Personalized Mobile Services Using Weighted Instance Based Learner for User Profiling

Ayse Cufoglu, Mahi Lohi
Department of Electronics and Computer Science
University of Westminster
London, UK
a.cufoglu@ieee.org,lohim@wmin.ac.uk

Colin Everiss
Department of Business Information Systems
University of Westminster
London, UK
C.G.Everiss@westminster.ac.uk

***Abstract*—Today, mobile device users receive a variety of services and information delivered to their mobile devices. Many of these are irrelevant, far from the user's satisfaction level and may likely be regarded as spam messages by the user. This results in the users to look for the relevant services by themselves, which would be time consuming and may cause inconvenience. User profiling has created opportunities for mobile service providers to provide a channel for user awareness as well as to achieve high user satisfaction. Apart from traditional collaborative and content-based methods, a number of classification and clustering algorithms have been used for user profiling. Instance Based Learner is a comprehensive form of the Nearest Neighbour algorithm and it is suitable for user profiling as users with similar profiles are likely to share similar personal interests and preferences. In our previous work we proposed a weighted classification method, namely Weighted Instance Based Learner, and evaluate its performance on user profiling. According to the simulation results Weighted Instance Based Learner, performs better than Instance Based Learner, on user profiling by reducing the error up to 28% on the selected dataset. In this paper, we aim to demonstrate how Weighted Instance Based Learner algorithm can be used for the user profiling for the provisioning of personalized mobile services. For this purpose a scenario has been proposed and implemented as a Java Mobile Application on NetBeans IDE 7.1.**

Keywords-User Profiling; Personalization; Machine Learning.

I. INTRODUCTION

Many works in the literature show that mobile recommendations have become very popular due to the growing diversity, availability and use of mobile information services [1]. Personalization of the mobile services is therefore an opportunity to help to improve the quality of service. Personalized mobile services aim to match users' requirements by considering when, where and how the users require the service to be delivered. The success of these applications relies on how well the service provider knows the user requirements and how well this can be satisfied. The user profile is the representation of the user and holds information about the user such as personal profile data (demographic profile data), interest profile data and preference profile data. These profiles are the outcome of the user profiling. In user profiling applications the major challenge is to build and handle user profiles. In the literature two fundamental user profiling methods have been proposed for this purpose. These are the collaborative and the content-based methods. It is also possible to use a hybrid of the two methods. The collaborative method has been built on the assumption that similar users, with respect to the age, sex, and social class, behave similarly, and therefore have similar profiles [2]. The content-based method, on the other hand, has been built on the concept of content similarity and assumes that users behave similarly under the same circumstances [2]. For example, the moreTourism, mobile recommendations for tourism [3], uses a hybrid method. The proposed recommendation system here takes into account the tags, provided by the users, to provide tourist information profiled for users with similar likes depending on the user profile (user tag cloud), location in time and space, and the nearby context. For example, nearby historical places and museums. Similarly in [4], Fernandez et al., proposed a tourism recommender system that offers tourist packages (i.e., include tourist attractions and activities) that best matches the user's social network profiles. Different from [3], the proposed hybrid system provides recommendations based on both the user's viewing histories (Digital Television (DTV) viewing histories received from the user's set-top boxes via a 2.5/3G communication network) and the preferences in the social network (i.e., preferences of the user's friends). Apart from the traditional profiling methods, a number of classification and clustering algorithms found applications within the user profiling process in personalization. This paper aims to show how the Weighted Instance Based Learner (WIBL) [5] classification algorithm can be used for the user profiling to provide personalized mobile services.

The rest of this paper is organized as follows: Section II provides information about the WIBL. Section III presents the scenario for this paper, while Section IV outlines a detailed overview of the system. Implementation of the scenario, is given in Section V. Finally Section VI concludes this paper.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-232-5

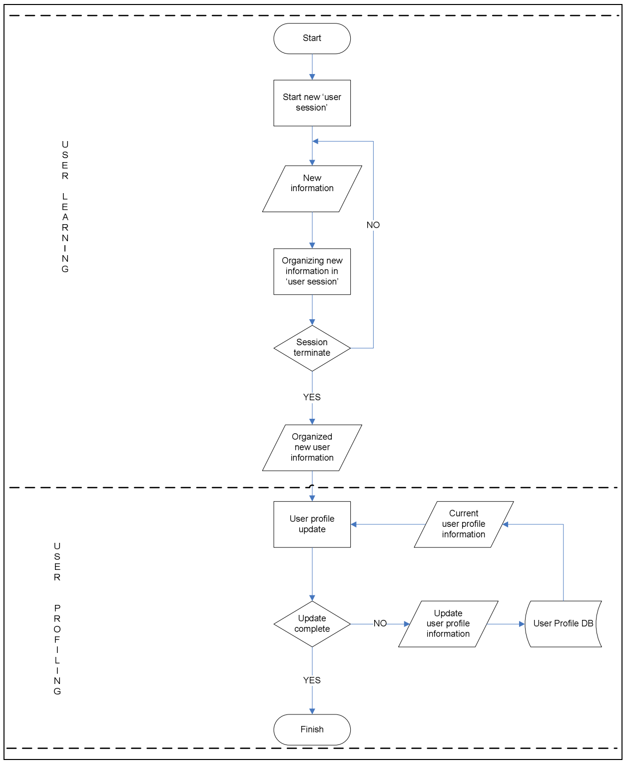

Figure 1. Flowchart of the user learning and user profiling.

II. Weighted Instance Based Learner (WIBL)

Instance Based Learner (IBL) is a comprehensive form of the Nearest Neighbour (NN) algorithm; which normalizes its attributes' ranges, processes instances incrementally and has a simple policy for tolerating missing values [6]. In contrast to IBL, the WIBL assigns weights to the attributes and considers the weighted distance of the instances for classification. Here, relevant attributes are aimed to have more influence on classification than irrelevant attributes. In WIBL the function that calculates the distance between test instance (new user) X_i and the training instance (existing user) Y_j is;

$$dist(X_i, Y_j) = \sqrt{\sum_{k=1}^{A} w_{k,l}(C_m)\ g(x_i(k), y_j(k))}\ , \qquad (1)$$

where

$$w_{k,l}(C_m) = P(C_m | f_k(l))\ . \qquad (2)$$

Here, l is equal to the value of the $x_i(k)$. Therefore, the selection of which weight is to be used for a particular attribute value is based on k and $x_i(k)$. Note that $g(x_i(k), y_j(k))$ is evaluated as it is in IBL [6]. In [5], presented simulation results illustrated that WIBL performs better than IBL on user profiling by reducing the error up to 28% on the selected dataset. For the simulations the dataset used was provided in [7], named 'Adult Data Set'. This dataset was created by Barry Becker via extracting information from the 1994 census database. From this dataset demographic information of 6000 instances (5000 training and 1000 test) have been adopted and used to create a complete dataset of user profiles for the simulations.

The WIBL enables the use of a user profile dataset of a single service with other services too. This is archived by feature weighting. Although the content of the dataset stays the same, the weighting values will be recalculated/updated when the dataset is used for another service.

III. Proposed Scenario

In this scenario, we focus on a mobile advertising service. Here, we introduce a personalized mobile advertising service called Discounts, Promotions and Deals (DPD). DPD advertising service provides discount, promotion and deal advertisements to the user according to the user's profile. Furthermore, for this scenario, DPD is concerned with the food industry, and a restaurant service called MyRestaurants, has been chosen. MyRestaurants can be used by the registered users only. The following user is assumed for this scenario.

Ren is a 30 years old Londoner. She is working as a property adviser in a company located in central London. She has got both an iPhone and a BlackBerry Smartphone, which have been provided by the company. She uses her BlackBerry for work related duties while her other mobile phone is a part of her personal life. Ren decided to subscribe for the personalized mobile advertising service, MyRestaurants. Through her mobile device, each of the advertisements are presented with the link where a user can follow for more information. Ren prefers to receive the advertisements everyday and prefers to check these out in the morning time. Subsequently, on Monday morning, around 9am on her way to work, Ren signs into the MyRestaurants service through her iPhone. She receives the advertisements. She is pleased with the one of a meal deal offer as the restaurant is very close to her work place and she has previously thought about trying out its food. Ren follows the provided link to book a table through the restaurant's mobile-web.

IV. System Overview

The following three subsections explain the user learning, user profiling and restaurant recommendation for this scenario. Figure 1 shows the flowchart of the user learning and profiling. User learning process starts whenever the user signs into to the MyRestaurants. Here, the system monitors user's feedback towards the given recommendations until user signs out from the system (i.e., session terminates). Following this, the new information from the learning process

53	<10	Male	Football	Rock	Political	Nature-Trekking	Married-civ-spouse	Private	11^{th}	Husband	Handlers-cleaners

Kusak Restaurant	Turkish	<30	order of 3-course meal for two comes with a free bottle of wine	New Cavendish Street London, W1W 6UW

Figure 3. Example of user profile information and restaurant profile information.

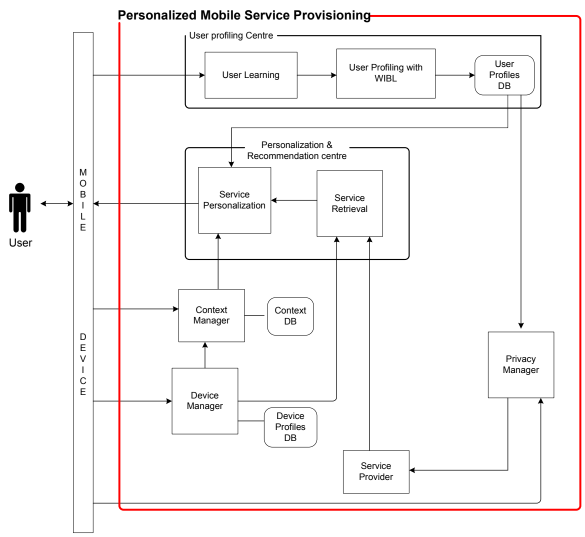

Figure 2. Architecture of personalized mobile service provisioning.

is used for the user profiling. In this process, a classification algorithm WIBL will update the user's profile information in the user profile dataset with using the information from user learning process. Following subsections give more detailed information on the aforementioned processes. Moreover, it was considered that an investigation of the user privacy issues, device management and context management is out of the scope of this paper.

A. User Learning

For this scenario, we assume that the information given by the user during the subscription is to be used for the initialization of the user's profile. Note that this corresponds to the directly/explicitly information gathering. The user's response (user feedback) to the provided services will then be used to update the user's profile implicitly. Each user is represented with three sets of user profile information that are interest profile, demographic profile and preference profile. It is worth pointing out that the location preference of the user will be kept in the user profile. Each user will have an identification (i.e., user-id and password) for the purpose of authentication for the service. Here, the system will automatically assign user a user-id and a password when she subscribes for the service. An initial password can be changed by the user following first sign in. After subscription and registration, the system continuously monitors user's feedback and behaviour towards the provided services to learn more about her (i.e., what services she likes, when and where). For example, monitoring Ren shows that she prefers to receive the advertisements every morning while travelling to work.

Figure 4. User enters her user-id and password to sing-in.

B. User Profiling

For this scenario, the WIBL [5] is used for the user profiling. Here, WIBL will assign different weights to the user's user profile attributes to increase the impact of relevant attributes in classification so as to define the user's service preferences more precisely. Which user receives what advertisements is decided by making use of the user's profile information and the class that the user belongs to. In this way, the same advertisements can be sent to the users that share the same class and these users receive the advertisements that most of the users in the same class showed a liking for.

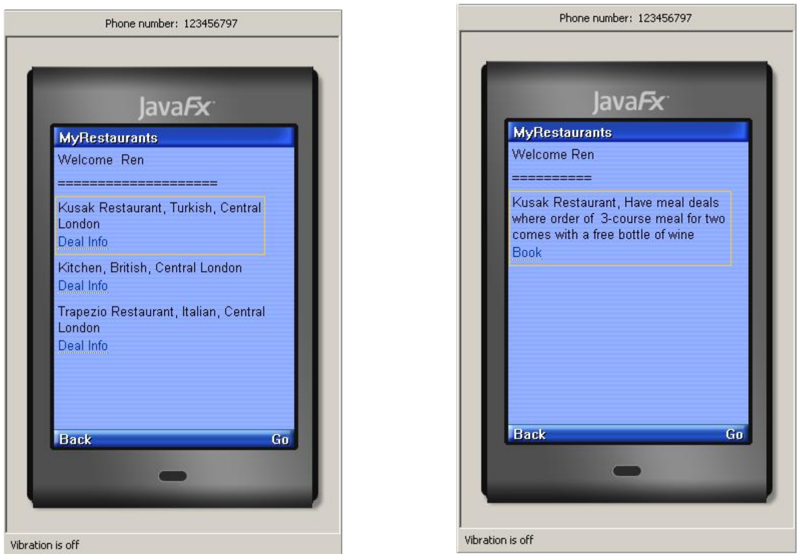

Figure 5. User's daily restaurant deals and detailed deal information.

User's location preference and user's current location are two important parameters for providing the right location based advertisements. For example, when it comes to the location based advertisements, Ren prefers the ones that are close to her work place so her location preference is 'work'. However, she is a property adviser and she needs to travel to different UK cities very often. Hence, when Ren is away, she will receive location based advertisements based on her user profile information and current location rather than her preferred location. The current location can be extracted from the GPS (Global Positioning System) information of the user's mobile device.

C. Restaurant Recommendation

For personalized mobile services, various architectures have been proposed [1][8][9][11]. Here, personalized restaurant recommendations are the outcome of the personalization process. In this scenario, the personalization process uses user profile information to personalize (filter) the restaurants to be recommended to the user. In Figure 2, detailed information of this process is shown. From this figure it can be seen that there are three inputs to the 'service personalization': user profile, service to be personalized, and current context information. Context information (i.e., location) and device capabilities are obtained from the mobile device. These are considered to be important for accurate user interface adaptation and personalization. Here, a privacy manager uses user's sign in information and user profile information to decide who can use the user profile information for what purpose, who the user is and if they have the right to use the provided service. It is worth pointing out that, like each user, each restaurant has to subscribe to MyRestaurants to be recommended to the users. This means that service provider acts like a bridge between users and restaurants.

Figure 3 is an example of some of the demographic, interest and preference information of a user in user profile with the following order: Age, Annual Income, Sex, Sport Interest, Music Interest, Book Interest, Leisure, Marital Status, Employment, Education and Profession. The WIBL uses this given data to predict the user's cuisine preferences. Here, user's cuisine preference is represented with its probabilistic distribution function, which enables the user to receive recommendations from different types of restaurants. These probabilities can change, based on the users feedback to the given recommendations. In this study, the user's clicks on the given recommendation is considered as a positive feedback. Here, the system counts each click on recommended restaurants and utilizes this information to update the user's cuisine preferences. Therefore, user's current information and new information are incorporated together to update the user profile information.

As mentioned previously, the user's location preference information (home, work or elsewhere) is also kept in the user profile and used for the location based restaurant recommendations. Gasson et al. [10] shows what kind of personal information can be obtained by monitoring a user's mobile device while in [1] it has been shown how the GPS data can be converted into text format. This method makes it possible to compare restaurants' location and user's location preference (or user's current location in case of elsewhere) to provide accurate recommendations. Similar to the user profile dataset, restaurant information is kept in the restaurants dataset. Figure 3 shows an example for the restaurant profile information. In this separate dataset, each restaurant is represented with the following attributes:

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-232-5

Name, Cuisine Type, Price, Deal Description and Location. Here, each of these are used to classify restaurants based on their cuisine types using IBL [6].

V. Implementation of the Proposed Scenario

This section implements the proposed scenario and shows the usage of a DPD-Restaurant application, named MyRestaurants, from the user's point of view. The scenario is implemented as a Java Mobile Application (Java ME) on NetBeans IDE 7.1. Note that for this scenario we assumed that user Ren is already subscribed for the service. Following her subscription, Ren started using the service. To check her restaurant recommendations she needs to sign into the system using her user-id and password (see Figure 4). Here, prompt information is compared with the information in the user's profile for authentication. Ren's successful sign in redirects her to the MyRestaurants main page. This main page displays two options: 'My Account' and 'My Deals'. The first option, 'My Account', redirects her to a new page where she can change her password, location preference and user-name. The user-name is different from the user-id and it is used for display purposes. In this scenario the user prefers her user name to be 'Ren'. 'My Deals', on the other hand, redirects her to a new page. This new page includes daily restaurant recommendations (see Figure 5). Each recommendation has a link, which provides more information about the deal and the restaurant (also see Figure 5). Here, if she wants, she can follow another provided link to make a booking.

VI. Conclusion and Future Work

In this paper, a real life scenario for personalized mobile services is presented. Here, the aim was to illustrate how WIBL classifier can be used for user profiling in mobile environments so as to provide personalized mobile services.

The IBL classifier is a comprehensive form of the NN algorithm. WIBL is a modified version of the IBL, where weights were assigned to the attribute values and classification of the users are done based on the weighted distances. In our previous work we showed that WIBL decreases the error rate of IBL up to 28% and therefore increases the classification accuracy performance on user profiling.

The proposed scenario provided for here, focuses on the mobile advertising service, namely MyRestaurants. This scenario is implemented as a Java ME on NetBeans IDE 7.1. Here, each user is represented with three sets of user profile information that are interest profile, demographic profile and preference profile. Similarly, restaurant information is also kept in the restaurants dataset. Detailed information on both datasets together with user learning, user profiling, restaurant recommendation and the implementation of the scenario have also been provided in this paper. Based on the work carried out here, it can be conclude the WIBL for user profiling can be used to provide more accurate personalized mobile services. In our future work, we would like to test our proposed scenario with real users. It will be also interesting to compare the performance of WIBL with different personalized mobile services.

References

[1] K. F. Yeung and Y. Yang. A proactive personalized mobile news recommendation system, IEEE Developments in E-systems Engineering, pp.207-212, 2010.

[2] G. Araniti, P. D. Meo, A. Iera, and D. Ursino. Adaptive controlling the QoS of multimedia wireless applications through user profiling techniques, IEEE Journal on Selected Areas in Communication, 21(10), pp.1546-1556, 2003.

[3] M. R. Lopez, A. B. B. Martinez, A. Peleteiro, F. A. M. Fonte, and J. C. Burguillo. MoreTourism:mobile recommendations for tourism, IEEE International Conference on Consumer Electronics, pp.347-348, 2011.

[4] Y. B. Fernandez, M. L. Nores, J. J. P. Arias, J. G. Duque, and M.I.M. Vicente. TripFromTV+:Exploiting social networks to arrange cut-price touristic packages, IEEE International Conference on Costumer Electronics, pp.223-224, 2011.

[5] A. Cufoglu, M. Lohi, and C. Everiss. Weighted instance based learner (WIBL) for user profiling, IEEE International Symposium on Applied Machine Intelligence and Informatics, pp.201-205, 2012.

[6] D. W. Aha, D. Kibler, and M. K. Albert. Instance-based learning algorithms, Machine Learning journal, 1(6), pp.37-66, 1991.

[7] A. Asuncian and D.J. Newman (Accessed: 2 July 2012), UCI Machine Learning Repository, Irvine, CA: University of California, School of Information and Computer Science, 2007.

[8] S. Loeb and E. Panagos. Information filtering and personalization: context, serendipity and group profile effects, Annual IEEE Consumer Communications and Networking Conference-Emerging and Innovative Consumer Technologies and Applications, pp.393-398, 2011.

[9] T. D. Pessemier, T. Deryckere, K. Vanhecke, and L. Martens. Proposed architecture and algorithm for personalized advertising on iDTV and mobile devices, IEEE Transactions on Consumer Electronics, 54(2),pp.709-713, 2008.

[10] M. N. Gasson, E. Kosta, D. Royer, M. Meints, and K. Warwick. Normality mining: privacy implications of behavioural profiles drawn from GPS enabled mobile phones, IEEE Transactions on Systems, Man, and cybernetics-Part C: Applications and Reviews, 41(2), pp.251-261, 2011.

[11] B. Mrohs and S. Steglich. Architectures of future services and applications for mobile user, IEEE Symposium on Applications and the Internet Workshops, pp.128-131, 2005.