http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 10 (6): 1186-1193, 2011
ISSN 1812-5638 / DOI: 10.3923/itj.2011.1186.1193
© 2011 Asian Network for Scientific Information

Design of a Shopping Assistance Agent System

Ming-Yuan Shieh, Juing-Shian Chiou and Chia-Jung Hu
Department of Electrical Engineering, Southern Taiwan University, No. 1, NanTai Street,
Yongkang District, Tainan City, Taiwan, Republic of China

Abstract: This study has proposed an agent system which consists of a humanized interface, a shopping behavior analysis system and an intelligent consulter system, for shopping assistance. Consumers could make an appropriate decision fastly from a plethora of cyber-information through the shopping assistant agent. The humanized interface which is designed as similar as a clever salesclerk, aims to provide the fittest purchasing suggestion for the user. The user will be asked to set up their personal profile initially; then the GA-based behavior analysis subsystem will infer the shopping tendency of the user according to his/her past consumption records and related consumer information. The consulter subsystem consists of a fitting interactive machine to lead the user to conclude a satisfactory transaction. From the proposed experimental results, it is seen that the intelligent agent system not only improves the shopping quality but also increases purchasing desires of consumers.

Key words: Shopping assistant agent, agent system, humanized interaction, genetic algorithms, behavior analysis, consulter system

INTRODUCTION

Following the remarkable advance on web technology, the cyber techniques are used more and more extensively. Cyber-shopping becomes a simple but popular application of e-commerce since web bookstore succeeded. All traders and managers are looking for possible web applications in their own business. A cyber store seems a good choice because of its variety and convenience. To run a well cyber store is just like to make the business never rest. It means that the trade and the competitive power will both be increased consequently after successful cyber commerce (Mehrjerdi *et al*., 2010).

At present, most of the products in cyber store are those with simple functions. They are always promised few after-sale services and usually as test products just for market survey. Sometimes, a cyber store promotes several hot products with lower prices to attract a crowd. As well known, cyber commerce occurs without any limit in time, place, or objects. Although it means that the lower price could be held by higher sales volume. However, it is noted that satisfactory service still dominates sales volume and further cyber commerce (Idris *et al*., 2009; Liao and Lee, 2010).

After analyzing the purchasing characters of consumers, we know that the price is not the unique factor. Shopping convenience is often more important than the price. Moreover, the ambiance of shopping, professional consultation, kindly salesclerk and habitual behaviors of customers are elements of a successful deal. We always can see that a store is thriving because of a careful marketing strategy. In the web, such rules still exist. That is why the intelligent agent system is proposed in this study.

Since cyber-shopping simplifies sales levels and reduces managerial costs, it results in insufficient consumer assistance. The ideas of cyber-shopping although provides the advantages of variety and facility, now they seem not existed in most cyber stores. The simplification of sales flow gives customers more autonomy, however in the meanwhile, customers have to seek shopping advices himself in internet and most of them are puzzled. Consumers receive information of products in abundance in the web but they are always not knowing what or how to choose a fitting product (Jarjes *et al*., 2010; Liang *et al*., 2010).

There are many studies of shopping agents (Chiou *et al*., 2010; Li and Fang, 2009; Yuan *et al*., 2010; Darooei and Khayyambashi, 2010; Yufeng *et al*., 2011) proposed during these years. Most of them are presented in the views of statistics and analysis. In this paper, an intelligent agent system is proposed to provide consumers shopping assistance. The system consists of a humanized interface, a shopping behavior analysis system and an intelligent consulter system. The humanized interface shorten the distance between

Corresponding Author: Juing-Shian Chiou, Department of Electrical Engineering, Southern Taiwan University, No. 1, NanTai Street, Yongkang District, Tainan City, Taiwan, Republic of China

customers and traders, moreover, it acquaints with customers. The agent records users' personal data to determine what kind sales way should be applied. The analysis of consumers' shopping behaviors is based on genetic algorithms in this system. Genetic algorithms learn and evolve from the survey of users' characteristics revealed in shopping. Since different consumers should be suited to different sales manners. The proposed consulter system in this study is designed for six kinds of consultative situations. While the shopping tendency of the user has been determined by the behavior analysis subsystem, the consulter subsystem will adopt the fittest interactive machine to lead the user then to conclude the transaction. The system is expected to be similar to a clever salesclerk and to offer consumers an efficient advice in cyber-shopping.

INTELLIGENT AGENT SYSTEM OF SHOPPING ASSISTANCE, IASSA

System architecture: The IASSA consists of three stages as shown in Fig. 1. The front stage is a humanized interface system; the user can enter the system through one of PDA, PC and cell phone in internet. The middle stage deals with the main operations of the IASSA which consists of analysis, consulter and learning machines. The style and need of the customer in shopping will be diagnosed by the Shopping Behavior Analysis Subsystem (SBAS) and the Shopping Demand Analysis Subsystem (SDAS). These two subsystem are based on the learning machine includes Genetic Algorithms (GAs) and Decision Tree Algorithms (DTAs). In which, GAs evolve the optimal weightings of the SBAS; DTAs train the SDAS then produce the most efficient decision tree for each shopping group. In the rear stage, there are data and knowledges in store in two databases for the machines in the middle stage.

The IASSA intends to classify consumers according to their shopping behavior firstly. Then, it leads the user by the SDAS and concludes his shopping demands from a series of questions and answers. Finally, the consulter system infers an appropriate advice from the shopping style and demands of the user. In this agent system, it is designed as a consultant on shopping. It helps the bushed user to seek for the real shopping demand by a simple but well-ordered Q and A. Like a seer, the IASSA knows the shopping behavior just because of the well-trained analyzer system based on GAs. It is worthy noted, there are the more customers use this system, the more personal data of shopping tendency will be recorded. The agent becomes smarter then concludes better analysis and policy.

System flow: Figure 2 illustrates the detailed system flow of the IASSA. Firstly, the user is requested to register the system while he intends to get some shopping assistance in the entrance page. Next, the member verifier system check out whether the ID and the password of the user are both correct, if not, recheck them until three times or correct. Following, a survey of shopping tendency will be held while the members is less than a fixed volume (two hundred in our study). The results of the survey are used as the initial evolution data of GAs. The shopping behavior analysis system will replace the survey while the members are enough such that the learning of GAs converges.

A survey of consumers' shopping tendency is designed by the results obtained from China Market Information Center (http://market.ccidnet.com /pub/report/show.html?reportid=2284). The investigation in China Market Information Center is based on 3781 samples, in which, each sample records the shopping manner of a consumer who lives in one of ten large cities of China. In the report, authors classify consumers into six groups. Every group has some unique features in comparison with

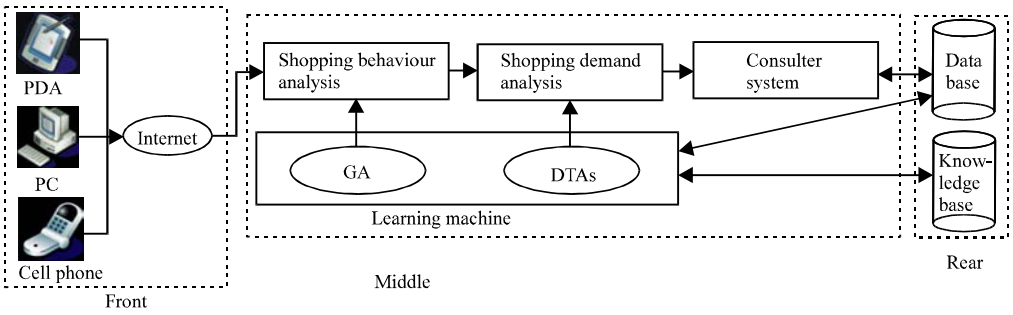

Fig. 1: The architecture diagram of the system

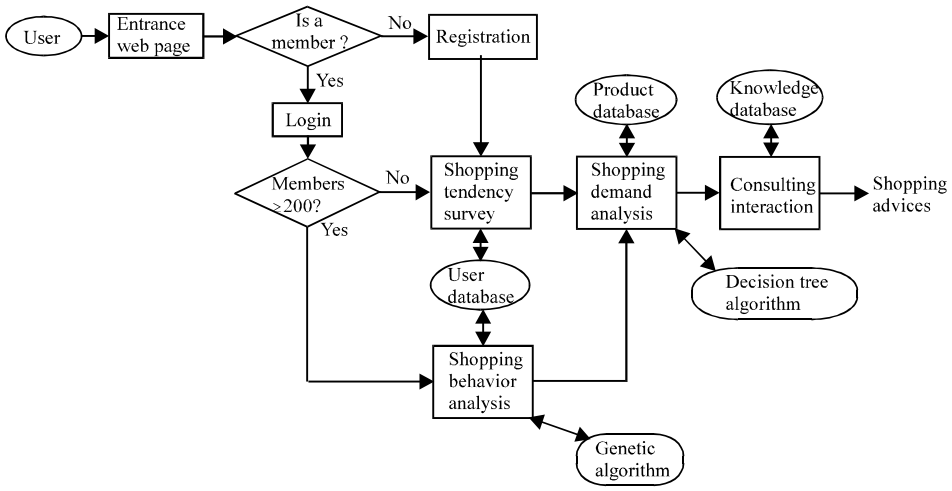

Fig. 2: The flow diagram of the IASSA

Table 1: Features of shopping groups

Shopping groups	Features		
	Character	Hobbies	Distinction
Playful	Like shopping	Pop music, Perfume, Western-style food, Coffee, Pub	Young female, Medium high earnings
Meticulous	Careful calculation and strict budgeting	Tea, smoking	Medium old people, lower earnings
Regular	Periodical shopping	Movie, deama series	Medium old women, all earningspossible
Mission	Fast shopping	Chat by mobile, Coffee, smoking	Male white-collar
Passive	Passive shopping	Tea, drink	No special
Apathetic	Careless in shopping	No special favor	Lower earnings

others. Table 1 shows a summary which the main features of each shopping group are listed in. Based on this report, one can easily design a survey to determine what relationship is existed among the consumers of the same shopping group.

The shopping demand analysis of the consumer follows the shopping behavior analysis. The demand analyzer will adopt a fitting consulter machine to lead the consumer answering a series of simple questions. Form these answers, the system filters off those inappropriate products firstly, then concludes several more suitable ones for the user. Using interactive interface, the user can choose one among these suggestions from the SDAS, or he adopts the unique shopping advice which is custom-made for him from the consulting system. During the operating process, all actions of the user will be stored. The leaning machine of the IASSA keeps on training the subsystems to get the better results. Humanized and intelligent interaction raise the shopping desires of consumers, moreover, the system also improves the adaptability of cyber-shopping.

SHOPPING BEHAVIOR ANALYSIS

Every customer has his own shopping favors. The uniqueness is not only shown in everyone's shopping viewpoint and behavior but also revealed in his purchasing decision. In our study, a survey is designed to analyze the tendency of people in shopping based on the report of China Market Information Center. From the results of the survey for two hundreds of people, one can find some relations between the shopping behavior and personal information such as interests, career, sex, age, etc.. The flow diagram shown in Fig. 3 illustrates the procedure for analyzing consumer shopping behavior.

Genetic algorithm learning: In order to induce the result of the survey, genetic algorithms are adopted to find out what relationship is existed between shopping feature and personal information of a consumer. Carefully observe the behavior of consumer, we find persons with different ages have unlike concepts in shopping. For example, a young girl would like to seek fashionable products but a mature woman usually buys a reliable one. Similar situations occur in interests, career, sex and earnings. It means if one gathers and analyzes enough shopping manners in statistical methods then one can conclude the consumer's tendency.

Since the consumers have been classified into six shopping groups and their personal information are also recorded by the shopping tendency survey. While a user finishes the survey, the system can analyze the relationship between shopping groups and his personal

Table 2: The statistical results in shopping tendency. The data obtained by the case of the user Jack

Variables	Playful	Mericulous	Regular	Mission	Passive	Apathetic	Possible weightings
Age	0.52	0.12	0.04	0.20	0.04	0.08	1
Constellation	0.55	0.05	0.14	0.15	0.08	0.03	3
Sex	0.40	0.10	0.20	0.04	0.25	0.01	1
Blood type	0.37	0.22	0.08	0.20	0.06	0.07	1
Marital status	0.28	0.10	0.10	0.40	0.07	0.05	1
Interests	0.44	0.07	0.19	0.10	0.17	0.03	3
Possible sums	4.54	0.90	1.41	1.59	1.17	0.39	

Table 3: The general case (weightings generated by Gas)

Variables	Playful	Mericulous	Regular	Mission	Passive	Apathetic	Weightings by GAs
Age	F_{11}	F_{21}	F_{31}	F_{41}	F_{51}	F_{61}	w_1
Constellation	F_{12}	F_{22}	F_{32}	F_{42}	F_{52}	F_{62}	w_2
Sex	F_{13}	F_{23}	F_{33}	F_{43}	F_{53}	F_{63}	w_3
Blood type	F_{14}	F_{24}	F_{34}	F_{44}	F_{54}	F_{64}	w_4
Marital status	F_{15}	F_{25}	F_{35}	F_{45}	F_{55}	F_{65}	w_5
Interests	F_{16}	F_{26}	F_{36}	F_{46}	F_{56}	F_{66}	w_6
Membership values	P(1)	P(2)	P(3)	P(4)	P(5)	P(6)	

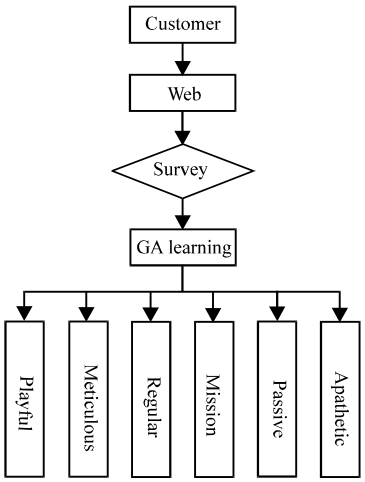

Fig. 3: The flow diagram of shopping behavior analysis

data such as age, sex, interests, career and constellation. For example, an investigated user Jack who is 22 years old, Pisces, young man and a college student likes sport, movie and shopping. Table 2 shows the statistical results in which the line 'age' lists the classificatory percentage of each group among all consumers whose age level are similar to the user. Following the same way, one can find out all elements listed in the table. It is worthy noted that every user must correspond with a unique table. Table 2 reveals the shopping tendency of the user in fact. If the weighting of each line is set as 1, 3, 1, 1, 1 and 3, respectively; then the sum of each row will gives 4.54, 0.9, 1.41, 1.59, 1.17 and 0.39. The maximum is 4.54 in row 'playful'; it means that the user should belong to the playful shopping group.

However, these weightings are not always known in every case. In other words, they should be determined efficiently by a learning machine. Genetic algorithm is a good choice which selects superiors, reproduces new generations, competes with others, evolves until converges into an optimal result.

We define that the population size of genes as 40, every gene consists of six chromosomes which indicate the normalized weightings whose values of 0 to 1. The sum, obtained by multiplying the weightings together the relative elements of each row shown in Table 3, can be regarded as the membership value of each shopping group. Theoretically, the largest value of these membership values must be found in the row of the group given by the shopping tendency survey for the user. Therefore, the performance index of GAs is given by the following equations:

$$\text{Performance Index} = \frac{P_{user} - P_{2nd}}{\sum_{i-1}^{6} P(i)} = PI \tag{1a}$$

$$P(i) = \sum_{j=1}^{6} \left(F_{ij} \cdot w_j\right) \tag{1b}$$

where, P(i) given by Eq. 1b represents the membership value of the ith shopping group; P_{user} means the membership value of user shopping group; and P_{2nd} is the second largest among all memberships. It is noted that the value of PI is positive only when the value of P_{user} is the largest. The larger PI indicates that the inference from GAs and the classification of the survey match better.

To avoid improper learning resulted from special cases, each gene is requested to examine 30 user cases at least which are sampled from all cases randomly. The fitness value of the gene is determined by the sum of these 30 PIs. The GAs selects the first twenty superiors survived. They reproduce, generate, mutate and evolve until the generation nearly changeless. The convergent gene, the winner, dominates the weightings of the SBAS.

Shopping behavior analysis system, SBAS: After the training by GAs, the SBAS has become a cybernetic system. It is because that the survey has replaced by a lookup table as similar as Table 3 with fittest weightings generated by GAs. While the user logins the system, the detail information about him/her will be found from the registered data. As follows, the personal information is responded to those statistics of user shopping tendency. It results in the lookup table, moreover by suitable computations; one can get the membership values of groups. The maximum indicates what shopping group the user belongs to.

SHOPPING DEMAND ANALYSIS

In this study, consumers are classified into six shopping groups. Since different group reveals different type of shopping behavior. The SDAS should set up six kinds of consulting modes for shopping demand analysis of consumers. The consulter system is trained by Decision Tree Algorithms (DTAs). The DTAs produces preference attributes in every mode according to the user shopping tendency. Every mode of the SDAS is trained steadily until finding the minimum tree. The flow diagram of the SDAS is shown in Fig. 4. The user will get some efficient suggestions on shopping from the system. It is just like a sweet consulter especially in the frigid cyberspace.

Decision tree learning machine: The decision tree learning machine aims to find the main attributes of products. Fine classified attribute trees can not only reduce the complexity of searching products but also shorten the time of shopping demand analysis. Most consumers desire to shop in an efficient way. The decision tree algorithms provide realizable generative rules; their computational volume is smaller than other learning algorithms relatively; they can deal with continuous sorting paragraphs; moreover, they show what paragraphs are important explicitly.

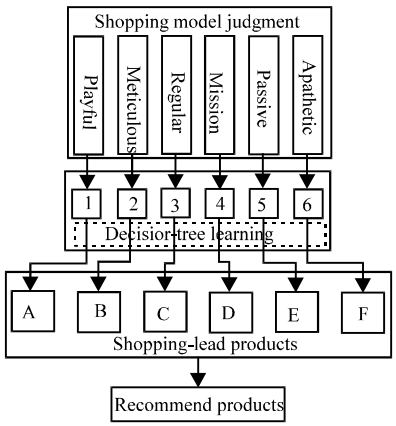

Fig. 4: The flow diagram of shopping demand analysis

There are two particularly influential families of rule induction algorithms: the ID3 family originated by Quinlan (1986) and the AQ family due to (Michalski and Chilausky, 1980). Decision tree algorithm can be regraded as a classification of tree-structure. Each node of tree indicates an attribute which is used to sort data. The attributes in proper order result in an efficient structure. Following a clear and ordered tree-structure, one can easily find the desired data from complex and enormous information. In order to make a fitting decision tree, we have to determine the key attribute. Testing set will be divided to several subsets depending on the key attribute. Every subset becomes a child tree then to find the next key attribute in every child trees repeated until all data are classified into leaves of the tree.

Top-down induction of decision trees: In the standard formulation of the algorithm, the database is known as the training set and each example in the database is known as an instance. It is assumed that there is a universe of objects, each of which belongs to one of a set of mutually exclusive classes. Each instance in the training set must be a description of an object, together with the class to which it belongs. Each object is described by the values of a collection of its attributes (the same attributes being used for each instance), each attribute taking one of a set of discrete, mutually exclusive values.

It is customary to depict the training set by means of a two-dimensional table, with one row per instance and each column holding the values of a particular attribute or the class to which the object belongs. The aim is to develop classification rules in the form of a decision tree which enables the class to which any object in the training set belongs to be determined from the values of its attributes.

Choosing the best attribute: Arranging the testing attributes aims to reduce the 'disorder' of every new child node as possible. It means that to decrease the times of testing action makes the kind of examples similar faster. Based on information theory, the information disorder of the event set S is given by:

$$\text{Entropy}(S) = \sum_{i=1}^{q} p_i \log_2\left(\frac{1}{p_i}\right) \quad (2)$$

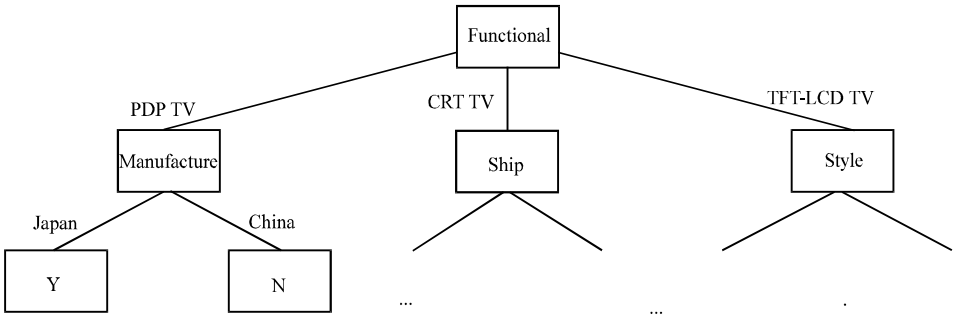

Fig. 5: An example of decision-tree on shopping TV

where, p_i is the probability of the ith event in S and q is the number of events in S. In order to choose the best attribute as a standard to reduce the disorder of information, the information gain is used. The information entropy gain can be defined as:

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} \text{Entropy}(S_v) \quad (3)$$

where, S is the event set before attribute choice; A is the attribute used for classification; Value(A) is the probablity of attribute A; S_v is the event set in which the value of attribute A is equal v in every event.

The basic algorithm can be described informally as follows. If all cases in the training set belong to the same class, then return the value of the class. If not, select an attribute A to split on firstly; next, sort the instances in the training set into non-empty subsets, one for each value of attribute A; finally, return a tree with one branch for each subset where each branch having a descendant subtree or a class value produced by applying the algorithm recursively for each subset in turn. It is important that in selecting attributes at the first step the same attribute must not be selected more than once in any branch.

Selection of the attributes of products: In this study, the decision-tree learning algorithm is used to find out the key attributes of products. These attributes will make a more suitable model for shopping demand analyses. The resultant system which is expected to behave like a clever salesman, leads the user and accomplishes the deal in a skilful way. Taking the purchase of a TV as the example, as well known, there are many kinds of TV products displayed in the mart and how the agent system knows what the user wants is. In this statue, a smart salesclerk might ask the customer firstly what type of the TV he wants to buy. By a series of ordered questions, the purchase demands of the TV will be concluded. Every question indicates an attribute of the product, in decision-tree algorithms, it is revealed as a node of testing. A good classification of nodes leads to a small tree.

Since the consumers are classified into six shopping groups and each group has its own shopping style. For instance, the playful type of customers like to venture on newness, they will pay much attention to the functional products. On views of those who are passive in shopping, the shipping warranty of products will be the most important thing. The differents in shopping tendency result in the different demand analysis models should be adopted. Figure 5 illustrates a possible decision tree of the playful shopping group in shopping a TV, in which, the nodes include the functional, manufacture, ship, style and others.

EXPERIMENTAL RESULTS

The experimental results of the intelligent agent system are shown in Fig. 6-8. Figure 6 illustrates the statues when a user operates the consulting interaction system to seek his shopping demands. In this case, the user follows the consulting guidances in the way of questions and answers firstly. Then, those products unfitted shopping demands will be filtered out by the SDAS as shown in Fig. 7. Since all users are requested to login in the system in registered member, their shopping tendency will be generated by the shopping behavior analysis system automatically. The resultant advice is concluded from the SBAS and the SDAS both the SBAS infers the shopping style of the user; the SDAS based on this inference leads the user in a suitable consulting model. An example of the shopping decision making in TV purchase is shown in Fig. 8, in which, the system produces four suitable products for the user.

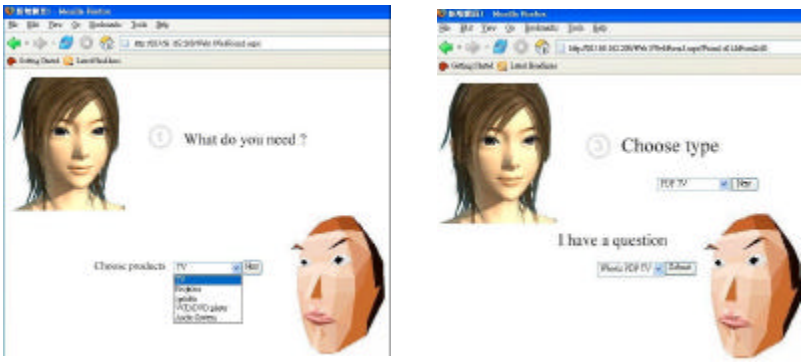

Fig. 6: The operation of consulting interaction

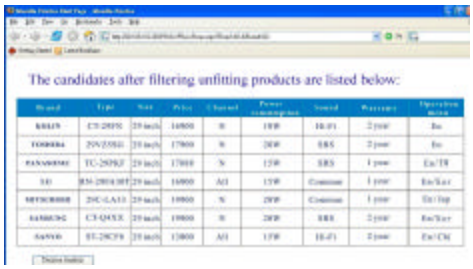

Fig. 7: The remaining products after demand analysis

Fig. 8: Resultant advices of the system

CONCLUSION

In this study, the authors adopt intelligent agent techniques to realize an cyber-shopping assistant system. The proposed system IASSA consists of three stages. The SBAS is designed to analyze the shopping tendency of consumers firstly. Then, the SDAS selects a suitable consulting model according to what shopping group the user is. After the user follows the guidances of consulting interaction system, the system will conclude the shopping advices. Genetic algorithms and decision-tree algorithms are used to generate and to improve the SBAS and the SDAS respectively. From the results of experiments, one can see that the proposed agent system is indeed able to provide proper shopping assistances.

ACKNOWLEDGMENTS

This work is supported by the National Science Council, Taiwan, R.O.C., under grand number NSC 99-2221-E-218-002 and NSC99-2221-E-218-021.

REFERENCES

Chiou, J.S., M.Y. Shieh, K.Y. Wang and Y.C. Hu, 2010. Design of hybrid based image recognition system for intelligent interactive robots. Int. J. Nonlinear Sci. Num. Simul., 11: 109-113.

Darooei, A. and M.R. Khayyambashi, 2010. Design and implementation of an agent-based trading mechanism. Inform. Technol. J., 9: 224-235.

Idris, M.Y.I., E.M. Tamil, Z. Razak, N.M. Noor and L.W. Kin, 2009. Smart parking system using image processing techniques in wireless sensor network environment. Inform. Technol. J., 8: 114-127.

Jarjes, A.A., K. Wang and G.J. Mohammed, 2010. GVF snake-based method for accurate pupil contour detection. Inform. Technol. J., 9: 1653-1658.

Li, X.S. and H.J. Fang, 2009. Optimal weights for consensus of networked multi-agent systems. Inform. Technol. J., 8: 77-82.

Liang, Z., X. Ma, F. Fang and S. Zhu, 2010. Improved monte carlo localization algorithm in a hybrid robot and camera network. Inform. Technol. J., 9: 1585-1597.

Liao, H.C. and P.C. Lee, 2010. A novel vision-based location authentication approach in a ubiquitous camera environment. Inform. Technol. J., 9: 1571-1584.

Mehrjerdi, H., M. Saad, J. Ghommam and A. Zerigui, 2010. Optimized neuro-fuzzy coordination for multiple four wheeled mobile robots. Inform. Technol. J., 9: 1557-1570.

Michalski, R.S. and R.L. Chilausky, 1980. Learning by being told and learning from examples: An experimental comparison of the two methods of knowledge acquisition in the context of developing and expert system for soybean disease diagnosis. Policy Anal. Inform. Syst., 4: 22-34.

Quinlan, J.R., 1986. Induction of decision trees. Mach. Learn., 1: 81-106.

Yuan, X., C.X. Zhao and Z.M. Tang, 2010. Lider scan-matching for mobile robot localization. Inform. Technol. J., 9: 27-33.

Yufeng, S., Z. Chunjie and L. Zheng, 2011. Fuzzy sliding-mode control for the swing arm used in a fourier transform spectrometer. Inform. Technol. J., 10: 736-747.