products, on the basis of some specified criteria. A personalized recommendation system supports individual consumers' decision making by considering their preferences or constraints. Various analytical techniques focus on important product attributes or contents, employ collaborative filtering, or anchor for item correlation analysis; for example, data mining remains a popular technique for collaborative filtering or attribute-based recommendation. Most previous research also concentrates on product search support that enables individual consumers to locate relevant, prospective products. However, for decisions that involve trade-offs among multiple criteria, few researchers consider the identification of user purchasing criteria or delicate comparisons during the decision process, especially online, which may be more challenging than traditional product searches (Schafer, Ben, & Riedl, 1999).

For many consumers, comparing decision alternatives involves various brands and models that differ in design, functionality, features, and appearance. Choosing an appropriate phone from a large set often requires a substantial amount of effort and time. Mobile phone selection usually involves a set of variables (e.g., design, functionality, features, and price) and thus can be modeled as a MCDM problem. Among the different MCDM methods or techniques, AHP supports complex decision-making tasks and has been applied to many domains, such as system (software) selection (Cebeci, 2009; Lai, Trueblood, & Wong, 1999; Lai, Wong, & Cheung, 2002; Yazgan, Boran, & Goztepe, 2009), investment risk assessment (Azis, 1990), automobile purchases (Byun, 2001), and decision support system (Cakir & Canbolat, 2008). Chen, Jeng, Lee, and Chuang (2008) use AHP to integrate group's preferences to facilitate a consumer-to-business transaction model. Moreover, AHP can be used to develop recommendation systems (Huang & Bian, 2009). It is capable of integrating a person's judgments in a multidimensional space to produce a single, overall ranking of the competing products. Hence, AHP provides a proper tool for developing a mobile phone recommendation system.

2.2. An overview of AHP

As a well-established, multi-criteria decision support technique, AHP can generate an optimal choice from a set of alternatives, in accordance with specified evaluation criteria or user-provided preferences (Mitra, 1995). This technique supports complex decision-making tasks in various domains, including information systems, management, finance, engineering, and environmental assessment. In general, AHP approaches a decision task from both qualitative and

quantitative perspectives and performs reasonably effectively in different decision-making tasks (Saaty, 1988; Saaty & Kearns, 1991).

To support consistent product comparisons and selections systematically, AHP generally includes three phases: decomposition, comparative judgment, and priority synthesis (Saaty, 1980, 1988; Saaty & Kearns, 1991). In the *decomposition* phase, it formulates a decision task using a hierarchical structure, such that the highest level represents the overall objective and lower levels denote the main evaluation criteria, sub-criteria, and alternatives. In the subsequent *comparative judgment* phase, AHP constructs a comparison matrix at each level according to the user's pairwise assessments of the criteria or sub-criteria under consideration. Finally, the *priority synthesis* phase calculates a composite weight (or score) for each alternative (e.g., product) from the preferences extracted from the matrix constructed in the previous phase. The resulting composite weights generate a relative ranking of the alternatives under examination (typically on a ratio scale), which then indicates an optimal alternative.

The general process for applying AHP to analyze a decision problem is as follows:

Step 1: Create a hierarchical structure of the decision problem by recursively decomposing it into a set of criteria. The decision problem is represented by objective (evaluation) criteria and alternatives. The objective is the root and explicitly states what is to be achieved or optimized. The criteria are derived from progressive decompositions of the target decision problem. The alternatives reside at the bottom of the hierarchy, denoting the competing products under consideration; that is, an optimal product will be selected from these alternatives. The objective and criteria jointly form a tree in which a criterion can be decomposed further into a set of sub-criteria until these sub-criteria are able to assess each alternative on the basis of its respective attributes. Domain experts usually construct the AHP hierarchy by determining the exact number of levels on the basis of their analysis and domain knowledge, as well as the complexity of the decision problem.

Step 2: Assess the relative importance of different criteria by pairwise comparisons, and then calculate the principal eigenvector of the matrix obtained from comparative assessments by the user. Each criterion corresponds to an essential product attribute.

Step 3: Transform the results into corresponding link weights in the AHP hierarchy and evaluate the consistency of the weights. A small consistency ratio (CR) is preferable. Saaty (1980) suggests repeating the pairwise comparisons until the CR ratio reaches 0.1 or lower.

Step 4: Use the resulting link weights to evaluate each alternative. For example, for this study, we generally multiply the weight of each criterion and its sub-criteria along each branch of the tree. The product of these weights represents the user preference, which can be mapped onto each attribute of the alternative under examination. The attribute scores of an alternative are comparable to those of other alternatives (e.g., competing products). For each alternative, we multiply the attribute scores and their relative branch weights to obtain a score that denotes the overall assessment of that alternative, then rank alternatives by their scores. The ranking reflects calculated user preferences or constraints.

Fig. 1 illustrates a sample AHP hierarchy. The root represents the objective; the leaf nodes denote alternative products A_i; and the decision criteria C_i and $C_{i,j}$ reside in between. The solid lines connect the criteria and the objective to form a decision tree, whereas the dashed lines link alternatives to the leaf nodes of
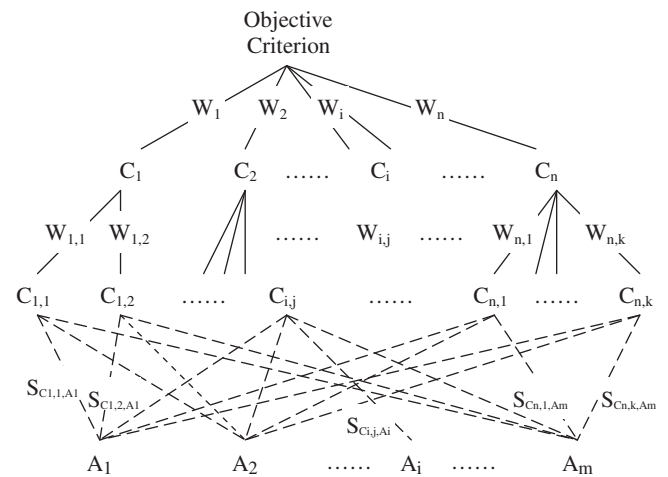

Fig. 1. A sample hierarchy structure of AHP models.

the tree. The link weights W_i and $W_{i,j}$ are determined by the user's inputs through the pairwise comparison of C_i and $C_{i,j}$, respectively. In the figure, $S_{C_{i,j},Ai}$ is the score of alternative A_i with respect to criterion $C_{i,j}$, and $W_{i,j} \times S_{C,j,Ai}$ is the portion of the score that the specific criterion receives in the overall evaluation of the competing products. The resulting value propagates upward until it reaches the root of the tree.

3. Constructing an AHP model for mobile phone selection

Effective AHP-based recommendation systems require an appropriate model construction that considers all essential attributes of the product and solicits their respective importance (weights) from the user. We follow the AHP procedures described in the previous section to obtain consumer preferences or constraints through a pairwise comparison process and build a recommendation system accordingly.

3.1. Constructing an AHP Hierarchy for mobile phone selection

In general, consumers are satisfied with a product when its properties fit their preferences (Jahng, Jain, & Ramamurthy, 2000, 2006). To develop an effective recommendation system, we must first identify the key product attributes that consumers use to select and purchase their mobile phones. We conduct extensive reviews of relevant product descriptions and documents (e.g., industry analysis reports, customer reports, leading business magazines, and major vendors' Web sites) and identify five essential selection criteria: brand, price, hardware feature/functionality, basic built-in functions, and extended built-in functions. Except brand and price, each criterion encompasses several sub-criteria. To determine the relative importance of these decision criteria and sub-criteria, we conduct a survey and obtain responses from 48 phone users from a randomly selected sample of 98 potential respondents (i.e., effective response rate of 49%). On the basis of their responses, we identify the top five sub-criteria for the hardware feature, basic built-in functions, and extended built-in functions. The internal consistency of our question items seems satisfactory, as suggested by a Cronbach's alpha value of 0.67 for hardware, 0.83 for basic built-in functions, and 0.88 for extended built-in functions. Fig. 2 depicts the resulting hierarchical structure of our AHP model for mobile phone selection.

Brand: Brand generally refers to the name, term, design, symbol, or other feature that identifies one provider's goods or service as distinct from those of other providers (Kotler, 2002). A brand
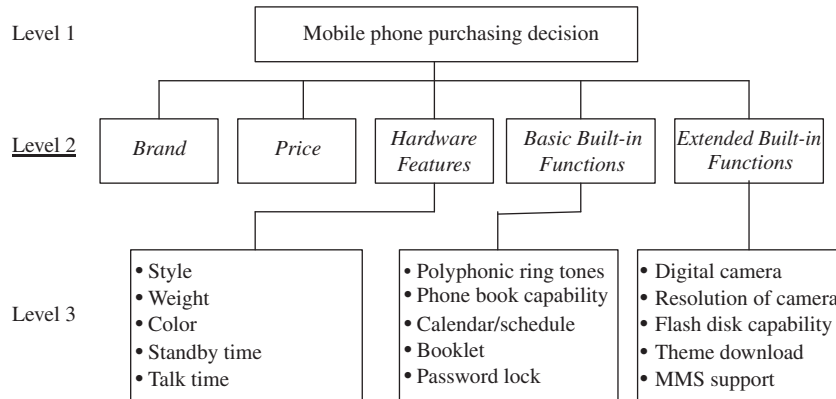
Fig. 2. Hierarchical structure of the AHP model for mobile phone selection.

may identify a product, a product family, or all products of a provider. Brand is an important attribute that signals the overall image of a product (Boyd & Mason, 1999). Our study includes five global brands (Nokia, Motorola, Samsung, SonyEricsson, and LG) and four additional leading brands in the Taiwanese local market: Siemens, BenQ, OKWAP, and Panasonic.[1]

Price: Price is a critical factor in most consumer purchase decisions and often serves as a quality indicator (Mitra, 1995). The importance of price in purchase decisions also may result partly from budget constraints.

Hardware features: Mobile phones often vary in design and feature. Our analysis indicates that consumers pay close attention to several hardware features in their phone selections, including appearance design, weight, color display and resolution, and battery capacity (affecting talk time and standby time).

Basic built-in functions: In addition to hardware features, consumers value software functionality. The basic built-in functions are essential to all models, though their specifics may differ between or among models (e.g., polyphonic ring tones, phone book, booklet, calendar/schedule, and password lock).

Advanced built-in functions: Advanced functions represent software features that exist in some but not all models for differentiation purposes. Our analysis suggests the importance of several advanced functions, including digital camera and resolution, flash disk extension capability, theme download, and multimedia messaging services (MMS).

3.2. Preference collection and normalization

After constructing the decision hierarchy, we must gather user preferences and the relative importance of these respective criteria or sub-criteria, for which AHP employs a systematic method that supports pairwise comparisons of different attributes. For example, we ask users to compare the color and weight of a mobile phone and determine which is relatively "more important." Using AHP, we can draw from users' preferences the weight of each criterion in the AHP hierarchy, ranging between 0 and 1.

Because the comparison is based on attribute pairs with different natures, we need to normalize the result before processing it further. We therefore transform the collected mobile phone attributes into a common 0–10 scale on which 0 denotes "worst" and 10 indicates "best." For example, prices of mobile phones range from $100 to $1,000. We adopt the positive trapezoidal fuzzy numbers approach to convert "price" into the common scale so that we can analyze the impact of price on consumers' decisions. Specifi-

cally, we use a positive triangular fuzzy numbers extension that requires three user-provided parameter values: the maximal price (c), the minimal price (a), and the most preferred price (b). The membership function $\mu\tilde{A}(x)$ is defined as follows:

$$\mu\tilde{A}(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leqslant x \leqslant b \\ \frac{x-c}{b-c} & b \leqslant x \leqslant c \\ 0 & x > c \end{cases}; \quad 0 < a < b < c.$$

For subjective attributes (e.g., brand, color, design style), we use a five-point Likert scale to collect consumer assessments. For dichotomous features (e.g., whether a mobile phone has a polyphonic ring tones, calendar/schedule, booklet, password lock, digital camera, theme download, or MMS), the presence of a feature receives a score of 10, and its absence obtains a score of 0. For attributes with objectively measurable features (e.g., weight, standby time, talk time, phone book capability, camera resolution, and flash disk capability), we rely on a rank list to demonstrate the relative

Table 1
Attribute scale normalization methods.

Mobile phone's attributes		Scale normalization methods
Brand		Five-point Likert scale
Price		Positive trapezoidal fuzzy numbers
Hardware	Style	Five-point Likert scale
Features	Weight	Rank and convert the order into arithmetical series from 0 to 10
	Color	Five-point Likert scale
	Standby time	Rank and convert the order in arithmetical series from 0 to 10
	Talk time	Rank and convert the order to arithmetical series from 0 to 10
Basic built-in functions	Polyphonic ring	Nominal scale (0 or 10)
	Phone book capability	Rank and convert the order in arithmetical series from 0 to 10
	Calendar/schedule	Nominal scale (0 or 10)
	Booklet	Nominal scale (0 or 10)
	Password lock	Nominal scale (0 or 10)
Extended built-in functions	Digital camera	Nominal scale (0 or 10)
	Resolution of camera	Rank and convert the order in arithmetical series from 0 to 10
	Flash disk capability	Rank and convert the order in arithmetical series from 0 to 10
	Themes downloading	Nominal scale (0 or 10)
	MMS support	Nominal scale (0 or 10)

---

[1] Together, the five global brands accounted for approximately 80% of the global mobile phone market in the second quarter of 2006 (Gartner Dataquest, 2006).

positions of different phone models; we then convert these positions into relative scores: 10 for the highest position, 0 for the lowest, and the remaining follow an arithmetical series order. For example, there are five attributes *A* through *E* and the rank list is *B, C, D, A, E*. In this case, attribute *B* receives a score of 10; 7.5 for attribute *C*; 5 for attrite *D*, 2.5 for attribute *A*; and 0 for attribute *E*.

Table 1 summarizes the different scales used in our study, together with their normalization.

4. Architecture design and system implementation

We design and implement a prototype in accordance with the AHP hierarchical structure described in Section 3. This prototype system is then compared with two benchmark systems, one with a rank-based analysis and another with an equal weight-based system.

The rank-based system determines the relative importance of an attribute by its relative rank among all attributes. Such systems are generally considered reliable and easy to use (Eckenrode, 1965). As Watson and Buede (1987) note, rank-based analysis

can solicit the relative weight of different attributes from a decision maker. For each investigated criterion, the rank-based system lists the attributes and asks subjects to identify the top five in descending order. We adopt the following weights for the resulting prioritized attributes: 5/15, 4/15, 3/15, 2/15, and 1/15. In another benchmark system, equal-weight, the user does not specify the attribute weight; instead, each attribute is equally important (i.e., receives an equal weight).

As shown in Fig. 3, the recommendation system follows a three-tier architecture: client, server, and backend database. On the client side, the attribute comparison module gathers a user's preference about the mobile phone attributes through pairwise comparisons. On the server side, the AHP calculation module performs the calculation of weights and scores. The back-end e-store database stores data and information about more than 200 mobile phones. The system was implemented in a Microsoft Windows environment using Internet Information Server (IIS) 6.0 as the platform and Microsoft Visual Studio.Net for the development of our front-end system. The back-end database was built in Microsoft Access. Fig. 4 illustrates a screen shot for the pairwise comparison
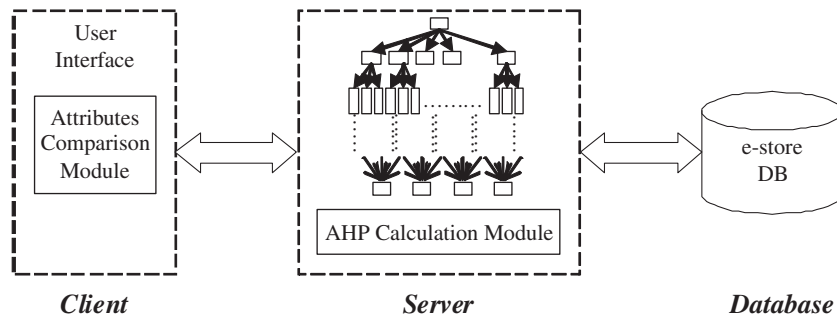

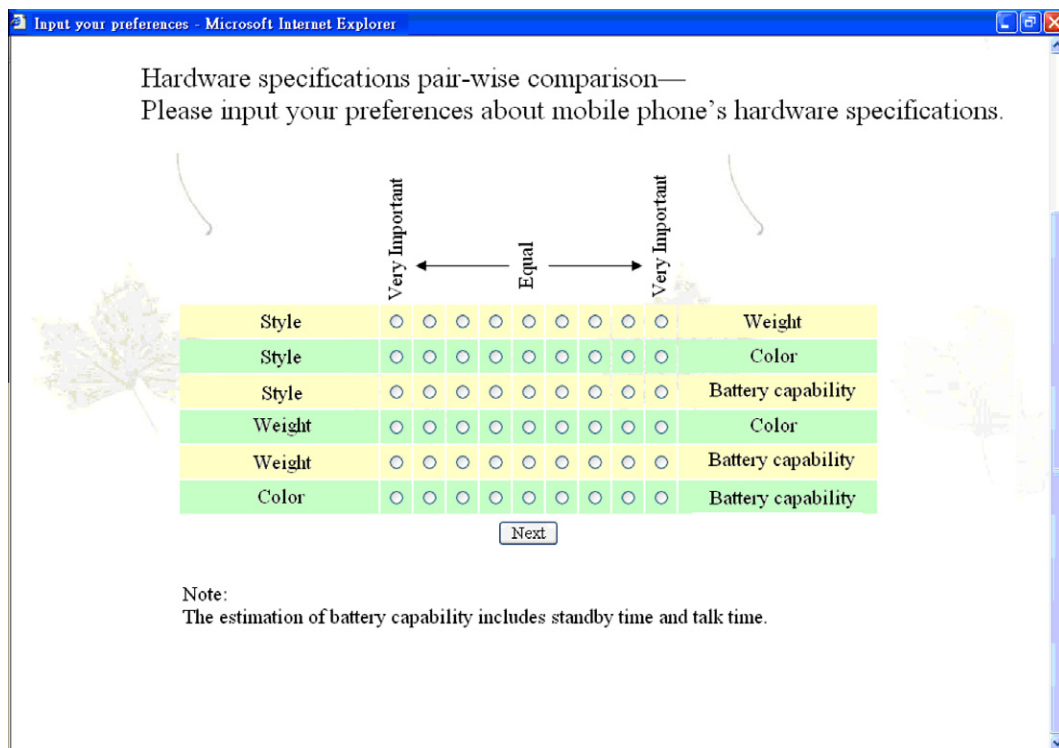**Fig. 3.** Architecture of the AHP-based recommendation system.

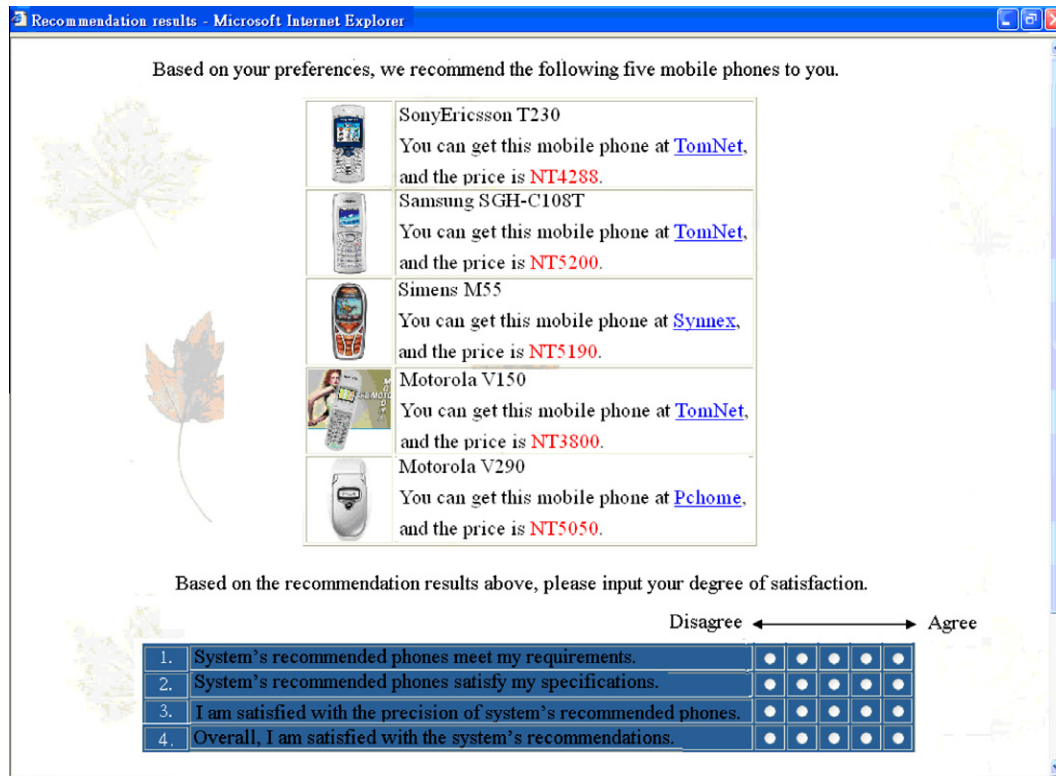**Fig. 4.** A sample screen for pairwise comparisons (hardware features).

Fig. 5. A sample screen for system recommendation and evaluation.

process, and Fig. 5 is the screen that presents the system's recommendations and solicits user satisfaction assessments.

The prototype system allows a user to perform pairwise comparisons of different mobile phones on the basis of the respective attributes, thereby obtaining his or her preferences or constraints. The e-store database, shown in Fig. 3, contains comprehensive data about different mobile phone models, together with appropriate search capabilities. The database schema consists of the attributes identical to those in Fig. 2. The system first collects preference data from each pairwise comparison and then converts them into a normalized scale, using the described normalization process. In turn, the aggregated preference score indicates the most appropriate phone models and makes recommendations accordingly.

5. Experimental evaluation of the AHP-based system

To evaluate the AHP-based recommendation system, we compared it with two salient methods (i.e., rank-order and equal weight-based), with respect to the user's satisfaction with the recommendation (i.e., content) and system use in general. Content satisfaction thus refers to the extent to which a customer is satisfied with the mobile phone models that a system recommends. System satisfaction, in contrast, denotes the degree to which a customer is satisfied with his or her use of and interaction with a recommendation system (e.g., ease of use, interface design) (DeLone & McLean, 1992, 2003). We adapt question items from existing literature to measure content and system satisfaction, with some modifications to fit our context (DeLone & McLean, 1992, 2003). To determine whether the AHP-based system makes better recommendations for product selection, we posit the following two hypotheses:

H1: Users are more satisfied with the recommendation by the AHP-based system than with that by the rank-based system.

H2: Users are more satisfied with their use of the AHP-based system than with that associated with the rank-based system.

Experimental design: We use a randomized between-groups experimental design. Subjects in the treatment group use the AHP-based system, whereas their control group counterparts use a rank-based system. In the treatment group, subjects assess the importance, in a pairwise comparison, of all attributes pertaining to each decision criterion in the AHP hierarchy, whereas the control group subjects evaluate only the top five criteria of consumers' purchase decision, supported by the rank-based system. We examine the content satisfaction and system satisfaction associated with each system, and compare them with the content satisfaction of an equal-weight system. For equal-weight system measures, users do not indicate the relative significance of the attributes, because each of them is considered equally important. The user interfaces and operations of the equal weight-system are identical to those of the AHP-based system or rank-based system; therefore, we do not compare the system satisfaction of the equal weight-system with that associated with the AHP or rank-based system.

Subjects: We compiled a mobile phone user pool from multiple sources and extended invitations by e-mail or phone to 500 randomly selected individuals. A total of 266 agreed to participate in the experiment voluntarily. We randomly assigned each subject to either the treatment or the control group. Most subjects had undergraduate or advanced degrees, used the Internet extensively (i.e., three hours or more a day), and had used mobile phones for at least four years.

Experimental system: The experimental system includes three modules: the AHP-based system, a rank-based system, and an equal weight system.

Dependent variables: Our dependent variables are content satisfaction and system satisfaction, consistent with the suggestions
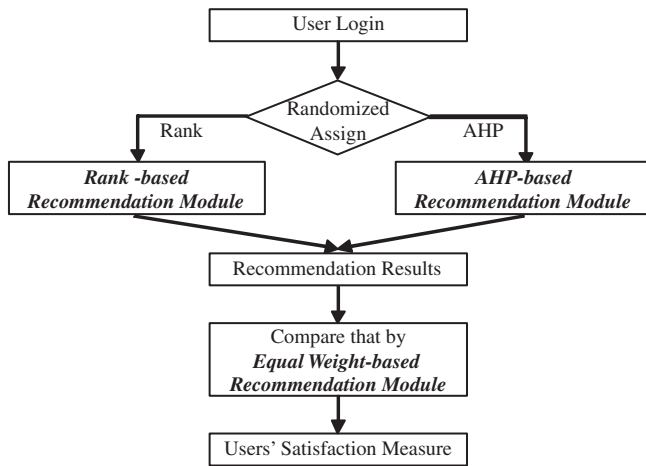
Fig. 6. Overall experiment flow.

Fig. 7. The experimental flow of AHP-based recommendation.

Fig. 8. The experimental flow of rank-based recommendation.

by Doll and Torkzadeh (1988), Jefferson and Nagy (2002), and Liang et al. (2006). We explicitly differentiate the user's satisfaction with the recommendation by a system from satisfaction with his or her use of the system, which can shed light on the key source of user satisfaction, that is, the system's recommendation capability, its design, or both.

Experimental flow: Fig. 6 shows the flow of our experiment. When they enter the password-protected experimental Web site, subjects receive an introduction to the purpose and procedures of the experiment, with a particular emphasis on the anonymity they retain throughout our data collection, analyses, and reporting. Each subject then completes a questionnaire to provide his or her demographic background before starting the actual experiment. We randomly assign them to the treatment or the control group. In the treatment group, subjects proceed through the pairwise comparison in the following order: hardware features, basic built-in functions, and extended built-in functions. The system then recommends five mobiles phone models (in descending order of preference) to each subject on the basis of his or her revealed preference. After receiving the recommendation, the subject reports his or her satisfaction with the recommended phone models and the recommendation system in general.

Fig. 7 shows the overall flow of subjects' use of the AHP-based system. We employ a rank-based system to provide the control group. Subjects identify the top five attributes in descending order according to their preferences; we then assign weights of 5/15, 4/15, 3/15, 2/15, and 1/15 to these attributes.

Fig. 8 shows the experimental flow for the rank-based system. We compare both the AHP-based and rank-based systems with an equal-weight-based system, which offers an appropriate baseline for our comparative analyses.

Pretest: To ensure the validity of the experiment, 14 volunteers participated in a pilot test before the actual experiment.[2] We used their responses and suggestions to make several (minor) changes to the experiment. First, we replaced talk time and standby time with battery capacity, which essentially determines both features. Second, we removed the resolution of the digital camera from the advanced built-in functions, because its inclusion often caused confusion among subjects during the pairwise

comparisons. These changes do not affect the overall experimental flow or the validity of our research.

6. Experimental results and discussion

Among the 266 participants, 14 failed to complete the experiment, so we discard their data. We thus have 132 subjects in the treatment group who used the AHP-based system and 120 subjects in the control group who used the rank-based system. Among the subjects using the AHP-based system, 8 display a CR that exceeds the commonly recommended threshold of 0.1 and therefore are excluded from our subsequent analyses, resulting in an effective sample size of 124.

Our subjects range between 20 and 40 years of age, and the gender distribution is approximately 7-to-3 in favor of men. As summarized in Table 2, we observe no significant between-group differences in gender distribution, education background, or average Internet usage. The subjects in the treatment and control groups also have comparable mobile phone experiences and report a similar frequency of changing mobile phones.

Analysis of reliability: To assess the reliability of our measurements, we examine their internal consistency on the basis of

---

[2] All pretest subjects are active phone users and are highly comparable to the subjects in our study demographically.
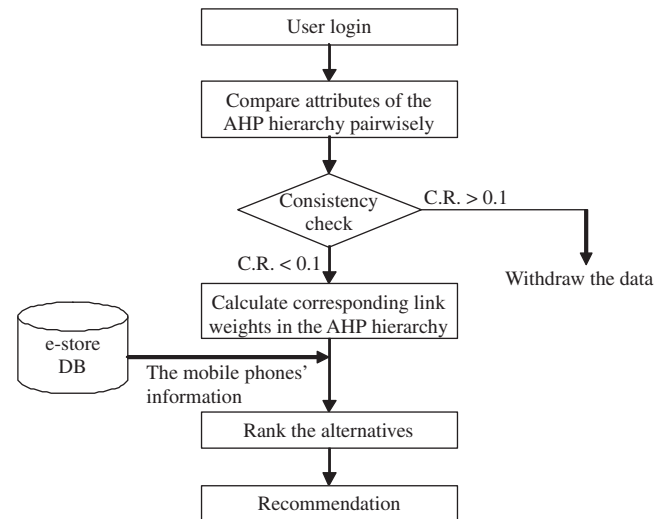
Table 2
Demographic information of experimental subjects.

Demographic dimension		AHP-based system		Rank-based system	
		Number	Percentage	Number	Percentage
Gender	Male	82	66.1	83	69.2
	Female	42	33.9	37	30.8
Age (in years)	Between 13 and 18	0	0	1	0.8
	Between 19 and 24	38	30.6	37	30.8
	Between 25 and 29	38	30.6	35	29.2
	Between 30 and 40	40	32.3	42	35.0
	Over 41	8	6.5	5	4.2
Education	High school	0	0	2	1.7
	University/College	64	51.6	52	43.3
	Master	58	46.8	65	54.2
	Doctoral	2	1.6	1	0.8
Daily internet usage (h)	Less than 2	7	5.6	10	8.3
	Between 3 and 5	35	28.2	32	26.7
	Between 6 and 8	33	26.6	34	28.3
	Between 9 and 11	24	19.4	21	17.5
	Over 12	25	20.2	23	19.2
Number of years in using mobile phone	Less than 1	0.8	0	0	
	Between 2 and 3	8	6.5	8	6.7
	Between 4 and 5	67	54.0	60	50.0
	Over 6	48	38.7	52	43.3
Number of times changed mobile phone	0	4	3.2	3	2.5
	1	17	13.7	15	12.5
	2	34	27.4	35	29.2
	3	34	27.4	35	29.2
	4	19	15.3	14	11.7
	More than 5	16	12.9	18	15.0
Total		124	100	120	100

Table 4
Analysis of convergent and discriminant validity.

Measurement item	AHP-based system		Rank-based system	
	Factor 1	Factor 2	Factor 1	Factor 2
S-1	0.86	0.31	0.87	0.22
S-2	0.86	0.23	0.82	0.27
S-3	0.90	0.23	0.90	0.16
S-4	0.71	0.51	0.71	0.31
S-5	0.33	0.86	0.31	0.77
S-6	0.34	0.75	0.34	0.77
S-7	0.22	0.89	0.10	0.84
S-8	0.23	0.87	0.22	0.82
Percent of variances explained	80.36%		73.65%	

Analysis of convergent/discriminant validity: We examine the convergent and discriminant validity of our measurement instrument through principle components factor analysis (Hair et al., 1998). As shown in Table 4, the factor loadings of the items measuring the same construct (i.e., content satisfaction or system satisfaction) are significantly higher than those that measure other constructs. The factors extracted from our analysis account for 80.36% of the variance in the AHP-based system and 73.65% in the benchmark system. Therefore, our instrument shows adequate convergent and discriminant validity.

Hypothesis testing results: As Table 3 shows, the AHP-based system achieves an average of 3.20 in content satisfaction and 3.53 in system satisfaction, whereas the rank-based system earns an average of 3.06 in content satisfaction and 3.66 in system satisfaction. This finding suggests the AHP-based system performs better than the rank-based system in generating appropriate recommendations but worse in terms of overall system use. One plausible explanation for this finding is that the AHP-based system requires more inputs and tedious pairwise comparisons, which hinder subjects' satisfaction with their use of the system.

To test H1, we perform a one-sample *t*-test to assess whether the average content satisfaction equals 3, the middle value on the five-point measurement scale. If the result of the one-sample *t*-test significantly exceeds 3, users apparently are satisfied with the recommendations by the system under evaluation. Table 5 summarizes our results, including the subjects' satisfaction with the recommendation by the respective systems, using the equal weight-based system as a comparative baseline.

We use a two-sample *t*-test to determine whether the content satisfaction resulting from the use of the AHP-based system equals that associated with the rank-based system. As shown in Table 6, we observe a higher content satisfaction through the use of the AHP-based system compared with the rank-based system; how-

Cronbach's alpha (Hair, Anderson, Tatham, & Black, 1998). Table 3 summarizes descriptive data for the items used to measure recommendation satisfaction and system satisfaction. The alpha values range between 0.86 and 0.92, considerably higher than the common threshold of 0.70. Thus, our measurement instrument exhibits satisfactory reliability.

Table 3
Analysis of content and system satisfaction and reliability scores.

Measurement item	AHP-based system			Rank-based system		
	Mean	S.D.	Cronbach's alpha	Mean	S.D.	Cronbach's alpha
S-1: The system's recommended phones meet my requirements	3.15	1.02	0.92	2.95	0.97	0.89
S-2: The system's recommended phones satisfy my specifications	3.23	0.96		3.16	0.92	
S-3: I am satisfied with the precision of the recommended phones by the system	3.07	0.96		2.93	0.99	
S-4: Overall, I am satisfied with the system's recommendations	3.35	1.06		3.21	1.06	
Average	3.20	0.89		3.06	0.87	
S-5: I find that the system's operations are clear	3.56	0.95	0.91	3.63	1.04	0.86
S-6: I find that the information presented by the system is clear	3.44	0.97		3.58	0.96	
S-7: The system's user interface is user-friendly	3.52	1.01		3.68	0.96	
S-8: Overall, I am satisfied with the system's ease of use	3.60	0.98		3.76	0.93	
Average	3.53	0.87		3.66	0.82	
This system produces better recommendations than those by the equal-weight based system	3.15	1.03		2.88	1.10	

Table 5
Analysis of content satisfaction using one-sample *t*-test.

Construct	AHP-based system		Rank-based system	
	t-Value	*p*-Value	*t*-Value	*p*-Value
Content satisfaction	2.48	0.01	0.80	0.43
Measurement item				
The system's recommended phones meet my requirements	1.59	0.11	−0.56	0.57
The system's recommended phones satisfy my specifications	2.70	0.01	1.89	0.06
I am satisfied with the accuracy of recommended phones by the system	0.84	0.40	−0.74	0.46
Overall, I am satisfied with the system's recommendations	3.64	0.00	2.15	0.03
This system produces better recommendations than those by the equal weight-based system	0.17	0.12	−1.24	0.22

Table 6
Analysis of content satisfaction using two-sample *t*-test.

Construct	Levene's test *p*-value	*t*-Value	*p*-Value
Content satisfaction	0.947	1.221	0.223
Measurement item			
The system's recommended phones meet my requirements	0.125	3.050	0.003
The system's recommended phones satisfy my specifications	0.385	3.142	0.002
Compared with my satisfaction with the recommended phones by the equal-weight based system	0.330	1.974	0.05

ever, this improvement is not significant statistically. When we analyze the content satisfaction at the item level, we observe statistical significance in two measurement items for the AHP-based system versus the benchmark system and the equal weight-based system. Thus, our data partially support H1 at the item level and suggest that the use of the AHP-based system may generate greater content satisfaction than does the rank-based system.

We also perform a one-sample *t*-test to test H2. The results, as shown in Table 7, indicate that the improvement in system satisfaction resulting from the AHP and the benchmark rank-based systems are both significantly positive with respect to the middle value. As summarized in Table 3, system satisfaction associated with the rank-based system is higher than that of the AHP-based system, but the difference is not significant statistically. Hence, we conclude that H2 is not supported; that is, the system satisfac-

Table 7
Analysis of system satisfaction with respect to the middle value.

Construct	AHP-based system		Rank-based system	
	t-Value	*p*-Value	*t*-Value	*p*-Value
System satisfaction	6.806	0.00	8.840	0.00
Measurement item				
I find that the system's operations are clear	6.53	0.00	6.64	0.00
I find that the information presented by the system is clear	5.07	0.00	6.50	0.00
The system's user interface is user-friendly	5.70	0.00	7.68	0.00
Overall, I am satisfied with the system's ease of use	6.89	0.00	8.97	0.00

tion resulting from the use of the AHP-based system is not higher than that associated with the rank-based system. According to our post hoc analysis, subjects need to provide fewer inputs when using the rank-based system compared with the AHP-based system. Therefore, the tedious pairwise comparisons required by the AHP-based system seem to make subjects' use of and interaction with the AHP-based system more difficult.

Overall, our findings show the viability and value of using the AHP technique to construct effective Web-based recommendation systems for mobile phone selection. As highlighted by our experimental results, customers are satisfied with the recommendations offered by the proposed AHP-based system but need to provide more inputs and assessments when using the system. In turn, this finding suggests the need to relax the stringent input requirements common to recommendation systems built on the basis of the AHP technique, as well as the demand for adequate interface designs that can improve users' satisfaction with using the system.

7. Conclusion

In this paper, we have presented the design, implementation, and evaluation of an AHP-based recommendation system that can help users select proper mobile phone models based on their preferences. Our experiment indicates that user satisfaction with the recommendation from the AHP-based system is higher than those from the rank-based and equal weight systems, but not better in terms of user satisfaction with the system due to its lengthy process of eliciting user preferences.

From a research perspective, our study illuminates the feasibility of using the AHP technique to support consumers' online mobile phone selections. In electronic commerce, we could apply this AHP-based system for product recommendations to consumers on the Web and thereby provide an intelligent transaction platform. If consumers are willing to provide their preferences, an online system could perform automated search-and-match functions to generate appropriate recommendations.

With regard to practice, we design and implement a viable Web-based recommendation system and empirically demonstrate its effectiveness. Our architecture design and implementation are scalable and can be extended to support similar product search/comparison problems. Overall, we show that the AHP method can be used as a core algorithm for automated recommendations in complex product search spaces. Our study also has implications for designing and using intelligent e-commerce software agents that offer autonomy, reactivity, and proactiveness (Wooldridge & Jennings, 1994) to support consumers' decision making (Liang & Huang, 2000). For example, an (intelligent) agent could automatically gather essential product information from identified Web sites and, using preferences previously collected from the consumer, identify products appropriate for him or her. The solicitation of the consumer's preferences or constraints would be greatly facilitated by the AHP process, which can be augmented with appropriate data mining techniques for automatic preference/constraint discovery. We use AHP to capture and represent the user's preferences or constraints and thus demonstrate the feasibility of using this technique to build computational algorithms that would enable software agents to discover individual consumers' preferences or constraints.

Our evaluation results point to several areas for extension in the design and implementation of an AHP-based recommendation system. First, the pairwise comparisons required by AHP-based systems are tedious and should be streamlined. Data mining may provide a promising remedy that analyzes the user's previous behaviors to discover his or her preferences or constraints automatically. In effect, researchers could develop more effective recommendation systems for complex product comparisons and

selections by integrating data mining and AHP techniques. Second, the transformation of qualitative attributes to quantitative measures represents a limitation of our AHP implementation, because consumers likely maintain qualitative criteria when making their purchase decisions. Although we propose a method for transforming qualitative attributes, its applicability and effectiveness require further extensions and evaluations.

References

Ahn, B. A. (2006). Multiattribute decision aid with extended ISMAUT. *IEEE Transactions on Systems, Man, and Cybernetics Part A - Systems and Humans, 36*(3), 507–520.

Azis, I. J. (1990). Analytic hierarchy process in the benefit-cost frame work: A post-evaluation of the Traus–Sumatra, highway project. *European Journal of Operational Research, 48*, 38–48.

Byun, D. H. (2001). The AHP approach for selecting an automobile purchase model. *Information and Management, 38*(5), 289–297.

Boyd, T. C., & Mason, C. H. (1999). The link between attractiveness of "extrabrand" attributes and the adoption of innovation. *Journal of the Academy of Marketing Science, 27*(3), 306–319.

Cakir, O., & Canbolat, M. S. (2008). A web-based decision support system for multi-criteria inventory classification using fuzzy AHP methodology. *Expert Systems with Applications, 35*, 1367–1378.

Cebeci, U. (2009). Fuzzy AHP-based decision support system for selecting ERP systems in textile industry by using balanced scorecard. *Expert Systems with Applications, 36*, 8900–8909.

Chen, D. N., Jeng, B., Lee, W.-P., & Chuang, C. H. (2008). An agent-based model for consumer-to-business electronic commerce. *Expert Systems with Applications, 34*, 469–481.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research, 3*(1), 60–95.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A 10-year update. *Journal of Management Information Systems, 19*(4), 9–30.

Doll, W. J., & Torkzadeh, G. (1988). The measurement of end-user computing satisfaction. *MIS Quarterly, 12*(2), 259–274.

Eckenrode, R. T. (1965). Weighting multiple criteria. *Management Science, 12*(3), 180–192.

Gartner Dataquest. (2006). *The market of mobile phones survey*. <http://www.find.org.tw/0105/news/0105_news_disp.asp?news_id=4453> Accessed 05.10.06.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis*. New Jersey: Prentice Hall.

Huang, Y., & Bian, L. (2009). A Bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the Internet. *Expert Systems with Applications, 36*, 933–943.

Jahng, J., Jain, H. K., & Ramamurthy, K. (2000). Effective design of electronic commerce environments: A proposed theory of congruence and an illustration. *IEEE Transactions on Systems, Man, and Cybernetics Part A – Systems and Humans, 30*(4), 456–471.

Jahng, J., Jain, H. K., & Ramamurthy, K. (2006). An empirical study of the impact of product characteristics and electronic commerce interface richness on consumer attitude and purchase intentions. *IEEE Transactions on Systems, Man, and Cybernetics Part A – Systems and Humans, 36*(6), 1185–1201.

Jefferson, T. I., & Nagy, T. J. (2002). A domain-driven approach to improving search effectiveness in traditional online catalogs. *Information and Management, 39*(7), 559–570.

Kotler, P. (2002). *Marketing management*. Pearson Professional Education.

Lai, V. S., Trueblood, R. P., & Wong, B. K. (1999). Software selection: A case study of the application of the analytical hierarchical process to the selection of a multimedia authoring system. *Information and Management, 36*(4), 221–232.

Lai, V. S., Wong, B. K., & Cheung, W. (2002). Group decision making in a multiple criteria environment: A case using the AHP in software selection. *European Journal of Operational Research, 137*(1), 134–144.

Liang, T. P., & Huang, J. S. (2000). A framework for applying intelligent agents to support electronic trading. *Decision Support Systems, 28*(4), 305–317.

Liang, T. P., Lai, H. J., & Ku, Y. C. (2006). Personalized content recommendation and user satisfaction: Theoretical synthesis and empirical findings. *Journal of Management Information Systems, 23*(3), 45–70.

MacCrimmon, K. R. (1973). *An overview of multiple objective decision making*. Colombia: University of South Carolina Press.

Mitra, A. (1995). Price cue utilization in product evaluations the moderating role of motivation and attribute information. *Journal of Business Research, 33*(3), 187–195.

National Communications Commission (NCC). (2006). *Statistic of mobile subscription*. <http://www.ncc.tw/chinese/Data-statistics/11.2/5-Mobile-3625.htm.a> Accessed 05.10.06.

Pyramid Research. (2006). *Total mobile subscription survey*. <http://www.pyramidresearch.com/images/pages/GlobalSubs.gif> Accessed 05.10.06.

Ryu, Y. U. (1999). A hierarchical constraint satisfaction approach to product selection for electronic shopping support. *IEEE Transactions on Systems, Man, and Cybernetics Part A – Systems and Humans, 29*(6), 525–532.

Saaty, T. L. (1980). *The analytic hierarchy process*. NY: McGraw Hill.

Saaty, T. L. (1988). *Decision making for leaders: The analytical hierarchy process for decisions in a complex World*. Pittsburgh: RWS Publications Inc..

Saaty, T. L., & Kearns, K. P. (1991). *Analytical planning: The organization of systems*. Pittsburgh: RWS Publications Inc..

Schafer, J., Ben, J. K., & Riedl, J. (1999). Recommender system in e-commerce. *Proceedings of the First ACM Conference on Electronic Commerce*, 158–166.

Taylor, S., & Todd, P. (1995). Assessing IT usage: The role of prior experience. *MIS Quarterly, 19*(4), 561–570.

Watson, S. R., & Buede, D. M. (1987). *Decision synthesis*. Cambridge: Cambridge University Press.

Wooldridge, M. & Jennings, N. R. (1994). Agent theories, architectures, and languages: A survey. In: M. Wooldridge, & N. R. Jennings (Eds.), *Proceedings of ECAI-94 workshop on agent theories, architectures, and languages, The Netherlands, Amsterdam*.

Yazgan, H. R., Boran, S., & Goztepe, K. (2009). An ERP software selection process with using artificial neural network based on analytic network process approach. *Expert Systems with Applications, 36*, 9214–9222.