Community-based Program Recommendation for the Next Generation Electronic Program Guide

Yu-Cheng Chen, Hsien-Chao Huang, and Yueh-Min Huang, *Member*, IEEE

Abstract — *With the popularization of digital TV and Internet Protocol TV (IPTV), more digital programs are being produced. In order to find the programs users want to watch, it is sometimes difficult and time-consuming for them to read the programs' information provided by Electronic Program Guide (EPG). Therefore, the next generation EPG composed of content search, content recommendation, and content management is expected to help users quickly, and easily find programs they are interested in. In content recommendation, this study proposes a Community-based Program Recommendation (CPR) for the next generation EPG. It analyzes and categorizes users with similar viewing habits into one community, then recommends programs to users to help them quickly find the programs they want to watch. The results show that this method can provide good precision and convenience to users*[1].

Index Terms — Recommendation, Personalized, and EPG

I. INTRODUCTION

Currently, there are over 220 million subscribers subscribing to Digital Video Broadcasting (DVB) [1]. The IPTV market is set to reach 311.5 million subscribers worldwide by the end of 2008 [2]. It is expected that more program resources will be available on DVB and IPTV systems for users' choices. It can be quite difficult for users to find the programs they are interested in from the vast amount of information available, and they can easily miss the programs they want to watch.

In the past, EPG provides program information to users. Such information includes program titles, program categories, program summaries, etc. Users utilize the information to find the programs they are interested in. However, more program resources result in information overload problem [3]. It is time-consuming for users to read the program information provided by EPG in order to find the program they are interested in. Therefore, the next generation EPG is expected to help users quickly find programs they want to watch.

The next generation EPG is composed of content search, content recommendation, and content management [4]. As shown in Figure 1 [4], Content Search returns the related program according to simple input by users, such as program key words, program categories, etc. Content Recommendation recommends suitable programs to users. Content Management allows users to modify the program's metadata. The metadata is utilized by Content Search and Content Recommendation.

As shown in Figure 2, the development of digital TV and IPTV lets users can watch TV programs with different kinds of terminals. Users can watch the programs on TV, computers, laptop, or on handheld devices, such as a mobile phone in the office or an outdoor location. Based on the above scenario, this study proposes a Community-based Program Recommendation method for recommending suitable TV programs to users. Its structure is shown in Figure 3. When users watch programs, the terminal will monitor and record their viewing habits daily. Viewing habits of all users are then fed back to the Community Server; such information is utilized to recommend programs that users may be interested in. Community Server divides all users into different virtual communities, according to their viewing habits; the community members all have similar viewing habits. Then, Community Server recommends programs suitable to the users of specific communities and feeds back the results. Finally, the terminal end collects users' responses toward its recommendations to provide better accuracy in its recommendations. The CPR method is detailed in Section III. The remainder of this paper is organized as follows. Section II introduces the related work. Section III describes the CPR method. Section IV presents the implementation of this system, and precision testing. Section V offers conclusions and proposals for future directions.

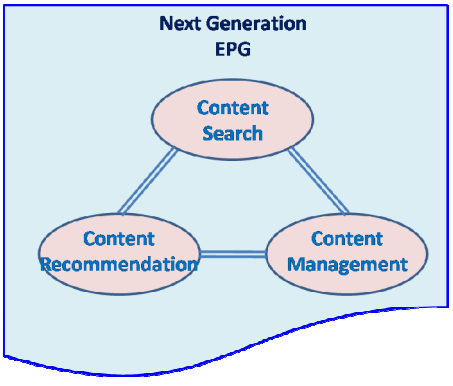

Fig. 1. The kernel functions of the next generation EPG.

[1] Yu-Cheng Chen is with the Department of Engineering Science, National Cheng Kung University, Tainan, Taiwan, R.O.C. (e-mail: cielo@www.mmn.es.ncku.edu.tw).

Hsien-Chao Huang is with the Department of Engineering Science, National Cheng Kung University, Tainan, Taiwan, R.O.C. (e-mail: hchuang@www.mmn.es.ncku.edu.tw).

Yueh-Min Huang is an Professor at with the National Cheng Kung University, Tainan, Taiwan, R.O.C. (e-mail: huang@mail.ncku.edu.tw).

Contributed Paper
Manuscript received January 15, 2009

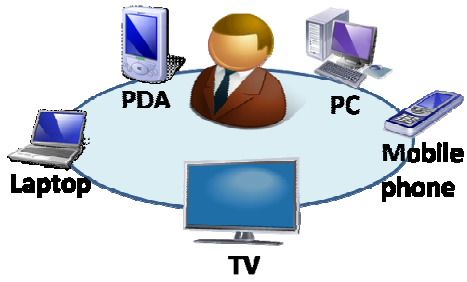

Fig. 2. Users can watch programs with different kinds of terminals.

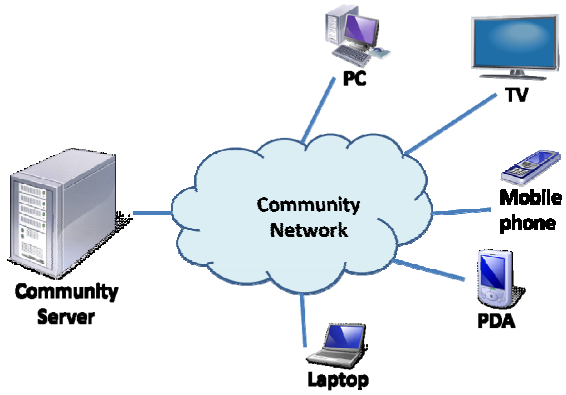

Fig. 3. Community Server collects users' viewing habits.

II. RELATED WORK

In the studies of recommendation methods, PTV [3] utilizes content-based and collaborative methods to provide a web-based personalized electronic program guide (PEPG) listing preferred program types, as input by users. Content-based methods recommend programs to users by comparing program contents against users' previous viewing habits. This collaborative method recommends appropriate programs to users according to their user group, who share similar viewing habits. TV3P [5] utilizes this content-based method to recommend programs to users. However, implementation of the Recommendation system uses an independent set-up box (STB) that could not be linked to a server; thus, Hongguang Z. [6] adopted a content-based method to recommend PEPG to users, through Grid EPG [6].

In above studies, users are required to input their basic data or rate programs to establish their user profile [3] [5] [6]. This study proposed a CPR method, which monitors users' viewing habits, and sends the data to the Community Server. Then, it divides users into different communities according to their viewing habits, and provides the users with choices of programs that interest them. The method proposed in this study allows users to save time when choosing programs, improves program selection efficiency, and does not require users to manually input extra information. It satisfies the requirement of the next generation EPG.

III. COMMUNITY-BASED PROGRAM RECOMMENDATION

A Community-based Program Recommendation for the next generation EPG is proposed in this study. Its structure is shown in Figure 4, which includes two parts: Terminal and Community Server. The terminal includes a Program Information Module, a User Profiling Module, and a Presenting Module. The Community Server is comprised of a Recommendation Module. The four modules are detailed as follows:

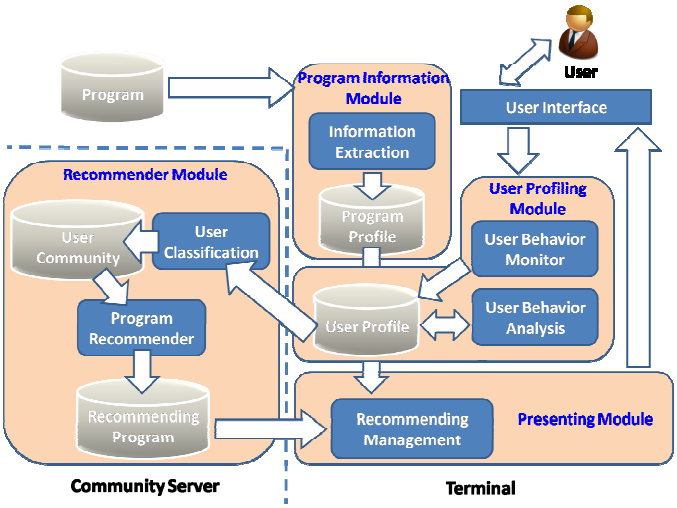

Fig. 4. To depict the function of each module in the Community-based Program Recommendation for the next generation EPG.

A. Terminal

1) Program Information Module

The Program Information Module receives programs transmitted from Program resources. Information Extraction interprets the program-related information referred to the DVB specification [7], such as program titles, program summaries, categories, etc., and stores the information in a Program Profile, which is provided to the Recommendation Module and Presenting Module.

2) User Profiling Module

The User Profiling Module collects users' viewing habits, and stores the information to the User Profile. When users switch channels or the TV program of the channel is changed, the User Behavior Monitor will count the viewing time of the previous program. If it is over 10 min, then it is recorded in the User Profile. In DVB system, TV programs can be classified in to 11 Level-1 types, such as Movie/Drama, News/Current affairs, Show/Game Show, etc. Each Level-1 type can be further divided into several Level-2 types. For example, Movie/Drama can be divided into 9 types: "movie/drama (general)", "detective/thriller", etc. [7]. User Profile records two types of information in XML format, with date intervals, as shown in Figure 5. The first type is total viewing time of the program type. Programs are categorized according to DVB-SI Level-1 types, and arranged in

sequence, i.e., Movie/Drama ☐ News/Current affairs, Special characteristics, and so on. The first number, 15, in Figure 5, indicates that the user watched the Movie/Drama type program on Oct 13, 2008 for 15 min. This information is provided to the Recommendation Module, which then divides users into different communities according to their viewing habits. The second type is (program title, program title, …, program title), a total of 24 elements per group. Each program title is a TV program viewed for the longest duration in each hour. If there is no effective viewing behavior in this time slot, then it is recorded "NULL". This information is provided to the Recommendation Module for recommending suitable TV programs to users. Detailed method will be described in next section. User Behavior Analysis normalizes the data recorded in the User Profile, and checks whether the users select the recommended program to update the User Profile.

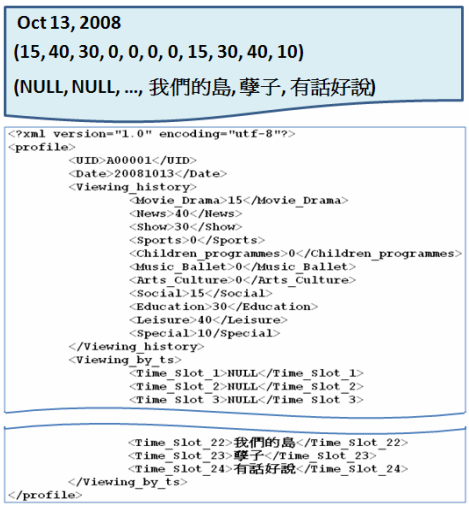

Fig. 5. Two types of data in User Profile and it's XML format.

3) Presenting Module

The Presenting Module receives recommended result S, as delivered from Community Server. The Recommending Management compares S against the Program Profile to determine which programs shall be recommended to users. The decision method is as follows: if one program completely conforms to S in a time slot, then this program will be recommended; otherwise, a program of the same type is recommended according to the Level-2 category in DVB. If neither condition occurs, then a program of the same type is recommended according to Level-1 category in DVB; otherwise, no program will be recommended in this time slot, indicating that there is no program, that users would be interested in, during this time slot.

B. Community Server

1) Recommendation Module

The Terminal sends the User Profile data back to the Recommendation Module at the Community Server end. User Classification divides users into different communities according to viewing habits; users with similar viewing habits are classified into the same community. The clustering basis is as follows:

> If $Sim(U_1, U_2) < \partial$
> U_1, U_2 were classified into the same community

$$Sim(U_1,U_2) = \cos(U_1,U_2) = \frac{U_1 \cdot U_2}{\|U_1\| \times \|U_2\|} = \frac{\sum_{i=1}^{11} c_{1i}c_{2i}}{\sqrt{\sum_{i=1}^{11} c_{1i}^{\,2} \sum_{i=1}^{11} c_{2i}^{\,2}}} \quad (1)$$

Equation (1) shows the similarities between U_1 and U_2, who are two different users. $U_1=(c_{11},c_{12},\ldots,c_{111})$, $U_2=(c_{21},c_{22},\ldots,c_{211})$, denote two different users' preferences to the 11 different TV program types. ∂ is a pre-set value. All users are classified into respective communities, and saved into the User Community. As mentioned in many studies, users with the same viewing habits would have similar preferences [3] [8] [9-11]. Therefore, it could deduce that users in the same community have similar preferences.

To recommend suitable programs to users of the same community, the User Profile records the program that each user watches for the longest duration in each time slot (1 hour), and this program is regarded as the most appealing content to the users, in this manner, a set of 24 TV program titles could be obtained. Among which, if the users have no effective viewing behaviors at a certain time slots, such as viewing for less than 10 min, or no viewing at all, then the program for this time slot is identified as "NULL". The programs in each time slot are deemed as 1 node, and arranged in the sequence of the time slots, in which, the first element is the TV program chosen at time slot 1 (00:00-01:00), the 24th element is the TV program chosen at time slot 24 (23:00-24:00). Therefore, identifying the programs that the users of each community are interested in can be regarded as forming an optimal viewing path from each node. The Ant Colony System [12]is used in Program Recommendation to solve this problem.

In real world, the paths that ants pass by would leave pheromones. When searching for paths, ants tend to move towards the location with the greatest concentration of pheromone. During the same period, more ants would pass through the shorter path, thus leaving more pheromone. Since the shorter path has more pheromone, succeeding ants would tend to choose the shorter path, rather than the longer path. The principle of the Ant Colony System is to allow the more advantageous path to be chosen, as shown in Figure 6.

Programs in each Time Slot are a set of programs that most all users, in the same community, view at this time. 1≦ N≦ Number of user, N=1 represents all users watching the same program, or all users who have no effective viewing behavior in this time slot. When N = number of users; this indicates that all users view different programs.

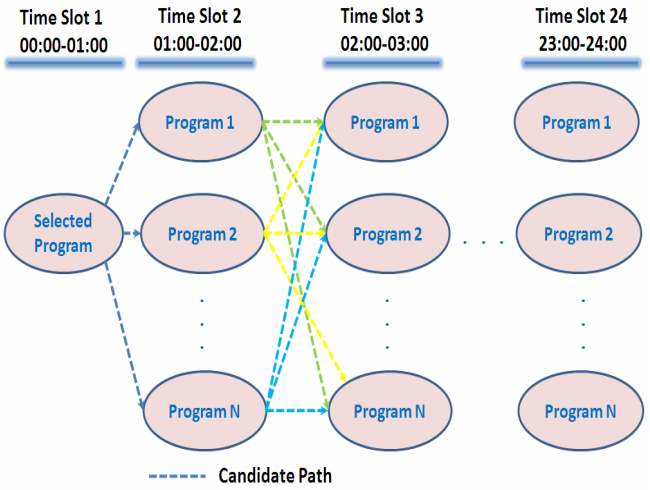

Fig. 6. To search the best viewing path from programs in each time slot.

The ant is randomly place in a node at Time Slot 1, then, ant k in node i tends to go to the next node j. The selection method is as follows:

$$s = \begin{cases} \arg\ \max_{u \in J_k(i)} \{\tau(i,j)\cdot\eta(i,j)^{\beta}\}, & if\ \ q \le q_0 \\ P_k(i,j) & ,otherwise \end{cases} \tag{2}$$

where, q is the random number between 0~1, q_0 is a pre-set parameter. Only when the probability is less than q_0 is the maximal probability path chosen. To consider the possibility of other paths (exploring new paths), if probability is greater than q_0, then path selection is by:

$$P_k(i,j) = \frac{\tau(i,j)\cdot\eta(i,j)^{\beta}}{\sum_{u \in J_k(i)} \tau(i,u)\cdot\eta(i,u)^{\beta}} \tag{3}$$

where, τ is the pheromone, τ(i, j) is the pheromone on the node i→j path, η is the initial pheromone of each path, calculated as follows:

$$\eta(i,j) = \frac{\sum_{c=1}^{N} Cnt(i,j)}{\sum_{c=1}^{N} Cnt(i,u)} \tag{4}$$

Cnt(i, j) denotes if user c has this path, if yes, then it is 1; otherwise it is 0. Cnt(i, u) denotes if user c has the path of node i → u; u denotes any one node at the next time slot, if yes, then it is 1; otherwise, it is 0. β is assumed 1. $J_k(i)$ is the set of nodes where ant k can still go. To avoid being trapped in local optimal solution, pheromone should be updated for the path as each ant passes by:

$$\begin{aligned} \tau(i,j) &= (1-\rho)\cdot\tau(i,j) + \rho\cdot\Delta\tau(i,j) \\ \Delta\tau(i,j) &= \eta(i,j) \end{aligned} \tag{5}$$

where, ρ is the evaporation rate, thus, after an ant passes the path, pheromone on that path will be the remainder (or initial value minus evaporated value), plus that left by the last ant. After the ant passes all paths, pheromone on the optimal viewing path is updated:

$$\begin{aligned} \tau(i,j) &= (1-\alpha)\cdot\tau(i,j) + \alpha\cdot\Delta\tau(i,j) \\ \Delta\tau(i,j) &= \begin{cases} \eta(i,j), & if\ (i,j) \in \ the\ best\ viewing\ path \\ 0 & ,otherwise \end{cases} \end{aligned} \tag{6}$$

The best viewing path S can be obtained from the following equation:

$$S = \min\{Sim(S,U_1), Sim(S,U_2), ..., Sim(S,U_N)\} \tag{7}$$

Then, randomly place an ant in one node of Time Slot 1. Repeat the above procedures, until reaching $Diff_s \leqq Diff_{target}$; $Diff_{target}$ is the pre-set value.

The results are fed back to all users of the community. As mentioned in Section III-A, whether users choose the programs recommended by the system is also recorded by the User Behavior Analysis, and the User Profile is updated to increase recommendation accuracy.

IV. Experimental Result

This study tested a Community-based Program Recommendation for the Next Generation EPG in two parts: Terminal and Community Server. At the Terminal end, USB Tuner supporting DVB-T was linked to a Pentium 4 computer, OS was Debian GNU/Linux SID, and Freevo (Open source project) [13] was used to run the Broadcast Interface Module, User Profiling Module, and Presenting Module by plug-in.

The terminal collected data on viewing behaviors of 100 users from May 1, 2008 to July 31, 2008. The system was executed on Freevo. In the User Profiling Module, the User Behavior Monitor would update the User Profile when any switch of programs was detected. The terminal then fed the User Profile data to the Community Server end, at 24:00 each day.

At the Community Server end, a Recommendation Module was executed on a Pentium 4 computer, with OS of Debian GNU/Linux SID. Python [14] was used to identify user clustering, execute the Ant Colony System, and feed back the new viewing path to users each day.

Finally, the Presenting Module at the Terminal end recommends suitable programs to users each day, according to the decision reached in part C of the previous section. Figure 7 is the recommendation displayed on the Terminal.

The accuracy of recommendation is calculated as follows [15]:

$$\Pr ecision = \frac{recommended \cap \text{int}erested}{recommended} \quad (8)$$

Fig. 7. The recommendation result. User's interesting programs are arranged by time. The broadcasting programs are previewed in right upper corner. The under rectangle shows information of the selected program in left upper rectangle.

In (8), "recommended" is the set of programs recommend by the system each day, and "interested" is the programs that users actually choose. The result is shown in the Figure 8.

Due to insufficient information in the first several days, the average accuracy of the first month was only 60%. However, after the first month, accuracy reached 75%. At the beginning of the third month, accuracy was maintained at 82%.

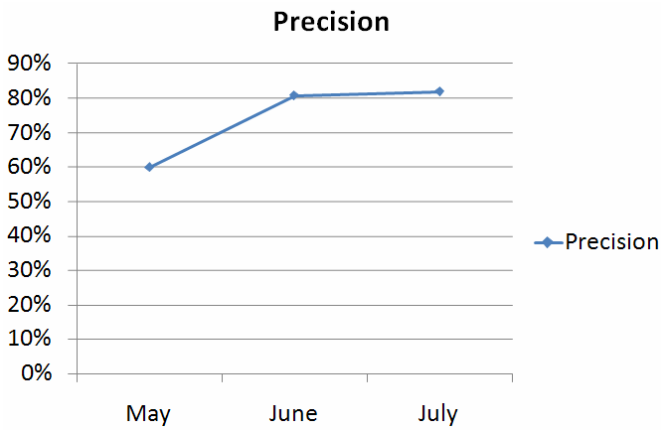

Fig. 8. The average of precision in each month.

V. Conclusion

The Community-based Program Recommendation for the Next Generation EPG, as proposed in this paper collects users' viewing habits, classifies the users with similar viewing habits into the same community, and utilizes the Ant Colony System to recommend suitable TV programs to users, thus allowing them to choose TV programs efficiently, lessening the time spent in searching through TV programs. The users are not required to manually input extra data or program preferences.

In consideration of actual system execution issues, the terminal simply collects user behavior data, and then the complicated computation is run at the Community Server end, and thus, this method could be easily executed on existing DVB and IPTV systems. This method satisfies the requirement of the next generation EPG. Users can quickly find the programs they want to watch.

The results show that the proposed system could provide good accuracy, however, the problem of poor accuracy during the initial stage still exists. Future studies will continue to improve this system by shortening the adjustment period or introduce other methods to make the system more complete and a better-fit for users' needs.

References

[1] DVB, http://www.dvb.org
[2] iSUppli, http://www.isuppli.com/Products/L2_IPTV.aspx
[3] P. Cotter, B. Smyth, "PTV: Intelligent personalized TV guides", 17th AAAI-2000, July 2000.
[4] P. Keys, "Advisory & Intelligence Service Program- Innovation Foresight Program", http://consult.nikkeibp.co.jp, Nov 2008.
[5] Z. Yu, X. Zhou, "TV3P: an adaptive assistant for personalized TV", IEEE Transactions on Consumer Electronics, Vol. 50, Issue 1, pp. 393 – 399(2004), Feb 2004
[6] H. Zhang, S. Zheng, J. Yuan, "A personalized TV guide system compliant with MHP", IEEE Transactions on Consumer Electronics, Vol. 51, Issue 2 , pp. 731 – 737, May 2005
[7] DVB Specification for Service Information (SI) in DVB Systems ETSI EN 300 468 v1.7.1, Dec., 2005
[8] Y.B. Fernandez, J.J.P. Arias, M.L. Nores, A.G. Solla, M.R. Cabrer: "AVATAR: An improved solution for personalized TV based on semantic inference". IEEE Transactions on Consumer Electronics, Vol. 52, Issue 1, pp. 223–231 , 2006
[9] P. Resnick and H.R. Varian, "Recommender Systems", Communictions of the ACM, Vol. 40, No. 3, pp. 56-58, Mar. 1997.
[10] J.A. Konstan, B.N. Miller, D. Maltz, J.L. Herlock, L.R. Gordon., and J. Riedl, "GroupLens: Applying Collaborative Filtering to Usenet News", Communications of The ACM, Vol. 40, No. 3, pp. 77-87, Mar. 1997
[11] J. L. Herlocker, J. A. Konstan, L.G. Terveen, J.T. Lidle, "Evaluating collaborative filtering recommender systems",. ACM Transactions on Information Systems , Vol. 22, No. 1, pp.5-53, 2004
[12] M Dorigo, L.M. Gambardella, "Ant colony system: a cooperative learning approach to the travelingsalesman problem", IEEE Transactions on Evolutionary Computation, 1997
[13] Freevo, http://freevo.sourceforge.net/index.php
[14] Python, http://www.python.org/
[15] Z. Yu, X. Zhou, D. Zhang, C.Y. Chin, X. Wang, and J. Men, "Supporting Context-Aware Media Recommendations for Smart Phones", IEEE Pervasive computing, 2006

Yu-Cheng Chen received the B.S. degree in department of information & computer engineering from the Chung Yuan Christian University, Taiwan, in 2003. He is a Ph.D. candidate at the National Cheng Kung University. His main research interests are in digital TV, recommendation system, algorithm, and embedded system.

Hsien-Chao Huang received the B.S. degree in department and institute of electronic engineering from the National Yunlin University of Science and Technology, Taiwan, in 2000. He is a Ph.D. candidate at the National Cheng Kung University. His main research interests are in digital TV, home network technologies, real-time operation system, and embedded system.

Yueh-Min Huang is a Distinguished Professor and Chairman of the Department of Engineering Science, National Cheng-Kung University, Taiwan, R.O.C. His research interests include multimedia communications, wireless networks, embedded systems, and artificial intelligence. He received his MS and Ph.D. degrees in Electrical Engineering from the University of Arizona in 1988 and 1991 respectively. He has co-authored 2 books and has published about 200 refereed professional research papers. He has completed 13 Ph.D. and over 100 MSES thesis students. Dr. Huang has received many research awards, such as the Best Paper Award of 2007 IEA/AIE Conference, Best Paper Award of the Computer Society of the Republic of China in 2003, the Awards of Acer Long-Term Prize in 1996, 1998, and 1999, Excellent Research Awards of National Microcomputer and Communication Contests in 2006. He also received many funded research grants from National Science Council, Ministry of Education, Industrial Technology of Research Institute, and Institute of Information Industry. Dr. Huang has been invited to give talks or served frequently in the program committee at national and international conferences. Dr. Huang is in the editorial board of the Journal of Wireless Communications and Mobile Computing, the Journal of Internet Technology, International Journal of Internet Protocol Technology, International Journal of Ad Hoc and Ubiquitous Computing, Journal of Security and Communication Networks and serves as an associate editor for Journal of Computer Systems, Networks, and Communications as well as International Journal of Communication Systems. He was the technical program chair of Symposium on Digital Life Technologies (SDLT2007) and general chair in 2008. He was the general chair of VIP2007. He is organizing the PCM2008 and ICST2008. Dr. Huang is a member of the IEEE as well as IEEE Circuits and Systems, Communication, Computer, and Computational Intelligence societies.