Spontaneous and Context-aware Media Recommendation in Heterogeneous Spaces

Daqing Zhang[1] and Zhiwen Yu[2]

[1]Institute for Infocomm Research, Singapore; [2]Nagoya University, Japan

***Abstract*—While mobile users move from one smart space to another, it is highly desirable for them to access the right media contents from the overabundant media information in the right form with their own devices. This paper deals with two important issues in pervasive media access: one is the spontaneous media access in heterogeneous environments, the other is the context-aware media recommendation in different spaces. A general platform for spontaneous media access and context-aware media recommendation has been proposed and implemented. The proposed hybrid recommendation algorithm shows quite good performance in heterogeneous environments.**

I. Introduction

Nowadays, smart phones with the capability of mobile phones and PDAs are becoming more and more popular. Unlike traditional mobile phones, which serve mainly for voice communication, smart phones can store not only personal information, such as contacts, phone numbers, and calendar information, but also richer media such as pictures and video clips. Equipped with network connectivity e.g. Bluetooth and WLAN, smart phones can be used to access media content in a variety of media modalities, e.g. video, audio, image, and text. This precipitates the need to recommend the right content from the overabundant media information, in the right form, to the right person.

A lot of efforts have been put in the design and implementation of smart environments, where media contents are organized in such a way they can be easily and intelligently accessed by the occupants. However, the spontaneous media access was still limited to using specialized hardware and software in specific environments. There is still no solution available for mobile users to use their own mobile devices universally accessing the media contents stored in an unfamiliar environment, based on context such as user preference, location, time, activity, terminal capability, and network condition.

In this paper, we would like to address two important issues in pervasive media: one is the spontaneous access in heterogeneous spaces with one's own mobile device, the other is the personalized media recommendation based on the environment media contents and mobile user's dynamic contexts such as user preference, location and device/network capability. The objective is to enable the mobile user to spontaneously access the right media contents in the smart spaces from the overabundant media information in the right form with one's own devices.

For spontaneous media access, there are two possible ways to make the media contents in heterogeneous spaces accessible to mobile users. One way is to get connected to a global service provider and access the media contents through the service provider; in this case, the service provider is supposed to aggregate all the contents and services in all the smart spaces. Due to the evolutionary nature of the smart spaces, it is unlikely to have such a powerful service provider in the near future. The other way is to enable the mobile user interact with individual smart space directly when the mobile user physically enters the space, that is, the mobile user can automatically discover and access the contents and services provided by the smart spaces. In the second case, the mobile user needs to get connected directly to individual smart space in an ad-hoc manner using local network connectivity and can "see" and access the associated media contents with his/her mobile device.

There have been several projects addressing the issue of spontaneous interaction with environment using mobile devices. The "personal home server" project [1] builds the service platform into the mobile device. When the mobile user enters a certain smart space, the services in the space are discovered via UPnP protocol and integrated by the personal home server in the mobile platform. The assumption for this solution is that the mobile device needs to "speak" the same language as the environment devices and services. Another project, called μJini Proxy [2] assumes that both the mobile device and the middleware framework use Jini to communicate with each other. The problem of this approach is that there is still no dominant service discovery protocol accepted by the community. Several other projects augment the environment with markers or tags [3][4]. By embedding the "tag" reader into the mobile platform, the mobile user can discover the devices and services associated with the tags. The constraint of this approach is that every mobile device and the entities in the environment need to be associated with additional hardware, which is again not easily standardized by the whole community.

This paper aims to impose minimum requirements on the personal access devices such as mobile phone or PDA. All the complexities of coordinating heterogeneous devices and protocols are hidden in the middleware and from the end user.

1550-2252/$25.00 ©2007 IEEE

In heterogeneous smart spaces, devices and services may be integrated in different ways, however, we intend to provide services to the end user in a simple and unified way. In another words, the mobile device doesn't need to install any special software or hardware, when entering an unfamiliar environment, it could automatically discover the services and access whatever services he finds interesting and relevant.

For personalized media recommendation based on context information, we classify context into three categories: **Preference Context**, the context information about user's taste or interests for media content. The preference context includes user requirements, user preference, etc.; **Situation Context**, the context information about a user's spatio-temporal and social situation. The situation context includes location, time, etc.; and **Capability Context**, the context of physical running infrastructure. The capability context includes terminal capability, network condition, etc.

Traditional recommender systems provide recommendations based only on user preference, which can be classified into content-based, collaborative, demographic-based, knowledge-based, and utility-based techniques. To improve recommendation performance, these methods have sometimes been combined as hybrid methods [5][6]. Context information has recently been incorporated in media recommendation systems where situation context, such as location and time, besides user preference, is used to generate media rating [7]. In the area of personalized multimedia delivery, some researchers have considered both user preference and device/network context to generate appropriate presentation to terminals [8][9]. However, none of them handle all of the above three categories of context. Furthermore, they often consider specific context information in a particular application, and seldom propose a formal and effective recommendation model or algorithms to deal with wide variety of context information.

Besides proposing a lightweight user interaction mechanism which enables spontaneous, context-aware service access, we also present a generic media recommendation model. The essence of the model is: context information, ranging from user preference and situation to device/network capability, is considered as input for both content and presentation recommendation, and a multidimensional output including content rating as well as media modality, format, etc. is generated as result of the recommendation. Since a single recommendation method cannot deal with all categories of context, we then propose a hybrid recommendation approach to synergize the content-based, Bayesian classifier, and rule-based methods which are designed to deal with the above-mentioned three kinds of context respectively. Based on the recommendation model and hybrid processing approach, we develop a context-aware media recommendation service, which can be offered in different smart spaces coupled with the media contents. By integrating the media recommendation service and the spontaneous interaction mechanism in a service platform in each smart space, spontaneous and context-aware media recommendation is achieved via popular mobile devices.

II. System Requirements

In order to enable the spontaneous and context-aware media recommendation in heterogeneous environments, a generic set of architectural features need to be supported in the mobile devices and environment. These features include:

A. Minimum assumption on mobile devices

One of the ideal features for spontaneous interaction is that the mobile device doesn't need to install any specialized software or hardware in order to interact with the services or devices in a certain environment. Even though mobile personal devices vary in functionalities and types, they all have wireless communication capabilities and posses certain applications. In order to communicate locally with the smart space, we assume that the mobile devices should have at least the built-in WiFi chipset to allow wireless connectivity and a web browser to access the web server hosting the services of the smart spaces. Apparently, this assumption is biased towards the popularity of WLAN hot spots, it excludes the usage of wireless cellular network for connecting the smart spaces, we purposely made such an assumption for the reason that: 1). The wireless cellular network won't provide enough bandwidth for rich media access; 2). It's still far for single service provider to aggregate and provision services in heterogeneous smart spaces.

B. Automatic discovery of services in heterogeneous spaces

Another ideal feature for spontaneous interaction with mobile devices is the dynamic and automatic discovery of the relevant services in the smart environments. Following the minimum assumption on mobile devices, it requests that when the mobile devices are detected within a certain smart space, the services associated can be automatically discovered and presented in the mobile devices.

C. Context management as services

In order to present the right services to the right person in the right place, context management is needed in each smart space. As specified in the introduction, relevant contexts might include the user preference stored and updated in personal mobile devices, the environment context related to the user acquired in the smart space such as location, orientation, etc., and the capability context regarding the device and network capability. The context manager should be able to acquire, process and provide contexts to specific context-aware applications in smart spaces. For example, "what is the user facing ?", "what's user's preference about watching movie ?" , "what is the screen size of the mobile device ?".

D. Media recommendation as a service

In order to support spontaneous media recommendation in a certain environment, a specific recommendation service needs to be provided and presented to the user when people enter the physical spaces and need it. For mobile users, the media recommendation service needs to be presented with the appropriate input and output interfaces. As the service resides in the environment, the input parameters and the output of the recommendation should be in the mobile device of the user, thus information exchange between the mobile device and the recommendation service is needed.

E. Information profile exchange between the mobile device and the recommendation service

Based on the input and output requirements of the media recommendation model, information exchange between the mobile device and the recommendation service needs to be carried out. As the web browser is the user interface shown in the mobile device and connected with the recommendation service in the smart space, thus the corresponding information profiles need to be defined and agreed upon.

III. System Architecture

The above-mentioned requirements are identified for supporting spontaneous and context-aware media recommendation in heterogeneous smart spaces. To address these requirements, we propose and implement a spontaneous and context-aware media recommendation framework (SCAMREF) which enables the media recommendation service accessible to any WLAN enabled mobile device. In particular, the media recommendation service can provide personalized and context-aware recommendation for the media contents stored in heterogeneous environments according to individual user's preference, situation and device/network capability.

The SCAMREF takes a simple client-server architecture as shown in Fig. 1. To make the media recommendation service accessible to the mobile devices, the framework only assumes that each personal mobile device has a web browser and WLAN capability, it doesn't require to install other additional hardware or software components in the mobile device.

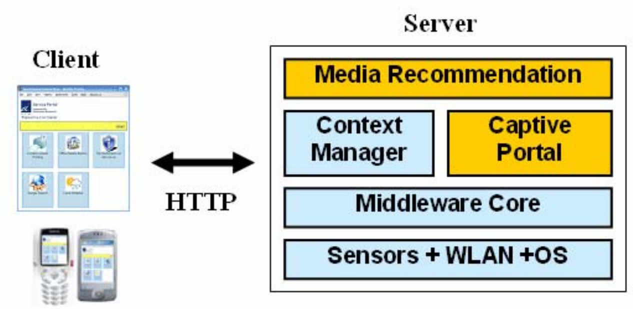

Fig. 1. SCAMREF client-server architecture

The server takes a service-oriented architecture and consists of five basic components: The lowermost layer combines the basic sensing, communication and operating system functionalities, it deals mainly with the context acquisition through sensors/devices and the WLAN communication. The middleware core refers to the service platform such as OSGi [10] or .NET, which can host a lot of reusable basic services and facilitate the integration of heterogeneous devices and services in different environment. We adopted the standard-based service platform for service reusability, interoperability and scalability, various devices and communication protocols are already supported.

The main function of captive portal is to direct the web browser of the mobile device to the media recommendation service of the specific smart space, it enables the automatic discovery and access of the service. With the captive and service portal mechanism, the mobile device only needs the minimum assumption to discover and access the media recommendation service.

The context manager is responsible for context representation, aggregation, reasoning, storage and query processing, our earlier developed "Semantic Space" [11][12] has been used to manage the context about users and the environment. The media recommendation service aims to recommend media contents when the user has no idea about the exact information such as the title, author, etc.. As the media recommendation is based on the user preference context, situation context and capability context, thus some inputs are provided by the personal mobile device. And the outcome of the recommendation should be delivered to the mobile device as well. In the system design, the user authentication and information exchange profile are associated with the recommendation service. When it gets accessed, it will invoke the associated authentication and information exchange process. As the web browser is the user interface between the user and the recommendation service, the W3C Composite Capabilities/Preferences Profile (CC/PP) has been adopted to exchange information between them. The detailed media recommendation mechanism will be described in the following section.

IV. Context-Aware Media Recommendation

A. *Multi-dimensional Recommendation Model*

In order to provide media recommendation for mobile users, we present a generic and flexible multi-dimensional recommendation model. It provides media recommendations over multiple dimensions to generate multi-dimensional output. We define user preference, terminal capability, location, time, etc., as context dimensions, and define modality, format, frame rate, frame size, score (similar to rating), etc., as QoS (Quality of Service) dimensions, which constitute the recommendation output.

Let CD_1, CD_2, …, CD_{N-1} be context dimensions, QD_1, QD_2, …, QD_M be output QoS dimensions, the recommendation model is defined as:

$$R : MediaItem \times CD_1 \times CD_2 \times \ldots \times CD_{N-1} \rightarrow QD_1 \times QD_2 \times \ldots \times QD_M$$

The input is an N-dimensional space comprising one dimension of *MediaItem*, and *N-1* dimensions of context. The output is an M-dimensional space including *M* QoS dimensions.

In general, the output consists of two parts: score and presenting form (e.g. modality and format). The score depends on preference and situation context, while the presenting form is determined by capability context, e.g. whether the terminal can play video and in which format.

B. *Hybrid Recommendation Approach*

Since a single recommendation method cannot deal with all categories of context, we propose a hybrid recommendation approach to determine the right content, in the right form, to the right person based on all categories of context. The recommendation process consists of four steps.

(1) Evaluates between media items and preference context by adopting VSM described in [13].

We model both the multimedia content and user preference as vectors. The *cosine value* of the angle between the two vectors is adopted as similarity measure between media item and preference context. The larger the *Similarity* is, the more relevant between the multimedia content and user preference.

(2) Evaluates between media items and situation context by utilizing Naïve Bayes classifier [14].

We group the values of each situation context dimension into classes. For example, we can divide a user's home location into three classes: Living room, Bed room, and Dining room; social activities into four classes: At party, At date, Accompanying with parents, and Alone. We evaluate the probability of a media item belonging to a class of a context dimension or a combined situation context, e.g. how much probability of the movie *Gone With the Wind* is viewed by the user in *Bed room*, *P(Bed room|Gone With the Wind)*. Suppose $C_1, C_2, \ldots, C_j, \ldots, C_k$ are k classes of situation context considered, the probability of media item $\vec{x}$ belonging to class C_j, that is, $P(C_j \mid \vec{x})$, can be calculated through statistical analysis of user viewing history. Given a class C_j, only the media items that have a high degree of $P(C_j \mid \vec{x})$ would be recommended.

(3) Evaluates between media items and capability context by using rule-based technique.

The modality, format, frame rate, frame size, etc., of the recommended item must satisfy the capability context. We adopt rule-base approach to infer appropriate form (such as modality, format, frame rate, and frame size) from capability context. Our current system applies the Jena2 generic rule engine [15] to support forward-chaining reasoning over the context. The rule below sets that even if the terminal device can display **`Video`**, **`Image`** and **`Text`**, as the available bandwidth is **`Middle`**, the suggested modality is **`Image`**.

```
type(?capability,    Capability),    type(?bandwidth,
Bandwidth), type(?modality, Modality),
  modalitySupported(?capability,   Video+Image+Text),
bandwidthLevel(?bandwidth, Middle)
=>modalityRecommended(?modality, Image)
```

(4) Measures the global evaluating by synergizing above three approaches.

The recommendation output consists of two parts: appropriate form (e.g. modality, format, and frame size) and score. The appropriate form is determined by rule-based evaluating on capability context. The score is composed of the *Similarity* between a media item and the preference context (achieved with VSM approach) and the probability of the media item belonging to the situation context $P(C_j \mid \vec{x})$ (obtained through Bayesian classifier approach). We use a weighted linear combination of these two sub-scores to calculate the overall score as:

$$Score = W_p \times Similarity + W_s \times P(C_j \mid \vec{x})$$

where W_p and W_s ($W_p + W_s = 1$, $0 \le W_p \le 1$, $0 \le W_s \le 1$) are weighting factors reflecting the relative importance of preference context and situation context. If only the preference context is taken into consideration, then W_p=1, W_s=0. On the contrary, if only the situation context is considered, then W_p=0, W_s=1. If both of the preference and situation context are taken into account, W_p and W_s are assigned two positive real numbers.

C. User Preference Learning

As the user preference is the key factor for media recommendation, thus how to effectively get user preference becomes crucial. One automatic approach to acquire user preference is through learning. User preference learning for media recommendation in mobile environment is extremely challenging due to several factors. First, the poor human-machine interactivity of mobile devices, especially the low-cost electronics such as phones and PDAs, causes user explicitly inputting his profile to be nearly impossible. Second, the limited computing power and storage of low-performance devices such as laptops make complex machine learning algorithm impractical. Finally, the mobility of user and devices causes insufficient feedback information that can be obtained for updating user preference.

Fortunately, people nowadays usually own a device with strong capabilities, such as office workstation, personal computer, etc. We therefore address the above challenges by proposing a collaborative centralized learning approach with the "strong capability device" been used in conjunction with the low-cost, low-performance, mobile user devices for user preference acquisition and update. Our approach collects fragments of user feedback information in different pervasive devices so as to build an abundant feedback repository, and then use a strong capability device to implicitly learn user preference from the feedback repository. It can also dynamically merge group user profile by considering interests of most of the members.

The implicit learning algorithm deduces and updates user preference by using a hybrid learning approach. It is designed to integrate user viewing history and user real-time feedback for preference learning. A time-driven learning algorithm is designed for update user profile through compiling statistical analysis on user viewing history, which is aggregated from all kinds of media playing devices, e.g. PC, television, PDA. It is built based on relevance feedback and Naïve Bayes classifier approach. The profile learning by utilizing user real-time feedback is an event-driven algorithm. When a real-time feedback (e.g. switching between channels or programs) happens, the learning algorithm will launch. The details of the user profile implicit learning approach were described in [16].

Services in ubiquitous environment (e.g. watching TV) are often consumed by a group of users, e.g. a family, roommates in a student dormitory, friends in a party, etc. Therefore, the common interest of the group users is needed for the purpose of group-oriented recommendation. The user profile merging algorithm can deduce the group preference by merging individual user preferences into a common one [17]. The key technology of the strategy is based on total distance minimization. Whenever a subgroup of users wants to access personalized services together, the profile merging strategy can dynamically generates the common profile for them.

V. SYSTEM EVALUATION

We have built a prototype for spontaneous media recommendation leveraging the OSGi service platform. In each logical smart space, we install the captive portal on a Linksys WRT54GL WLAN router. On entering the smart space, the web browser of the mobile device is automatically directed to the service portal showing the offered media recommendation service.

We mainly evaluated our prototype by measuring the time spent on recommendation. The time factor is crucial, because long delay will cause negative user experience. We measured the recommending time on a Dell workstation with 2.4 GHz Pentium 4 CPU, and 512 MB RAM running Windows XP by varying the number of movie metadata (also movies) ranging from 40 to 200. When a user was looking for a desired movie, the recommendation engine would be launched. It calculated the similarity between the user preference and each movie metadata, determined the probability of each movie viewed in a specific location and time, and then inferred the displaying modality of the selected content. The rhombus dot line of Fig. 2 showed the result. We observed that the recommending was computationally intensive and as the amount of metadata (or movies) grew, the recommending time would increase proportionally to the size of the movie database.

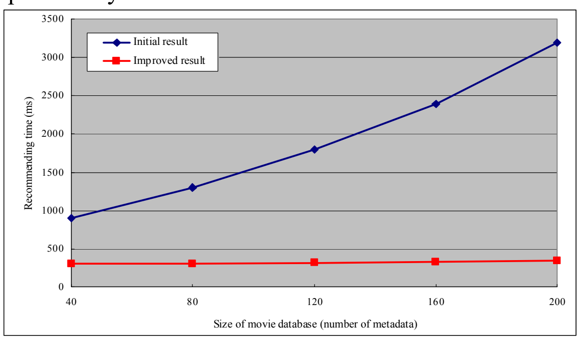

Fig. 2. Recommending performance

We found that if the movie database was very large, the system would not work very well. To achieve scalability when handling extremely large datasets, we split computations associated with recommendation into offline and online phases. The offline phase measured the similarity and determined the probability before hand. The online phase, upon user's recommendation request, inferred the displaying modality of the selected content. Since the offline computation was always running in the background, and its time was transparent to the user, thus only the online time was crucial for the performance evaluation. The improved result was shown as the square dot line of Fig. 2. This time, the recommendation response time was much less and nearly constant. According to most of our invited testing participants (7 of 9 real users), the time spent on obtaining recommendation was quite acceptable (0.3s).

From the user's perspective, we also evaluated the quality of recommendation. The result was encouraging with most *precision* values around 0.8, and *recall* values ranging from 0.63 to 0.75.

VI. Conclusion

We have proposed and implemented a platform for spontaneous and context-aware media recommendation in heterogeneous spaces. It has the following distinct features:

- Enabling spontaneous discovery of services in heterogeneous environments with a "minimum assumption" on mobile devices.
- Allowing context-aware media recommendation based on the media contents in a specific environment and the mobile user's context.

References

[1] NAKAJIMA, T. and SATOH, I., "A software infrastructure for supporting spontaneous and personalized interaction in home computing environments". Personal and Ubiquitous Computing, pp 379-391, 2006.

[2] LEE, C., etc., Universal Interactions with Smart Spaces. IEEE Pervasive Computing, Vol. 5, Jan. 2006, pp. 16-21

[3] RUKZIO, E., etc., "An Experimental Comparison of Physical Mobile Interaction Techniques: Touching, Pointing and Scanning", Eighth International Conference on Ubiquitous Computing (Ubicomp 2006), California, USA. 17-21 September 2006.

[4] HINCKLEY, K., etc., "Sensing Techniques for Mobile Interaction", ACM UIST 2000 Symposium on User Interface Software & Technology, CHI Letters 2 (2), pp. 91-100.

[5] Burke, R. "Hybrid Recommender Systems: Survey and Experiments", *User Modeling and User-Adapted Interaction*, 12(4), 2002, pp. 331-370

[6] Balabanovic, M. and Y. Shoham. "Fab: Content-based, collaborative recommendation", Communications of the ACM, 40(3), 1997, pp. 66-72

[7] Adomavicius, G., et al. "Incorporating Contextual Information in Recommender Systems Using a Multidimensional Approach", ACM Transactions on Information Systems, 23(1), 2005, pp. 103-145

[8] Tseng, B.L., et al. "Using MPEG-7 and MPEG-21 for Personalizing Video", IEEE Multimedia, 11(1), 2004, pp. 42 – 52

[9] Steiger, O., et al. "MPEG-based Personalized Content Delivery". IEEE Intl Conf. on Image Processing (ICIP 2003), Spain, 2003

[10] Open Service Gateway Initiative (OSGi), http://www.osgi.org

[11] Xiaohang Wang, et al. "Semantic Space: An Infrastructure for Smart Spaces". IEEE Pervasive Computing, 3(3), 2004, pp. 32-39.

[12] Yu, Z., et al., "Supporting Context-Aware Media Recommendations for Smart Phones, IEEE Pervasive Computing, Vol. 5, No. 3, July-September 2006, pp. 68-75

[13] Zhiwen Yu. "TV3P: An Adaptive Assistant for Personalized TV". IEEE Transactions on Consumer Electronics, 50(1), 2004, pp. 393-399

[14] Joachims T., "A probabilistic analysis of the rocchio algorithm with TFIDF for text categorization". Proc. of the 14th Intl Conf. on Machine Learning, pp. 143-151, 1997.

[15] J.J. Carroll, et al. "Jena: Implementing the Semantic Web Recommendations". Proc. of 13th Intl Conf. on World Wide Web. New York. 2004.

[16] Zhiwen Yu, et al. "A Hybrid Learning Approach for TV Program Personalization", The 8th International Conference on Knowledge-Based Intelligent Information & Engineering Systems (KES 2004), pp. 630-636.

[17] Zhiwen Yu, et al. "TV Program Recommendation for Multiple Viewers Based on User Profile Merging", User Modeling and User-Adapted Interaction, 16(1), March 2006, pp. 63-82.