Towards a Context-Aware Photo Recommender System

Fabrício D. A. Lemos[1 2], Rafael A. F Carmo [3], Windson Viana[1 3], Rossana M. C. Andrade[1 2 *]
[1] Group of Computer Networks, Software Engineering and Systems (GREat)
[2] Master and Doctorate in Computer Science (MDCC)
[3] UFC Virtual Institute (IUFCV)
Federal University of Ceará (UFC)
{fabriciolemos, rossana}@great.ufc.br
{rafael, windson}@virtual.ufc.br

ABSTRACT
The main challenge of recommender systems is to be able to identify and recommend items that have a greater chance of meeting the interests of their users, which generally have a very subjective and heterogeneous nature. It is imperative, then, that recommender systems, from the identification of each user's profile, could recommend personalized items. However, the user's profile is not enough for the system to be able to completely identify the user's interests. The use of the system in a different context from the usual may cause an unsatisfactory result for the recommendation, requiring it to be adapted to a new context. This paper presents the MMedia2U, a prototype of a mobile photo recommender system that exploits the user's context and the context when the photo was created as a means to improve the recommendation. Three context dimensions area exploited: spatial, social and temporal. We describe the similarity measures used for each dimension and the results of the system evaluation by 13 users following a Gold Standard approach.

Keywords
Recommender Systems, Context-Awareness, Ubiquitous Computing.

1. INTRODUCTION
One of the most important challenges in Information Systems is *information overload*. Recommender Systems try to cope with this problem by helping people in retrieving information (ex: videos, TV programs, routes, images, people, etc.) that may match their preferences and intentions. Recommender Systems try to identify items, from the information corpus, that have a greater chance to meet the wishes of its users [4][6]. However, the characterization of user's preferences and intentions is a complex task that, as rule, needs user intervention to be fulfilled correctly. Another issue of Recommender Systems is related to user's context. The use of the system in a different context than usual may cause an unsatisfactory result for the recommendation, since preferences and intentions can be influenced by the user's context (location, trajectory, time, activity, etc.). Context-awareness refers exactly to the capacity that a system has to detect user's situation and guide the system behaviour accordingly [5]. Nowadays, mobile devices improvements and home sensors technologies allow a better user's context characterization and this information can be useful to improve recommendations.

CARS-2012, September 9, 2012, Dublin, Ireland.

In this scenario, this paper presents the Mobile Media to You (MMedia2U), a photo recommender system that suggests images previously annotated with contextual information. In order to execute the recommendation, MMedia2U explores the current context of the user acquired by his/her mobile device. This domain is interesting since there are available a large number of images on sharing sites such as *Flickr*[1] and *Picasa Web*[2]. Many of these pictures are captured by mobile devices that store the location, date, time, and other contextual information which can be explored for recommendation.

2. Context-Aware Recommender Systems (CARS)
With the dissemination of ubiquitous computing concepts, context-awareness has become a very important research field. Its ideas have been used to increase efficiency and usability of Information Systems, particularly, those systems accessed from mobile devices [5]. Context-awareness in recommender system has been motivated from research, which recognizes the dependence of user long-term needs on time, location, and any information about the physical environment surrounding the user [6][10]. Context-awareness introduces an additional level of personalization since it takes into account the influence of the external environment of the user on his/her appreciation of the products or items. Recommender systems can take benefits from the context-awareness by considering not only the characteristics intrinsic to each item and user, but also the characteristics of their current situations (both user and item). For example, gathering context, a restaurant recommender system will be able to adapt its recommendations for restaurants that are next to the user, open and that have available seats for the amount of people who are with the user (e.g., his/her family).

In the last years, some researchers have showed the feasibility of implementing CARS (e.g., news, movies, music, and services). Adomavicius et al. [6], for instance, implement a recommender system of movies that takes into account the user context (e.g., if the user is going to watch the movie at home or in the movie theatre) and it has attested improvements in the precision and recall of recommendations. The authors also propose a classification of context-aware systems into two categories. The **first category** contains systems that use contextual information as

[1] www.flickr.com/

[2] http://picasaweb.google.com/lh/explore

*Sponsored by CNPq under Productivity Scholarship DT-2

criterion for filtering items. For instance, COMPASS [1] is a tourism guide that recommends Points of Interest (POI) taking into account the current context of the user. Kaialeido Photo [2], in turn, performs contextual annotation of images and allows the user to filter the albums according to preset categories (e.g., an event where the photo was created). In this system, the user has to explicitly specify the filters to be used. Columbus [11] is simpler; it is a mobile application that displays georeferenced photos taken near the user's location.

The **second category** comprises systems that use the contextual information at the time that the user evaluates an item. In addition to the ratings and characteristics of items and users, systems of this category also take into account contextual information (e.g., location) during recommendation. The work described by Adomavicius et al. [6] is an example of a recommender system in this category. The system records time that a user watched a movie, stores both the score given by the user to the film and the context in which the movie was watched (e.g., at home, with his wife). Thus, for example, a film that was well rated by users in a given context, are more likely to be recommended to a user who is in a similar context. In the domain of multimedia content, the system C2_Music [7] incorporates contextual information on music recommendation. In general, the behaviour of this system is similar to the movie recommender system, in which the song evaluation is enriched with the context in which it was heard (e.g., day of week, weather conditions).

Such systems depend, however, on a historical database describing which items were evaluated by others users (or the user himself/herself) in similar contexts to the current user's context. Another problem occurs when a new item is added to the collection. As this item has not yet been used, it will be difficult to recommend it in accordance with recommendation techniques such as collaborative filtering (so-called cold start problem). The work described in this paper differs from the studies aforementioned since it takes into account both the context of users and the context in which the items were created. The hypothesis is that photos taken in a given context *c* may be of interest to users who are in similar contexts to *c*. Then, we do not need, in a first moment, of a historical database of recommendation evaluations. MMedia2U uses a knowledge-based recommendation method trying to avoid the cold start problem of collaborative filtering [8].

3. MOBILE MEDIA TO YOU (MMedia2U)

In the system presented in this paper, users receive recommendations of photos created in contexts similar to current users' context. This similarity computes three contextual dimensions (spatial, social, and temporal). The system has as target two types of users. The first type are those who are in an unusual context (e.g., visiting a tourist sight for the first time) and they can enjoy the pictures recommended to have references to activities or new places to explore. The second group contains users that have already been in this similar context and that the recommended photos may give a new vision and perspective of the situation they find themselves. Rost et *al.*[11] have noted that georeferenced images can influence in a positive and playful way the exploration of space by these two categories of users.

3.1 Context Modelling

A fundamental part in the development of a context-aware system is the definition of what information should compose the "context", since elements that describe the contextual information depend on the system tasks, and on the system capacity to observe this information. This definition is associated with the creation of a context model, in which are established the elements that compose its description and how it should be represented (e.g., using ontologies, XML, objects). Fig 1 shows our context model represented as OWL-DL ontology. Our model has four dimensions: spatial (location and points of interest), social (e.g., personal information and activity being performed), temporal (date and time) and computational (mobile device).

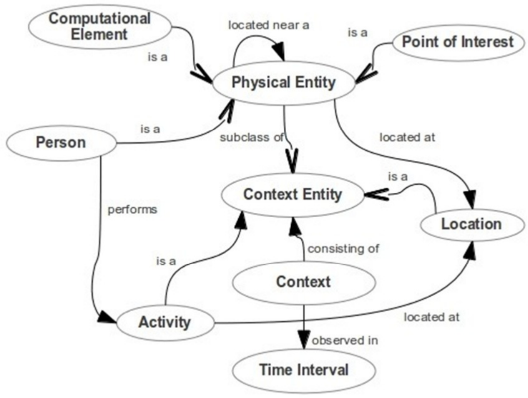

Fig 1- Our Context Model.

In MMedia2U, these dimensions were explored in the acquisition of knowledge about users and photos. These dimensions are already exploited in context-aware management of photos [3]. They have been proved to be relevant in organizing and finding personal photos, which is an indicator that can also be exploited in recommender systems of this type of multimedia document. In MMedia2U, the location attribute is extracted from the user's mobile device (e.g., GPS). Other attributes such as place description (e.g., shopping, beach, etc.) can be derived from freely available web services such as *GeoNames*[3] e *WikiMapia*.[4] Date and time considered are the time of use of the system. In the current version, the activity needs to be informed by the user and can be chosen from among the options presented or reported manually. Examples of activities are: sports, festivals, and landscapes.

3.2 Similarity measure

Similarity measure is used in the system in order to retrieve those photos created in contexts more similar to the user's current context. The algorithm developed is an adaptation of traditional knowledge-based techniques [13], which uses the user context as indicative of their preferences and the context of items (i.e., photos) as a representation of its features. In our system, the context of items is the context in which the photos were created.

The similarity is calculated between the context of the user U and the context of an item I using the following formula:

$$\text{similarity}(U, I) = \sum_{c \in CONTEXT} w_c * \text{sim}_c(U, I) \tag{1}$$

In Formula 1, the similarity is calculated without the need of training data. In this case, *c* is an attribute belonging to the dimensions of the context model (e.g., location); w_c is the weight of influence of attribute *c* (e.g., location has a weight of 50%) and sim_c is the similarity function for attribute *c*. Those pictures that have the highest value of similarity are the ones recommended to

[3] http://www.geonames.org/

[4] http://wikimapia.org/

user U. The function sim_c is particular to each type of context and application domain.

Each context model dimension must have a method to calculate its similarity. The location similarity, for instance, can be calculated by measuring the distance between the place where the picture was taken and the current user's location. The similarity for activity is calculated by comparing the activity or occasion that the user is found and the activity or occasion in which the image was generated. Some of the activities mapped in image shoots and their similarities are shown in Table 1.

	Shopping	Party	Leisure	Sports
Shopping	1	0	0	0
Party	0	1	0,5	0
Leisure	0	0,5	1	0,5
Sports	0	0	0,5	1

Tab. 1 – Activity Similarity

Formula 2 is used with numeric context attributes. The similarity is defined by how close two values are.

$$\mathrm{sim}_c(U,I) = 1 - \frac{|V_c(U) - V_c(I)|}{max(c) - min(c)} \quad (2)$$

In Formula 2, $V_c(U)$ and $V_c(I)$ represent, respectively, the values of context c for the user and the item; $max(c)$ and $min(c)$ represent, respectively, the maximum and minimum values for the compared attribute of context c (e.g., for c = hour of the day, $min(c) = 0$, and $max(c) = 12$). The similarity between dates can compare the various attributes related to the moment the photo was shot and the moment the user is found. Some attributes compared were: the hour of the day, day of the week and month of the year. The date attributes were compared individually to analyze the influence of each one (e.g., Has the similarity between hour of the day a greater influence on the choice of the user than the similarity of the day of the week?). The similarity between the months of the year can be calculated by Formula 2, adapting it to cyclic values (e.g., the distance between January and December is 1, instead of 11).

3.3 System Architecture

We designed the MMedia2U following a client-server architecture for mobile computing that is based on RESTful Web Services. This design allows mobile devices, even those with low processing capabilities, make use of our recommender service. Fig 2 presents the execution flow of the system.

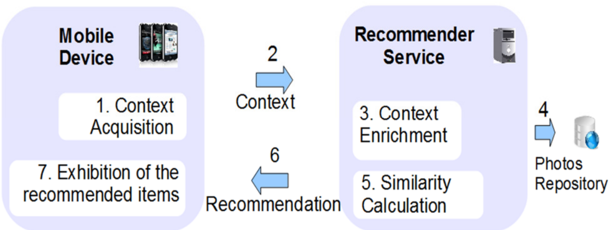

Fig 2. Execution flow of the MMedia2U recommendation.

In step 1, a mobile application is responsible for gathering user's context. Some types of information (e.g., location) can be acquired from sensors (e.g., GPS). Other types rely on information passed by the user (e.g., current activity). In the second step, the mobile application accesses, from a HTTP call, the Web-Service provided by our recommender system, and it informs the user's current context. MMedia2U server receives the request, and performs an enrichment of user context data. The metadata stored in the repository of photos are scanned and compared with the current user's context. In step 5, MMedia2U computes a photo ranking according to the results of similarity measures. The ranking contains the photos URLs and their metadata.

3.4 Photo Corpus

MMedia2U has a repository of photos for recommendation. These photos need to be associated with contextual information in order to be compared with the current context of the mobile users. The photos should have as metadata the location where they were taken and the activity of the photo's author at the time of their creation. At first, we expect to use photos from Web 2.0 applications, such as Flickr and Picasa Web. However, we find many errors in the metadata of these photos (e.g., time, location). In order to evaluate the recommendation method without annotation errors, we built a repository of photos from images of Picasa Web. Manually, we corrected and increased the metadata returned by the Picasa Web Service. Then, we incorporated the new metadata into the photo file by using IPTC and EXIF headers. Examples of enrichment are the inference of the activity from the photos description, and the day of week according to the shot date. We hope the evolution of multimedia content management systems, such as CoMMeDiA [3], will reduce the effort to enrich this kind of image metadata.

3.5 MMedia2U Mobile Application

The mobile application was developed for the *Android*[5] platform, compatible with devices that have version 2.2 or higher. Fig. 3 shows an execution flow of the mobile application.

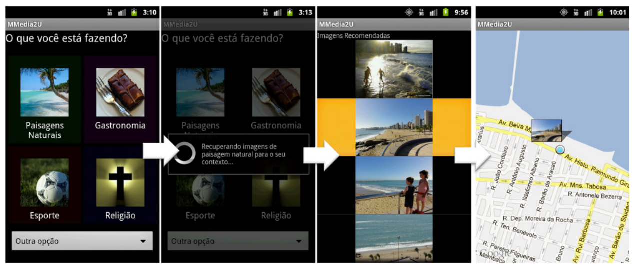

Fig 3. Execution flow of the mobile application.

The user chooses the "activity/interest," then, the system captures the current context (location and date/time), and it sends to the server. This, in turn, returns a list of recommended images (third screen). If the user selects a photo, he can see its position and the distance between him and the place where the photo was captured. Clicking on the picture located on the map, its content is displayed in full screen, and its metadata can be also viewed.

4. EXPERIMENTS

Evaluating a recommendation system is a hard task, due to the property that an item's relevancy has a strong personal nature and is complex to be measured. This difficulty is enhanced when exist a lack of historical evaluation data, which makes large-scale studies very costly and difficult to be run. In the case of CARS, the complexity is even bigger, since we need to range the possible contexts of real situations (i.e., places, daily situations, etc.). As we did not have a historical data about recommended photos, we created a *Gold Standard*, which consisted of photos evaluated by

[5] http://developer.android.com/

users in a certain context. The objective was both to use the *Gold Standard* to compare the performance of our recommendation and to use it as historical data.

While building this *Gold Standard*, it was asked for a group of 13 users to evaluate photos from 8 different contexts, each one consisting of a stage of evaluation. In each stage, one context (e.g., shopping in some stores on the seaside of the city of Fortaleza[6] during the evening) was presented to the user, who had to visualize a set of photos and choose those that seemed to be more appealing for him/her, taking into consideration the context he/she was suggested. The photos chosen by the users were included in the *Gold Standard* and provided the historical base in which the recommendation of MMedia2U was evaluated. The degree of success on recommendations was then evaluated by the ratio of chosen photos that are in the *Gold Standard* (e.g., if in a given combination of user and context, 10 photos are recommended and 6 of these are in the *Gold Standard* of the same combination, then the recommendation precision is 0.6).

In each stage of evaluation, a mean of 100 photos were visualized by the users. 20 were taken in similar contexts to the one showed to the user and 80 were different in some dimensions of the context (e.g., same activity but very distant location). Five of 13 users didn't know the place chosen as the location for the users (the seaside of Fortaleza). Ten users were Computer Science graduate students aging from 23 to 28. All of them use mobile phones every day. For some of the users, we have presented 8 photo collections, while for others we have only presented a subset of it, performing 66 simulations.

4.1 QUANTITATIVE RESULTS

The objective of the experiment was to evaluate the following hypotheses:

(H-I) It is possible to make satisfactory recommendations of georeferenced photos without prior knowledge of the user profile, considering only its current context;

(H-II) The context in which the photos were taken is relevant in making recommendations; and

(H-III) The usage of a context model considering various contextual dimensions may lead to an improved recommendation comparing to the result of one which uses only one context attribute (e.g., location).

In order to verify these hypotheses, first, the algorithm was run with different weights for each dimension without a previous training data. Another implementation used the weights obtained by training the algorithm using a 7-fold method (same way of adjusting the parameters) [5]. Weights adjustment in the training data was performed by linear regression using the least squares method. We compared the precision changing the amount of recommended photos. For this analysis, we also used random choices as base statement since we were unable to find other photo recommendation algorithms that use context information.

Table 2 shows the average precision of our recommendation algorithm in relation to four sizes of recommendation lists (Top 3, Top 5, Top 10 and Top 20). The average precision of the algorithm without training, assigning equal weights to all similarity measures, was 0.54 for the Top 5 (5 recommended items). Using the weights obtained by the least squares, the precision of the Top 5 was 0.55. Recommending pictures at random, without the ranking generated by the recommendation algorithm, the average precision was 0.28. This precision is relatively high since some users have chosen more than 30 photos for a specific context (e.g., a user in particular has selected half part of the corpus).

The last two rows of the table show the average precision when using the calculation of similarity of only one of the contextual dimensions. Combination I, which got the best results, was the one we assign twice the importance of Activity in relation to Location and four times in relation to temporal attributes.

Comparing the results obtained without training (Combination I and Equal Weights) in relation to random experiments, one can see that it is possible to have much higher precision than random choices (agreeing with the hypothesis **H-II** and **H-I**). Moreover, the gain over the random method was not big when considering only one contextual dimension (e.g., only location or only activity), leading us to believe that a model of full context is essential for a good context-aware recommendation (hypothesis **H-III**).

Fig 4 shows the F measure (harmonic mean) analysis for the six recommendation algorithms. Equal weights and Combination I had the better results for Top 3. Regarding the Top 5, Top 10 and Top 20 lists, Combination I and Least Square were the most effective. For instance, Combination I (0,410 for Top 20) was two times better than Random algorithm (0,201).

We used the Student t-test in the precision values for the Top 3 obtained by the algorithms compared to the results obtained by random selection of photos. This showed that the combined use of contextual attributes is significantly better than randomization (95% degree of confidence, probability < 0.0001). In addition, the test showed that the use of only one attribute (activity or location, for example) is not significantly better than the random method. Finally, the comparison with the results of Combination I and Least Squares resulted in performance differences not statistically significant, which may not indicate the need for training to improve the precision of this kind of CARS.

	Top 3	**Top 5**	**Top 10**	**Top 20**
Equal weights	0.56	0.54	0.45	0.42
Least Squares	0.54	0.56	0.51	0.44
Combination 1	0.56	0.56	0.51	0.47
Random	0.29	0.28	0.26	0.25
Localization	0.28	0.30	0.30	0.29
Activity	0.35	0.33	0.33	0.34

Table 2. Mean values of obtained precisions.

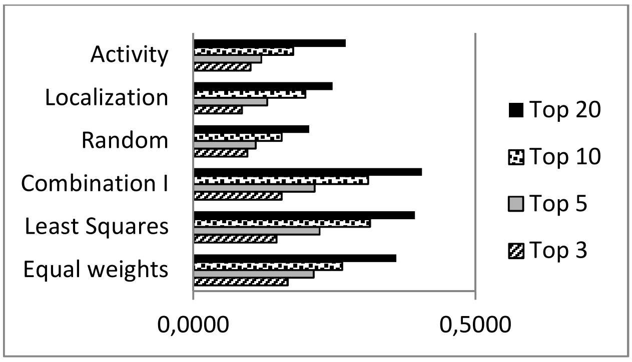

Fig 4- Mean values of f-measure

6 http://en.wikipedia.org/wiki/Fortaleza

4.2 USERS QUESTIONNAIRE

At each stage of the experiments, a questionnaire was applied to the users so that relevant factors were investigated in the implementation of CARS for photos. One of the factors to be investigated was the relevance of a mobile system for photo recommendation and, in the case of existence of such system, if it would be interesting to recommend photos taken into account the current user's context. He/she was asked whether, in this specific context, the user would like to receive recommendations of photos taken in a similar context. 74% of users answered "yes" for this question. When asked whether the recommendation of pictures generated in similar contexts would be interesting, the level of interest was 100%. Another point investigated was the relative weight of each contextual dimension. Each user was asked what contextual dimension, including location, activity, date and time; they would prefer to be taken into consideration to build the set of photos. Eight of the users said they think the proximity of the location of the photo is the most important factor to increase their interest. Five of the users responded that the activity associated with the picture portrayed by the activity, which he/she is playing at the moment is the most important factor. No users found the similarity between date and time the most important factor.

5. CONCLUSION AND FUTURE WORK

This work presented MMedia2U, CARS for contextual photos. During the system development, we specify a context model to cope this application domain and we adapted a knowledge-based technique to incorporate the context information. MMedia2U allows recommendation of photos even those that have never been evaluated by users. The recommendation is performed only from the context in which the photos were created. This allows users, without a history of use, to receive recommendations based on their current context. The recommendation mechanism was validated by the construction of a *Gold Standard*. The average precision achieved by the algorithm allowed us to conclude that, for the data used, context-awareness can bring gains in the photo recommendation compared to a random list. It is important to note that even weight combinations without training phases (the *Comb I* and *equal weights*) achieved satisfactory results. This strategy can be used to reduce the drawback of cold start problem. In addition, the user's survey suggests that systems of this nature are interesting to users.

In this moment, we cannot generalize our performance results (dependence of users and the images corpus). However, it serves as a good indication of the quality of the prototype recommendation. Analysis of these results was limited to the size of the corpus (655 photos, with 335 from two tourist areas of Fortaleza, city in Brazil). We think the increase of the database could decrease the precision of the algorithm or present concentration of photos in certain places. In such cases, new search filters and clustering algorithms should be used to solve these challenges. Nevertheless, the results are already encouraging the construction of context-aware photo corpus of city sights (e.g., points of interest in host cities of 2014 World Cup).

As future work, we want to increase the image corpus by using an evolution of the CoMMeDiA system [3]. We aim to ensure the accuracy of contextual information, and take benefits from the automatic context acquisition. We also expect to evaluate MMedia2U with users (tourist or otherwise) in a real mobile situation. In a new version of the system, we would like to integrate clustering algorithms in the rank results, and allow users to add words of interest in order to refine the recommendations.

ACKNOWLEDGEMENTS

This work is a partial result of the UbiStructure project supported by CNPq (MCT/CNPq 14/2011 - Universal) under grant number 481417/2011-7.

6. REFERENCES

[1] Van Setten, M., Pokraev, S. and Koolwaaij, J. 2004. Context-aware recommendations in the mobile tourist application compass. In *Adaptive Hypermedia and Adaptative Web-Based Systems*, 235–244

[2] Xia, S., Gong, X., Wang, W., Tian, Y., Yang, X., and Ma, J. 2010. Context-Aware Image Annotation and Retrieval on Mobile Device. In *Second International Conference on Multimedia and Information Technology*, 111-114.

[3] Viana, W., Miron, A., Moisuc, B., Gensel, J., Villanova-Oliver M. and Martin, H. 2011. Towards the semantic and context-aware management of mobile multimedia. In *Multimedia Tools and Applications*, v. 53, 391-429

[4] Jannach, D., Zanker, M., Felfernig, A. and Friedrich, G. 2010. Recommender Systems, an Introduction. In *Cambridge University Press*, first edition

[5] Fabiana G. Marinho, Rossana M.C. Andrade, Cláudia Werner, Windson Viana, Marcio E.F. Maia, Lincoln S. Rocha, Eldânae Teixeira, João B. Ferreira Filho, Valéria L.L. Dantas, Fabrício Lima, Saulo Aguiar, MobiLine: A Nested Software Product Line for the domain of mobile and context-aware applications, Science of Computer Programming, Available online 18 May 2012, ISSN 0167-6423, 10.1016/j.scico.2012.04.009.

[6] Adomavicius, G. and Tuzhilin, A. 2010. Context-Aware Recommender Systems. In *Recommender Systems Handbook: A Complete Guide for Research Scientists and Practitioners*, Springer Press, first edition

[7] Lee, J. S. and Lee, J. C. 2007. Context awareness by case-based reasoning in a music recommendation system. In *Proceedings of the 4th international conference on Ubiquitous computing systems*

[8] Burke, R. 2000. Knowledge-Based Recommender Systems. In *Encyclopedia of Library and Information Systems,* no. 32, 180-200

[9] Albanese, M., d'Acierno, A., Moscato, V., Persia, F. and Picariello, A. 2011. A Multimedia Semantic Recommender System in Cultural Heritage Applications. In *Semantic Computing (ICSC), 2011 Fifth IEEE International Conference on* , pp.403-410, 18-21

[10] Boutemedjet , S. and Ziou , D. 2010. Using Images in Context Aware Recommender Systems*, In Proc of 1st Int'l Workshop on Adaptation, Personalization and Recommendation in the Social-semantic Web*

[11] Rost, M., Cramer, H., and Holmquist, L. E. 2011. Mobile exploration of geotagged photographs. In *Personal and Ubiquitous Computing*

[12] Siersdorfer, S. and Sizov, S. 2009. Social recommender systems for web 2.0 folksonomies. In *Proceedings of the 20th ACM conference on Hypertext and hypermedia* (HT '09)

[13] Mcsherry, D. 2003. Similarity and Compromise. In *Proceedings of the Fifth International Conference on Case-Based Reasoning.*