Context Aware Recommendation of Location-based Data

Karol Waga, Andrei Tabarcea, Pasi Fränti

Speech & Image Processing Unit, School of Computing, University of Eastern Finland, Joensuu

***Abstract*—There is much information available, but the problem is how to find which is relevant. We present a context aware recommendation system, which recommends relevant location-based data. We study its performance within MOPSI service that includes fixed form maintained database and free form user collection.**

***Keywords*—Recommendation, relevance, context-aware, location, content, network, time, location-based applications**

I. INTRODUCTION

Recommendation systems are important research and are in scope of interest of both universities and companies [1]. Recommendation systems produce personalized search results by performing analysis of user actions [3]. Such systems can be used, for example, for recommending similar products in online stores, music or videos which may be of interest of particular user, and advertisements targeted to specific audience. Recommendation system takes into account additional information about user, which is called context. Examples of contexts we identified are user's location (distance to the service), identity (age, gender, hobbies and language), social network, history of activities, time, technical resources (network accessibility, bandwidth), and the purpose of use (work, leisure time).

Location is very important attribute of our data. Mobile technology is increasing its popularity and availability and it allows collecting location data [9]. Furthermore, mobile phones are one of the main devices for information access [19]. Because of technical limitation, such as bandwidth and screen size, recommendation system can be used to reduce amount of information presented to user. Recommendation system can consider user's location for recommending the nearest service (ATM, restaurants, pubs, tourist sights and social events). However, the nearest service may not be the most interesting for user and other factors should also be considered.

Our goal is to design a context-aware recommendation system based on the four aspects of relevance (content, time, location and social network) discussed in [7]. For recommendation we use the MOPSI services[1] geo-tagged database, which contains user-generated photo collection and service database. Our solution is designed and implemented as a prototype solution within MOPSI (see Fig. 1), as a case study. The MOPSI project implements various location-based services and applications such as mobile search engines, data collection, user tracking and route recording. It has applications integrated both on web and in mobile phones.

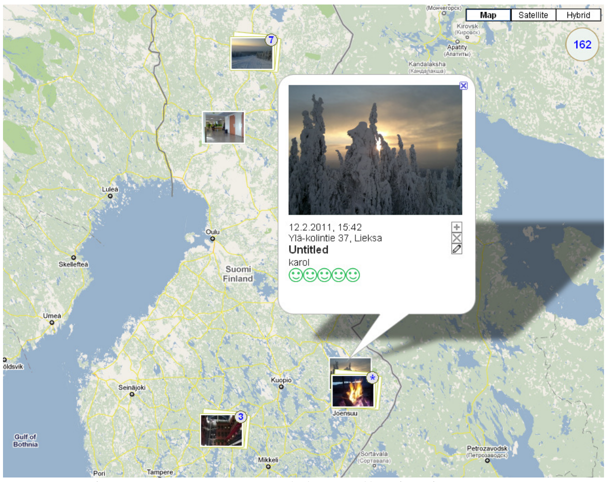

Figure 1: User generated data collection in MOPSI.

II. RELATED WORK

Content-based multimedia information retrieval has been described in [16]. Exploiting synergy between the various media, including text and context information is described as one of the major challenges for the future. Our recommendation system has the goal of combining textual, context and multimedia information in order to get relevant results.

Recommendation systems produce personalized search results relying on a variety of contextual information. Main approaches are *collaborative filtering*, *content-based filtering* and *knowledge-based recommendations*. We aim at as little user interaction as possible during the recommendation process, and because our data is mostly user-generated, collaborative filtering is the most suitable approach in our case. According to [6], collaborative filtering is not sufficient in modern recommendation system because of limited contextualization; therefore we use hybrid approach, similarly as in [4]. The advantage of using hybrid recommendation when applying for recommending web pages was emphasized also in [10].

An intelligent multimedia browsing system was described in [2] by combining the use of patterns with low-level features. Their main contribution is using the browsing behavior of users combined with features of objects to provide recommendations for content-based multimedia retrieval. We also monitor behavior of users (recent searches, searches in the particular area) for better

[1] http://cs.joensuu.fi/mopsi

recommendations. In specific, our photo recommendation problem is similar in a sense that we also recommend from a multimedia database (photo, descriptions, locations). Although a multimedia database is used in both cases, the main difference is that we use location and our recommendation is mostly location-oriented. Furthermore, they use only the previous behavior of the current session (there is no explicit login), whilst we store previous behavior history for every user.

Most of the existing recommendation systems are tested using restaurants and tourist targets. Our system aims to be more general and can be summarized by the phrase "What is interesting in the area?". In [17], a map-based and context-based personalized recommendation system was proposed using Bayesian Networks, which is applied to restaurant recommendation. This is similar to our service, but we do not use Bayesian networks because of lack of flexibility and the user does not explicitly input his profile and preferences.

The method in [22] has two part recommendation system which contains the location-based data (traffic, road-condition, map, service information) and value-added data provided by users (ratings, comments, blogs and tags). They focus on how to gather value-added data provided by users and how to extract it from Web 2.0 websites (ratings sites, comments, blogs). We extract our value-added data from explicit user ratings and from user previous behavior.

Browsing data and image semantics is used in [14] for recommending purchasable items. One of our future goals is to use image semantics as an extra criterion in our photo recommendation. In [20], a recommendation system of shops based on users' past location history was described. The frequency of shop visits by users are the input of an item-based collaborative filtering algorithm and the predictions are narrowed down based on user movement and geographical conditions in the city. Our MOPSI system also tracks down user movement (explicitly, on user request) and our future goal is to use this tracking data in determining the area of interest of the user and assigning extra relevance score to the items in the database that are frequently visited or photographed by other users. Reference [11] shows a good example on how to effectively use social network in order to provide better recommendation by defining and implementing a trust-based model. The key concept is trust between users and two trust metrics are defined: *personal trust* and *person-group trust*. We also use the social network as a relevance criterion, although the current solution is still in a very early stage.

Location-based recommendation systems using mobile devices have been described in [12], [15], [13], [21], who tested it either for restaurant or tourist recommendation using various hybrid techniques and contextual information.

Reference [5] has similar motive to ours by proposing a system to suggest tourist destinations based on visual matching and minimal user input. The input is a photo of desired scenery or a keyword describing a place of interest and the output is a place that shares the visual characteristics. The system uses a large-scale geo-tagged web photo collection (Flickr). The queries are compared against representative tags or representative images in order to discover interesting tourist destinations. Their main contributions are: an efficient clustering of the geo-tagged image database, proposal of using representative images for having faster retrieval, and designing a flexible interface to allow user to enter either keyword (such as architecture, beach, flower.) or image to describe its interest. Although our recommendation is more general, there main similarity is that we also use a geo-tagged image database as input. In addition, our system can also take keywords as input, although this uses a search algorithm. However, this is not the focus of this paper and has been documented in [8].

III. SYSTEM DESCRIPTION

In this section, we provide description of what our system actually recommends and how it uses the four main aspects of relevance as the context.

A. *What is recommended*

Our aim in MOPSI is to recommend interesting places in user's surrounding. In the service, we have two databases that are used for the recommendation (see Fig.2). First database contains trusted services verified by administrators. These services represent variety of categories from restaurants, bars, and cafeterias, through grocery stores, pharmacies, and ATM machines, to car repairs, and museums. Service data include location, contact information, and relevant keywords.

Second database contains photos users have taken using mobile phones and uploaded real-time with several related information, such as location, time, and description. Both photos and services (referred as *items* from now onwards) are rated by users. Moreover, photos can picture any place, which is found interesting by the user.

B. *Recommendation methods*

In our recommendation system, we give personalized recommendations by combining various paradigms of recommendation systems. We combine collaborative filtering with information about user profile and context.

Having these two sources for recommendation, the challenge is how to select the most relevant items to users. First we define context for each recommendation request. In our previous work we identified four aspects of relevance: location, content, time, and network [7]. Location is physical place of the user represented by geographical coordinates (latitude and longitude). Content in MOPSI is determined currently based on the description of the photos and keywords attached with the services. Time is considered only for photos and measures age of photo and the season (of the year) when photo was taken. Network is utilized via ratings given by other users to items and it constitutes an integral part of the system based on collaborative filtering. Considering these context in mind, we create profile for each user of MOPSI.

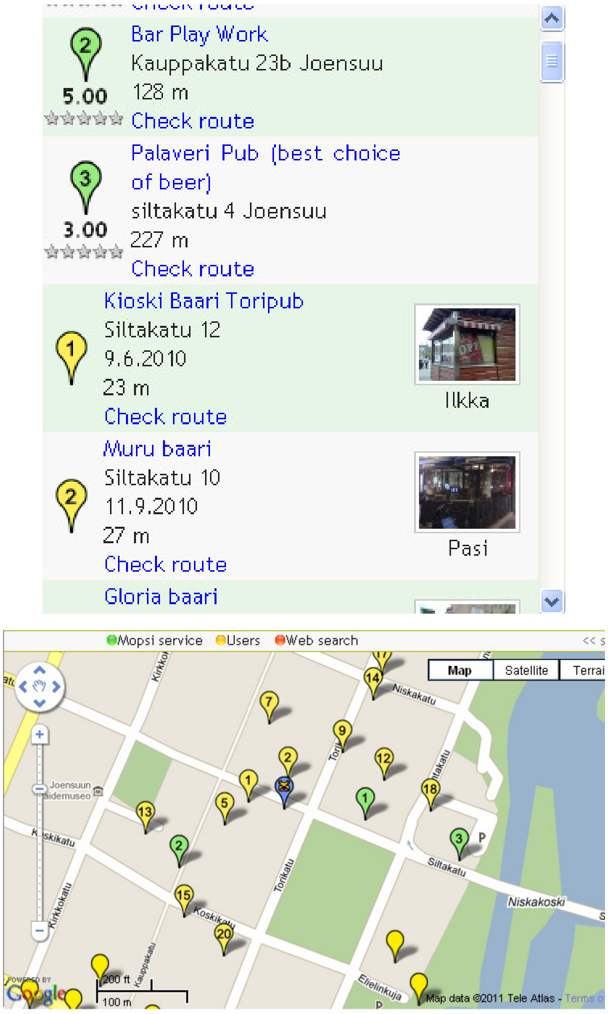

Figure 2: Data from both databases visualized as a list (top) and on map (bottom) with user position marked with blue bubble

User profile contains user behavioral data, such as location and previous usage of data, i.e. how user interacted with the system. Currently this is measured by the keywords user has performed earlier searches and visited locations.

C. System description

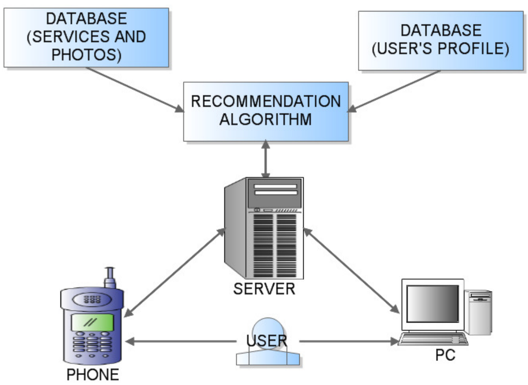

Figure 3: System diagram

Users can access the recommendation system through MOPSI website and mobile application. Mobile access to the service is important, since this is the most natural environment where the system can be most beneficial in real life. It is key functionality of the system that user can ask for recommendation in any location when he can easily visit the suggested places immediately. When recommendation request is sent to server, after computing, list of recommended items is returned to user as illustrated in Fig. 3.

D. System implementation

In this section, we describe in details how we implement the system. Summary of the algorithm is followed by description of the scoring function used by the algorithm.

1) Recommendation algorithm

Recommendation algorithm consists of three major steps. Its input is the username and location of the user. The steps of the recommendation algorithm are outlined in Fig. 4. First step is to select potential items that are to be considered. We use location as the criterion for this pre-selection. Items that are far from user are considered irrelevant and are skipped already at this stage. The selected items are then scored in the second step. The scoring function will be described in details in Section IV.

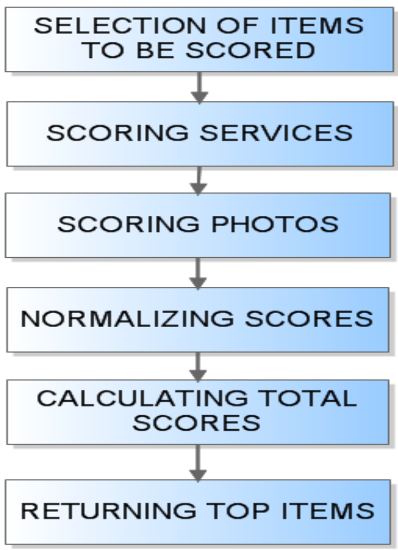

Figure 4: Scoring function schema

Third step is to prepare final list of the recommended items ordered according to the scores received in the second step. The system outputs the final recommendation list, which consists of 20 items of the highest scores.

2) User interface

User interface is provided both for the MOPSI mobile application and the website. In web, the recommendation function is embedded on the MOPSI main page where user can request recommendation by pressing a single button (see Fig. 5). The results are visualized on screen in two ways. On the left, there is scroll list of the recommended items including title, street address and distance. Photo results include the description user has entered (if available), street address, date, distance, photo thumbnail and author. All the results are marked on the map visible on right side of the screen. Services are marked with green bubbles and photos with yellow bubbles. User location is marked by blue bubble with "*Mopsi dog*" icon.

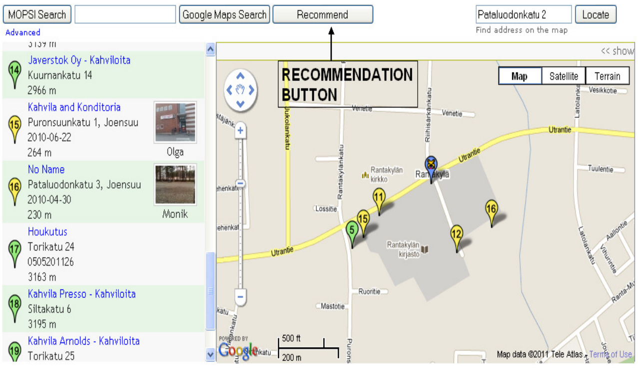

Figure 5: Web interface for the recommendation system.

Figure 6: Mobile interface for the recommendation system. MOPSI Services screen in main menu on the left. Recommendation results list on the right.

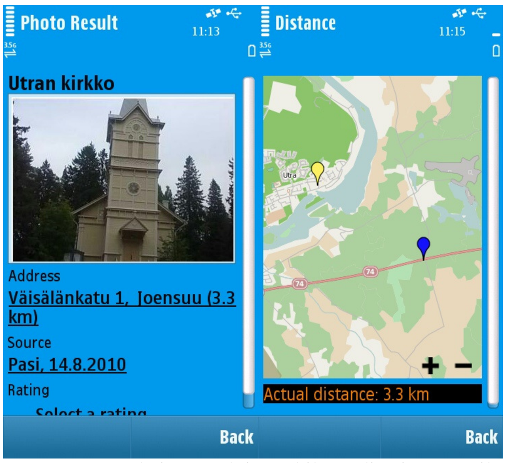

Figure 7: Recommendation result in mobile application. Details of the item are shown on the left. Item location on map is shown on the right.

In mobile application, the recommendation function is embedded in MOPSI Services screen (see Fig. 6) where user can request recommendation by pressing a single button. Application will show list of results including name, distance and type (service or photo) of each item (see Fig. 6). It is possible to see details of every result by selecting it on the list. Details of service results include title, street address, distance and list of keywords. Photo results include description user has entered (if available), photo thumbnail, street address, distance, author and date (see Fig. 7). By clicking on the address field, user can see map with item and its location marked.

IV. SCORING FUNCTION

Services are scored using contextual information about search history, location, and explicit rating. Photos are scored based on the same three factors and also on time. These factors are discussed next. We explicitly use two of the four aspects of relevance (location and time), whilst the other two factors (search history and rating) combine content, social network and time.

A. Search history

We define two search histories that are based on previous user behavior: *general* and *user-specific*. For services we take into consideration both the service name and the associated keywords, and for photos, we use the description which is assigned by the user.

The *general* history records keywords used for searches by all MOPSI users. It is used in three ways. Firstly, to check if any of the keywords associated to the service in question has been searched in nearby locations (S_N). Secondly, to check if any of the keywords has been searched recently (S_S). Thirdly, to check if the keyword has high frequency within all search requests (S_F).

User specific history records keywords that a given user has been used before for searches (S_U). Keywords of services and photos that are found in the history list are given 3 points each. For example, let us consider a service with keywords *café* and *restaurant,* and a user who has searched for *restaurant*, *bar*, *café* and *sauna* in the past. In total, this service gets 6 points for user specific history since two matched keywords were found.

Total score of search history consists of the following components:

$$S_H = S_N + S_S + S_F + S_U \tag{1}$$

where S_N, S_S, S_F, and S_U are the raw counts for keyword matches in nearby locations, within recent time, frequency of keywords in search history, and searches done by current user, respectively.

B. Location

We calculate the distances between each recommendation item and the user's location and define it as location score. By use of distance, we introduce location relevance aspect to the system.

C. Rating

Users can rate photo and services through the web or mobile interface. Services in MOPSI database have been rated by users in scale of 0 to 5 and the rating for photos is cumulative, using a thumbs up/thumbs down system (e.g. a photo liked by 5 users and disliked by 2 has a score of 3). The average score, in the case of services and the total score, in the case of photos, represents the rating score.

D. Time

Time is also a very important aspect of relevance. Photo relevance decreases with time, as the places or views capture by user may suffer changes over the years. Also, the season when the photo is taken is very important for the relevance, as winter activities, for example, cannot be recommended during summer.

More recent photos in user collection are considered more relevant than old ones and the newer the photo is, the higher score it receives. Additional difference is that the score is also influenced by time of the year when photo was taken.

We define following time thresholds (points given in brackets): 1 week (10), 1 month (7), and 1 year (4). Secondly, photos are classified into one of the four seasons of the year (winter, spring, summer, autumn). If the recommendation request is performed during the same season as the photo was taken, it is given additional 10 points. Thus, total score based on time is for each photo calculated as follows:

$$S_T = S_A + S_Y \tag{2}$$

For example, photos that were taken 4 days ago, and in the same day are scored S_A=10. The photo was also taken in the same time of year thus S_Y=10. By use of time for scoring photos, we introduce time relevance aspect to our system.

E. Total score

All the above scores are normalized to the scale [0..1] using the following formula:

$$N = \frac{S - MIN(S)}{MAX(S) - MIN(S)} \tag{3}$$

where S is the raw score, N is the normalized score, , and MIN(S) and MAX(S) are the minimum and maximum scores for each of the criterion respectively.

Final score of each service is then calculated using the following formula:

$$S_{SERVICE} = w_H N_H + w_L N_L + w_R N_R + 1 \tag{4}$$

where N_H stands for the normalized score for search history, N_L for location, and N_R for rating; w_H, w_L and w_R are weights for the corresponding scores. A constant of one point is added in order to promote services for recommendation, because they are assumed to originate from a trusted source and therefore more relevant than older photos from user collection. The location score is multiplied by two emphasize nearby locations.

Final score of each photo item is calculated in the same way as services, having an additional time score:

$$S_{PHOTO} = w_H N_H + w_L N_L + w_R N_R + w_T N_T \tag{5}$$

where N_T stands for time score and w_T is the weight for the time score.

Currently all weights are set to 1 except for the location weight which is set to 2 in order to emphasize importance of location.

V. PRELIMINARY EXPERIMENTS

As stated by [18] and [2] evaluation of user satisfaction is important part of the system evaluation. Moreover, evaluating user subjective opinion towards the system helps to improve various aspects of the system [18]. We conducted several experiments where we collected subjective opinions on recommended items. In the tests we have chosen several locations in Joensuu and analyzed usefulness of the recommendation there. Tests have been limited to Joensuu area, because user data collection is richest and most diversified in the town. There were locations with different characteristics selected. We checked whether the recommendation is relevant for different type of users such as tourists and locals. We also evaluated services and photos that were scored but not recommended in order to find out what relevant results may have been overlooked by the implemented scoring function.

Recommendations in downtown always give many cafeterias, pizzerias and bars taken from service database. In addition, the system provides additional recommendations such as sport places and shops taken from photo collection. Experiments have shown that all factors have impact on recommendation results. For example, in Papinkatu 3 (area with many block of flats and services nearby) there are many eating places recommended, but they are chosen based on rating and search history. Same example show well that our recommendation system chooses relevant photos from the collection, such as shop, kiosk and mailbox, whilst general photos of streets, houses and people are skipped, although located nearby.

In other districts of Joensuu service collection is smaller, and more of the recommendation results are photos. In some locations, such as Latolankatu 12, which is mostly a residential area, many photos in the very near proximity are selected although their content completely lacks relevance to most users. These photos are chosen due to lack of enough services and better photos available at the area. It also revealed that the location was given too much weight in this particular case, and it should be redefined more adaptively in future.

In other test cases, recommendations outside downtown area are mostly useful. For instance in Utrantie 75 (residential district with single houses and recreational areas in forest and on riverside) many recreational areas are found and recommended. Best bars from center of Joensuu are recommended in addition to the local bars and cafeterias nearby. This example also revealed that direct distance calculation is not sufficient as some results across the river are included despite there is no bridge nearby in this area to access these recommended services. Using navigational distance based on road network could help in this case. This example demonstrated the relevance of seasonal time score as photos taken in spring one year ago were selected instead of winter time photos taken 3 months ago.

VI. DISCUSSION AND CONCLUSIONS

In conclusion, our system gives useful recommendation and selects relevant items. There some exceptions where behavior of the system is not satisfactory. However, the system fails to give useful recommendation in specific, untypical cases, for example when user generated photo collection is very dense and limited to test photos with useless content or when there is no information about particular area in our data collection.

Although we already have useful results, there is still room for further improvements. First, user profile can be extended to include other factors besides the search history. In Facebook, people reveal personal data such as age and gender, which can be used for additional factors. For example, age can influence the choice of restaurant. Young people usually prefer fast food and older ones more expensive restaurants. Moreover, Facebook user's profile contains information about hobbies and favorite types of food. For example, people who like pizza should be given recommendation of nearby pizzerias instead of Chinese restaurants. Therefore, connection of user's profile in MOPSI with user profile in Facebook would give potentially much more information to build user profile.

Scoring function can also be improved further. Selection of items to be scored is the first area of improvement. Current solution with static distance threshold and maximum limit of items to be scored cannot adapt between areas with extremely low or extremely high density of items available in MOPSI data collection. In case of low density, there is small number of items scored (sometimes no items at all). In case of high density, some relevant items may not be chosen for scoring at all as happened with the Latolankatu example.

Furthermore, with growing number of items available in MOPSI we may need to increase speed of the recommendation algorithm. When MOPSI account will be integrated with Facebook, user's profile need to be systematically updated based on Facebook profiles. Weights for the scores should be adjusted in future, because experiments suggested that the importance of location was emphasized too much in some cases.

Calculation of time score should also be revised, however, it has to be done so that it preserves current behavior of the system in cases when time score improves recommendation results. Promoting services over user collection photos by using constant one is not always beneficial. It is worth to consider time score for services so that bars get more points in the evening, while lunch places in lunch time.

REFERENCES

[1] Adomavicius, G. & Tuzhilin, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transaction on Knowledge and Data Engineering 17(6), 734-749, June 2005.

[2] Albanese, M., Chianes, A., d'Acierno, A., Moscato, V., Picariello, A. A multimedia recommender integrating object features and user behavior. Multimedia Tools and Applications, 50(3), 563-585. December 2010.

[3] Birukov, E. Blanzieri, E.& Giorgini, P. 2005. Implicit: An agent-based recommendation system for web search. International Conference on Autonomous Agents and Multi-Agent Systems. Utrecht, The Netherlands, 2005.

[4] Burke, R. Hybrid Recommender Systems: Survey and Experiments. User Modelling and User-Adapted Interaction, 12(4), November 2002.

[5] Cao, L., Luo, J., Gallagher, A. A Worldwide Tourist Recommendation System Based on Geotagged Web Photos. IEEE International Conference on Acoustics Speech and Signal Processing. Dallas, Texas, USA, 2010

[6] Ducheneaut, N., Partridge, K., Huang, Q. Collaborative Filtering Is Not Enough? Experiments with a Mixed-Model Recommender for Leisure Activities. International Conference on User Modeling, Adaptation, and Personalization. Trento, Italy, 2009.

[7] Fränti, P., Chen, J., Tabarcea, A. Four Aspects of Relevance in Sharing Location-based Media: Content, Time, Location and Network. International Conference on Web Information Systems and Technologies. Noordwijkerhout, The Netherlands, May 2011.

[8] Fränti, P., Tabarcea, A., Kuittinen, J., Hautamäki, V. Location-based search engine for multimedia phones. IEEE International Conference on Multimedia & Expo. Singapore, July 2010.

[9] Ge, Y., Xiong, H., Tuzhilin, A., Xiao, K., Gruteser, M. & Pazzani, M. J. 2010. An Energy-Efficient Mobile Recommender System. ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington, DC, July 2010.

[10] Göksedef, M., Sule G.-Ö. Combination of Web page recommender systems. Expert Systems with Applications, 37(4), April 2010.

[11] Hinze, A., Quan, Q. Trust- and Location-Based Recommendations for Tourism. Confederated International Conferences. Vilamoura, Portugal, 2009.

[12] Horozov, T., Narasimhan, N., Vasudevan, V. Using location for personalized POI recommendation in mobile environments. International Symposium on Applications on Internet. 2006.

[13] Kang, E.-Y., Kim, H., Cho, J. Personalization Method for Tourist Point of Interest (POI) Recommendation. Knowledge-Based Intelligent Information and Engineering Systems. 2006.

[14] Kwon, O. B. "I know what you need to buy": context-aware multimedia-based recommendation system. Expert Systems with Applications, 25(3), 387-400. October 2003.

[15] Lee, B.-H., Kim, H.-N., Jung, J.-G., Jo, G. Location-based Service with Context Data for a Restaurant Recommendation. 2006.

[16] Lew, M. S., Sebe, N., Djeraba, Ch., Jain, R. Content-Based Multimedia Information Retrieval: State of the Art and Challenges. ACM Transactions on Multimedia Computing, Communications and Applications, 2(1). February 2006.

[17] Park, M.-H., Hong J.-H., Cho, S.-B. Location-based Recommender System Using Bayesian User's Preference Model in Mobile Devices. Ubiquitous Intelligence and Computing, 1130-1139. 2007.

[18] Pu, P., Chen, L. A User-Centric Evaluation Framework of Recommender Systems. ACM Conference on Recommender Systems. Barcelona, Spain, September 2010.

[19] Ricci, F., Mobile recommender systems, Information Technology & Tourism, Vol 12, No. 3, p. 205-231. 2011

[20] Takeuchi, Y., Sugimoto, M. CityVoyager: An Outdoor Recommendation System Based on User Location History. Ubiquitous Intelligence and Computing, 4159, 625-636. 2006.

[21] Tung, H.-W., Soo, V.-W. A Personalized Restaurant Recommender Agent Service for Mobile E-Service. IEEE International Conference on e-Technology, e-Commerce and e-Service. Taiwan, 2004.

[22] Yang, F., Wang Z.-M. A Mobile Location-based Information Recommendation System Based on GPS and WEB2.0 Services. WSEAS Transactions on Computers, 8(4), April 2009.