Association for Information Systems
AIS Electronic Library (AISeL)

PACIS 2004 Proceedings

Pacific Asia Conference on Information Systems (PACIS)

December 2004

Personal Recommendation in Mobile Environment

Wan-Shiou Yang
National Changhua University of Education

Jia-Ben Dia
National Changhua University of Education

Follow this and additional works at: http://aisel.aisnet.org/pacis2004

Personal Recommendation in Mobile Environment

Wan-Shiou Yang
Department of Information Management
National Changhua University of Education
wsyang@cc.ncue.edu.tw

Jia-Ben Dia
Department of Information Management
National Changhua University of Education
simon7870565@yahoo.com.tw

Abstract

With a fervent interest in the development of M-Commerce applications, we are engaged in a TAMC project, whose principal goal is to develop technologies for supporting targeted advertisement in mobile environment. The first investigation, the PR system, which aims at recommending vendors' web pages, which include offers and promotions, to interested customers, is reported here. Therefore, in this paper, we present the developed location-aware framework and constructed approaches. Simulated evaluation and implementation experience are also discussed.

Keywords: M-Commerce, Recommender systems, Data mining

1. Introduction

Mobile computing, where users provided with PDAs, cellphones or laptops are free to move while staying connected to service networks, has proven to be a true revolution (Stafford and Gillenson, 2003; Brunato and Battiti 2003). Exploring the promises of mobility to design new applications and services, which automatically accommodate customer's shopping needs with location dependent vendor offers and promotions, has generated considerable excitement among both practitioners (HP, www.hp.com; Nokia, www.nokia.com; Samsung, www.samsung.com) and academics (Sun, 2003; Tasasewish, 2003; Tewari et al., 2002; Brunato and Battiti 2003).

With a fervent interest in the development of such applications and services, we are engaged in TAMC (Targeted Marketing in M-Commerce) project. The project, whose principal goal is to develop technologies for supporting M-Commerce marketing services, is a marriage of three series investigations. The first investigation involved the design and construction of a personal recommendation system, termed as the *PR* system. The second investigation addresses the advertisement allocation problem, and the third investigation focus on the interface design in the context. The progress of the first investigation is reported in this paper.

The first investigation, the *PR* system, aims at recommending vendors' web pages, which include offers and promotions, to interested customers. The *PR* system adopted a location-aware framework so that customers can receive the information of their preferred vendors which are in their neighborhood. The core of the *PR* system is a recommendation mechanism, which carries out analysis of consumer's history and position so that venders' information are able to be ranked in order of the likelihood with which they match the preferences of a customer. Various characteristics of mobile environment are taken into account, and the endeavor has resulted in a recommendation system that is suitable in M-Commerce.

2. Literature review

In E-commerce, we have seen the emergence of many recommendation systems that intend to provide personal recommendation to various types of products and services, including news and emails (see http://www.netperceptions.com/ for a commercial site and (Goldberg, 1992; Lang, 1995; Konstan, 1997; Billsus, 1999) for research prototypes), Web pages (see http://my.yahoo.com/ for a commercial site and (Balabanovi'c, 1997; Terveen, 1997; Pazzani, 1997; Armstrong, 1997) for research prototypes), books (see http://www.amazon.com/ for a commercial site and (Mooney, 2000) for a research prototype), albums (see http://www.CDNow.com/ for a commercial site and (Shardanand, 1995) for a research prototype), and movies (see http://movies.eonline.com/ for a commercial site and (Alspector, 1998; Breese, 1998; Basu, 1998; Ansari, 2000; Pennock, 2000; Schafer, 2001) for research prototypes).

The first type of recommendation techniques was called the *content-based approach* (CACM, 1992). A content-based approach characterizes recommendable items by a set of content features and represents users' interest profile by a similar feature set. Then, the relevance of a given content item and the user's interest profile is measured as the similarity of this recommendable item to the user's interest profile. Content-based approaches select recommendable items that have a high degree of similarity to the user's interest profile.

Another type of recommendation technique, the *collaborative approach* (or sometimes called the *social-based approach*), takes into account the given user's interest profile and the profiles of other users with similar interests (Shardanand, 1995). Specifically, the collaborative approach looks for relevance among users by observing their ratings assigned to products in a training set of limited size. The "nearest–neighbor" users are those that exhibit the strongest relevance to the target user. These users then act as "recommendation partners" for the target user, and collaborative approaches recommend the target user items that appear in the profiles of these recommendation partners (but not in the target user's profile).

We observe that traditional recommendation techniques are not suitable in mobile environment. Firstly, both content-based and collaborative approaches require customers' provision of rating scores on selected items such that both positive and negative examples are available for analysis. For the immediately foreseeable future, the wireless devices typically used will be limited in input and output capability. Requiring customers to rate some products before making recommendation is not realistic. Secondly, M-Commerce is unique in its location-aware capability (Stafford and Gillenson 2003). Mobile computing adds a relevant but mostly unexplored piece of information- customer's position- to the recommendation problem. Personal recommendation in M-Commerce has the opportunity and necessity to take location into account. We therefore propose a location-aware framework for recommending venders' web pages in mobile environment.

This paper is structured as follows. The overall architecture of the *PR* system is described in Section 3. Detailed design and construction approaches are described in Section 4. Preliminary evaluation and implementation experience are discussed in Section 5. Finally, Section 6 summarizes this paper and points out our future directions.

3. System architecture

Generally, positioning systems fall into one of two categories. In *centralized architectures*, such as Active Badge (Harter and Hopper, 1994), Active Bats (Ward et al., 1997), and PARCTab (Want et al., 1996), the infrastructure consists of receivers deployed in some places, with end-users beaconing out data. Client's location is determined and held on servers. On the contrary, in *decentralized architectures*, such as Cricket (Priyantha et al., 2000) and RADAR (Bahl and Padmanabhan, 2000), the infrastructure consists of beacons deployed in some places, signaling to clients their locations. Hence, client's location is determined and held on a personal device.

There were numerous interviews (Barkhuus and Dey, 2003), reports (Weiser et al., 1999), and books (Garfinkel, 2001) describing people's unease over the potential for abuse of the privacy-sensitive location-aware systems. These concerns suggest that privacy may be the greatest barrier to adoption of location-aware services (Hong et al., 2003). Compared with centralized approach, decentralized approach gives end-users greater choice over whether to disclose their location data to others. Therefore, in this research, we adopt decentralized approach. Customers take control over their location data and only send out to the server as they need the recommendation service.

The overall architecture of the *PR* system is shown as Figure1. On the client side, there are two components. The first component is the off-the-shelf internet browser, and the second one is the location manager which estimates client's position. As customer needs the recommendation service, the request sends with the instant position to the server. On the server side, the core is the recommendation engine which consists of two components: off-line and on-line subsystems. The off-line subsystem maintains a database, WEB ACCESS, which logs data about what pages have been visited by each customer. The off-line subsystem also analyzes the logged data and derives each customer's interest profile. As receiving service request, the on-line subsystem generates a list of possibly interesting web pages by employing customer's interest profile and the instant position provided by the location manager.
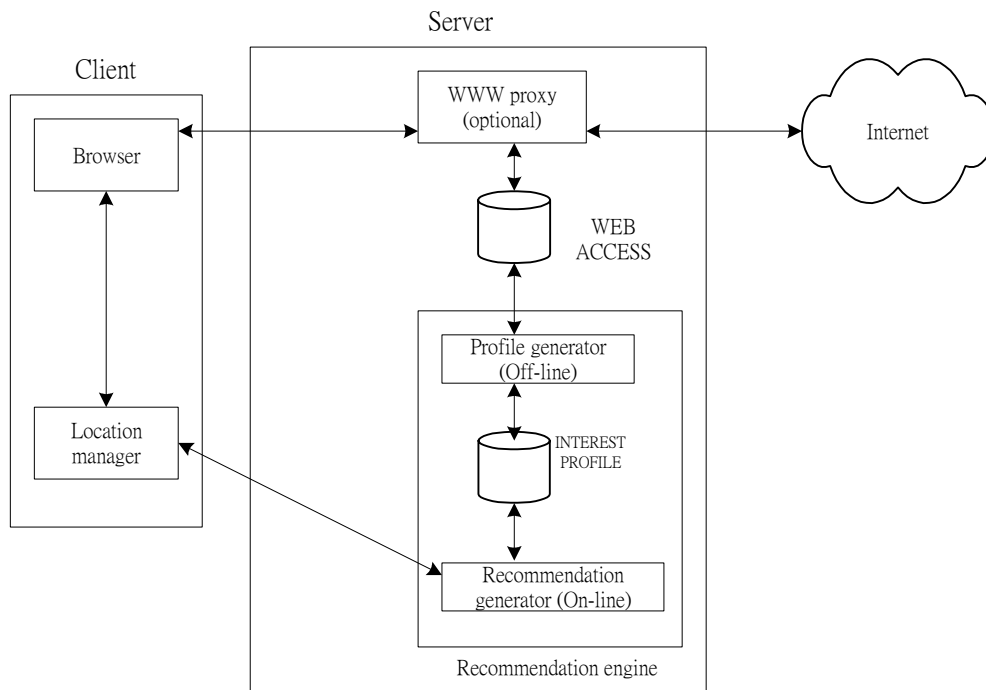
Figure1. Architecture of *PR* system

A screenshot of the client-side subsystem is shown as Figure2. After sending a service request, the customer receives a recommendation list containing links to various venders in the bottom frame. The positions of recommended venders are also marked in an e-map shown in the top frame to help the customer realizing venders' locations. Once a link in the bottom frame being clicked, the corresponding web page will be shown in the bottom frame.

Figure2. A screenshot of the client-side *PR* system

4. Proposed approaches

In this section, we first describe the off-line learning tasks conducted in the recommendation engine of the *PR* system, and then close with the on-line generation of a recommendation list.

Identification of interest profile

Once the WEB ACCESS database is populated with past customer accesses to pages, its data can be used to estimate customer's interest profile. The *PR* system collects the visited web pages and applies a simple information extraction method (Kushmerick, 1997) to it. Information extraction (IE) is the task of locating specific pieces of information from a text, thereby obtaining useful structured data from unstructured text. In our system, specifically, it involves parsing the raw pages, removing punctuations and prepositions, and grouping stemming words into generalized terms. The system adopts the bag-of-words model and creates a vector of terms for each web page, in which each cell indicates the frequency with which each term occurs in the page.

Each customer's interest profile, which is learned from the pages that the customer has visited, is also represented as a vector of terms. The system creates a vector of terms for each customer by summarizing the pages that he has visited. Therefore, suppose a customer c has visited a set S of web pages, the interest profile of customer c is estimated by Equ1, where Vector(w) denoted the term vector of page w.

$$CP(c) = \sum_{w \in S} \text{Vector}(w) \qquad\qquad (Equ1)$$

Generation of recommendation list

Based on the learned result (i.e., customer's interest profile), the *PR* system estimates customer's interest on vender's web pages. The interest on a given page is primarily measured as the similarity of this page to the customer's interest profile. The factor of distance between customer and vender is then considered. We posit that, for a customer, a farther vender is less likely to be visited with an exponential decay. Therefore, the interest of customer c on vender's page w at instant position is estimated by Equ2, where $\lambda \in [0,\infty]$ is a parameter and $\mathrm{Dis}(c, w)$ denotes the Euclidean distance between customer c and the web page w[1]. The *PR* system generates a recommendation list containing the top-N web pages to the requested customer.

$$\mathrm{Interest}(c, w) = \frac{\mathrm{Cosin}(\mathrm{CP}(c), \mathrm{Vector}(w))}{e^{\lambda \times \mathrm{Dis}(c,w)}} \qquad \text{(Equ2)}$$

It is noticed that λ is a parameter for representing customer's sensitivity to location. Clearly, if λ is small, customer's interest on a given page decrease mildly. In this case, it is more likely that a customer will choose a vender in a distant location. While in the case that λ is large, customer's interest will decrease dramatically, and it is less likely that customer will choose a distant vender. In order to find a desired λ for each customer, a λ learning procedure is also designed in this research. The λ learning procedure starts with a randomly chosen λ_0, and repeatedly makes gradual changes on it to find a desired parameter setting.

LMS (Least Mean Square) strategy is used to determine the magnitude of gradual change. Once customer c requests the service, a set S of recommended pages is generated. Suppose page $w \in S$ is the farthest page that customer c selected. We can define a disjoint partitioning of S as $S_S \cup S_{NS}$, where S_S is the set of the pages closer than w to customer c and S_{NS} is the rest set of pages. Take a review on the exponential decay function $f(x) = \dfrac{1}{e^x}$. For a given customer c and a given page w, the desired output of the function should be close to 1 if $w \in S_S$ (i.e., relatively near to customer c) and be close to 0 if $w \in S_{NS}$ (i.e., too far to customer c). Therefore, for estimated λ, the square error on comparing the estimated output and the desired output of the exponential decay function is $E = (e^{-\lambda \times Dis(c,w)} - d)^2$, so that

$$\frac{\partial E}{\partial \lambda} = -2(e^{-\lambda \times Dis(c,w)} - d) \times e^{-\lambda \times Dis(c,w)} \times Dis(c, w) \qquad \text{(Equ3)}$$

---

[1] Since the goal of the *PR* system is to accommodate customer's shopping needs with location dependent vendors, only pages created by location-dependent vendors are considered in this research. Accordingly, each page has a physical location coinciding with its vender.

The magnitude of gradual change in this research is hence computed as a negative multiple of

$\frac{\partial E}{\partial \lambda}$ shown as Equ4, where d is 1 if $w \in S_S$ or 0 if $w \in S_{NS}$.

$$\Delta \lambda = (e^{-\lambda \times Dis(c,w)} - d) \times e^{-\lambda \times Dis(c,w)} \times Dis(c,w) \qquad \text{(Equ4)}$$

The detailed pseudo code of the λ learning procedure is shown as Figure3. This procedure is implemented and embedded in client's location manager component. Each time, as the customer requests the recommendation service, the current λ is send with client's position to the server for generating a recommendation list. After receiving the generated pages, λ will then be recomputed according to customer's navigation behavior and held on client's device for next recommendation service.

λ-learning(λ: real; S: a set of pages): real
{
    Partition S into two sets S_S and S_{NS}
    For (each page w in S)
    {
        If ($w \in S_S$) then
            $\lambda = \lambda + (e^{-\lambda \times Dis(c,w)} - 1) \times e^{-\lambda \times Dis(c,w)} \times Dis(c,w)$
        Else
            $\lambda = \lambda + (e^{-\lambda \times Dis(c,w)} - 0) \times e^{-\lambda \times Dis(c,w)} \times Dis(c,w)$
    }
    Return λ
}

Figure3. The pseudo code of the λ learning procedure

5. Evaluation and discussion

We currently concentrate on the on-line workload of the *PR* system, and test it in a simulated environment. *K* sites, each of which associates with a web page, have been placed across a 10000m * 10000m square in random positions. We consider a service station in which customers, each of which associates with a random characteristic value and a vector of terms, arrive in accordance with a Poisson process with rate β. Upon arrival, a customer either enters service if the recommendation engine is free at that moment or else joins the waiting queue. When the engine generates a recommendation list for a customer it then either begins serving the customer that had been waiting the longest if there are any waiting customers, or,

if there are no waiting customers, it remains free until the next customer's arrival. In addition, there is a fixed time- one hour- after which no additional arrivals are allowed to enter the system, although the server completes servicing all those that are already in the system.

We first investigated the effects of the number of the web sites, ranging from 10000 to 50000 at 10000 increments, on the average servicing time of request. 10 experiments are executed and the average result is shown as Figure4(a). In general, the average servicing time grows linearly as the number of sites increases. It is expected since the *PR* system requires scanning the whole web page database for each service request to build a recommendation list. Also, we investigated the joint effects of the number of the web sites and the request arrival rate β, ranging from 2 to 10 at 2 increments, on the average time a customer spends in the system (i.e., including servicing time and waiting time). 10 experiments are executed and the average result is shown as Figure4(b). In the left-bottom area of Figure 4(b), the average time a customer spends in the system grows mildly. However, as the servicing time exceeds the time period between two requests (i.e., the inverse of the arrival rate β), as data shown in the right-top area, the average time a customer spends in the system grows dramatically. This is because a great number of requests are accumulated in the waiting queue, and accordingly result in the huge increase of the waiting time.
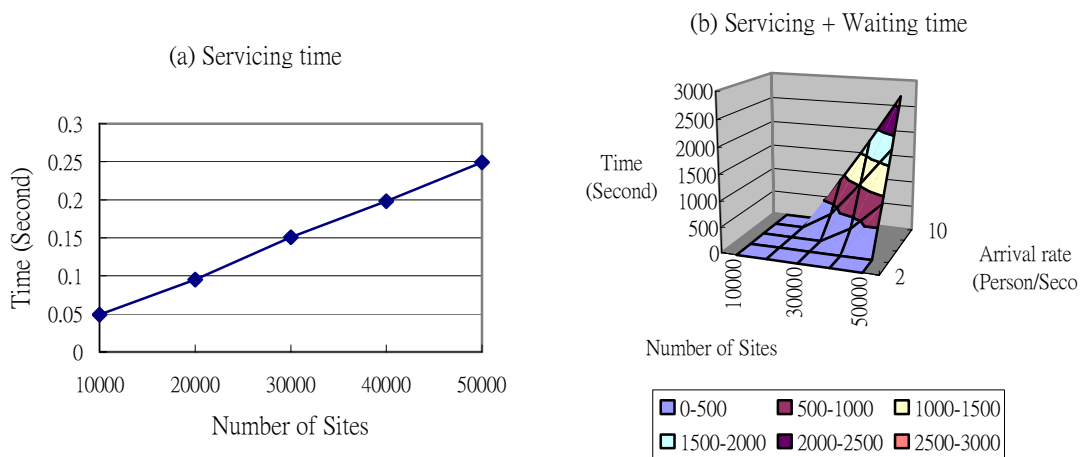

Figure4. The execution time of the experimental *PR* system

From the simulated evaluation, it is clear that the scalability problem is critical for the success of location-aware recommendation system. Customers often expect to receive the generated recommendation list in a real time manner, and thus it is necessary for the system to deal with the problem of efficiently finding interesting pages, particularly in the heavy traffic case. We therefore recognize this scalability problem worthy for further research in our project. It is

conceived to use some locality-based data structures and searching strategies to ease the work of scanning the page database.

Also, a real-world pilot study is performing in our project. We have five subjects; all of them are undergraduate students. The involved subjects are shown a brief, two-minute demonstration of the system; then they are provided with PDAs implemented the system; and they are asked to use the device for three months. At the first month, only the browser component is activated to log what pages have been visited by the subject. At the second and third months, all components are activated to provide the recommendation service. After this pilot study, we will involve more subjects with various demographic characteristics to test the effectiveness of our system.

Also, though the decentralized approach we adopted is considered as a more flexible privacy mechanism, subjects still exposed their privacy concerns in the pilot study. Mostly, they concerned that, though location is calculated locally, in the connected case, the information is still transmitted through an access point, across a network service provider, and to the recommendation service provider. There is still possibility that location data be revealed without control. Therefore, studying some privacy-preserving schemes and techniques to integrate into our architecture and hence improve the privacy level of our system is another direction of our research.

6. Conclusion

In this paper, we have addressed the issue of location-aware personal recommendation. This paper reflects our effort to formulate a system that amalgamates the information abundance of the internet with the tangible richness of physical shopping. We have discussed how the system being conducted to exploit the functionality afforded by a powerful location-aware architecture. Using the simulated model, we also estimated the efficiency of our current system. Hereafter, we will investigate the scalability and privacy problems as well as more extensive real-world tests in our project.

Reference

Alspector, J., Kolcz, A., and Karunanithi, N., "Comparing Feature-based and Clique-based User Models for Movie Selection," *Proceedings of International Conference on Digital Libraries*, 1998.

Ansari, A., Essegaier, S., and Kohli, R., "Internet Recommendation Systems," *Journal of Marketing Research*, 37(3), 2000.

Armstrong, R., Freitag, D., Joachims, T., and Mitchell, T., "WebWatcher: A Tour Guide for the World Wide Web," *Proceedings of International Joint Conference on Artificial*

Intelligence, 1997.

Bahl, P., and Padmanabhan, V.N., "RADAR: An In-Building RF-Based User Location and Tracking System," *IEEE INFOCOM*, 2000.

Balabanovi'c, M., and Shoham. Y., "Fab: Content-Based, Collaborative Recommendation," *Communications of the ACM*, 40(3), 1997.

Barkhuus, L., and Dey, A.K., "Location-based services for mobile telephony: a study of user's privacy concerns," *Proceedings of the International Conference on Human-Computer Interaction*, 2003.

Basu, C., Hirsh, H., and Cohen, W., "Recommendation as Classification: Using Social and Content-Based Information in Recommendation," *Proceedings of AAAI Symposium on Machine Learning in Information Access*, 1998.

Billsus, and Pazzani, "A hybrid user model for news story classification," *Proceedings of International Conference on User Modeling*, 1999.

Breese, J., Heckerman, D., and Kadie, C., "Empirical analysis of predictive algorithms for collaborative filtering," *Tech. Report, MSR-TR-98-12*, Microsoft Research, 1998.

Brunato, M., and Battiti, R., "PILGRIM: A Location Broker and Mobility-Aware Recommendation System," *Proceedings of International Conference on Pervasive Computing and Communications*, 2003.

(1992), "Special issue on information filtering," *Communications of the ACM*, 35(12), 1992.
Garfinkel, S., "Database Notion: The Death of Privacy in the 21st Century," *O'Reilly & Associates*, 2001.

Goldberg, D., Nichols, D., Oki, B., and Terry, D., "Using collaborative filtering to weave an information tapestry," *Communications of the ACM*, 35(12), 1992.

Hater, A., and Hopper, A., "A Distributed Location System for the Active Office," *IEEE Network*, 8(1), 1994.

Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L, and Riedl, J., "GroupLens: Applying Collaborative Filtering to Usenet News," *Communications of the ACM*, 40(3), 1997.

Kushmerick, N., Weld, D., and Doorenbos, B., "Wrapper Induction for Information Extraction," *Proceedings of International Joint Conference on Artificial Intelligence*, 1997.

Lang, K., "Newsweeder: Learning to Filter Netnews," *Proeedings of International Conference on Machine Learning.*, 1995.

Mooney, R., and Roy, L., "Content-Based Book Recommending Using Learning for Text Categorization," *Proceedings of International Conference on Digital Libraries*, 2000.

Pazzani, M., "A Framework for Collaborative, Content-based and Demographic Filtering," *Artificial Intelligence Review*, 13(5-6), 1999.

Pennock, D., Horvitz, E., Lawrence, S., and Giles, C., "Collaborative Filtering by Personality Diagnosis: A Hybrid Memory- and Model-Based Approach," *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, 2000.

Priyantha, N.B., Chakraborty, A., and Balakrishnan, H., "The Cricket Location-Support System," *Proceedings of the International Conference on Mobile Computing and Networking*, 2000.

Schafer, J., Konstan, J., and Riedl, J., "E-Commerce Recommendation Applications," *Data Mining and Knowledge Discovery*, 5(1), 2001.

Shardanand, U., and Maes, P., "Social Information Filtering: Algorithms for Automating 'Word of Mouth'," *Proceedings of International Conference on Human Factors in Computing Systems*, 1995.

Stafford, T. F., and Gillenson, M. L., "Mobile Commerce: What It Is and What It Could Be," *Communications of the ACM*, 46(12), 2003.

Sun, J., "Information Requirement Elicitation in Mobile Commerce," *Communications of the ACM*, 46(12), 2003.

Terveen, L., Hill, W., Amento, B., McDonald, D., and Creter, J., "PHOAKS: A System for Sharing Recommendations," *Communications of the ACM*, 40(3), 1997.

Tewari, G., Youll, J., and Maes, P., "Personalized location-based brokering using an agent-based intermediary architecture," *Decision Support Systems*, 34(12), 2002.
Wan, R., "The PARCTAB Ubiquitous Computing Experiment," *Mobile Computing*, 1996.

Ward, A., Jones, A., and Hopper, A., "A New Location Technique for the Active Office," *IEEE Personnel Communications*, 4(5), 1997.

Weiser, M., Gold, R., and Brown, J.S., "The Origins of Ubiquitous Computing Research at PARC in the Late 1980," *IBM Systems Journal*, 38(4), 1999.