Proposal for

FQAS'09 Special Session on Personalization, Preferences, Context and Recommendation

Proposed by
Juan M. Fernández-Luna, Juan F. Huete, and Pablo Castells

Introduction and Motivation

The goal of enhancing IR models and methods towards user-aware and context-aware models has raised increasing interest in the research community, and is being identified as a key step in order to cope with the continuous growth of information environments (repositories, networks, users) worldwide. The notion of context refers to any dynamic condition occurring at the time when an information retrieval task takes place, and which may be relevant to fully define and understand a user need. Several context-sensitive retrieval algorithms exist in the literature; most of them based on statistical language models to combine previous queries and clicked document summaries with the current query, for better ranking of documents. Relevance feedback [Roc71], and implicit feedback [STZ05], exploit contextual user input as a source of information to complement explicit user queries and guide the retrieval process. When the notion of context focuses on persistent user characteristics and preferences, it is usually referred to as an issue of personalization [MS04]. A significant body of research in the last two decades has paid attention to the problem of personalizing information access and delivery, commonly addressed under such names as information filtering, collaborative filtering, recommender systems, or personalized IR, with variations in approach and perspective [HSS01]. From different angles, the problem has been a major research topic in fields such as IR, User Modelling, and Machine Learning. In general, personalizing the retrieval of content involves knowing something about the user beyond her last request, and taking advantage of this knowledge in order to improve the system response to the actual user need. In an increasingly demanding and competitive market, room for such improvement exists often nowadays, to varying degrees, in common retrieval scenarios, either because the request is vague or because there is no explicit request at all. The research activity in this area has been paralleled by a comparable interest towards making such techniques commercially profitable. Popular online services such as Google [BL05], [ZKF05] or Amazon [LJB01], [SLZ05] are nowadays exploiting some basic personalization features.

Most of the proposed techniques in this area keep and process long records of accessed documents by each user, in order to infer potential preferences for new documents. The known or predicted user preferences are commonly represented in terms of thematic categories [LYM04], sets of words [MS04], sets of documents [JW03], or even richer semantic descriptions in terms of domain concepts. User preferences can be explicitly provided by users themselves, or be automatically obtained by monitoring user actions in the retrieval system [MS04]. In the latter case, the system can rely on arbitrary interaction history, including past queries, the documents that the user has chosen to view, and even how a user has accessed a document. In some cases, the system does not retain an explicit representation of user interests, but these are implicit in classification models or functions, such as decision trees, rules, and other classifiers. Based on the relevant

knowledge about user interests that a system is able to capture, different techniques can be used to personalize search results, which come into play at different points of a retrieval system. User preferences have been used to reformulate user queries, by expanding or refining the information in the query [Roc71], [STZ05], or to adapt search results, by re-ranking [MS04] or clustering the results by ranked categories [LYM04]. A long series of variations of the popular PageRank algorithm for Web search have been proposed as well, where a personalized initialization vector results in biased PageRank values towards individual user interests [JW03]. In scenarios where information (preferences, history) of a large number of users is centrally available to the personalization system, it is also possible to apply collaborative filtering techniques in a way that thousands of users benefit from each other's experience without even getting acquainted [AT05]. The effect of all these personalization strategies is reflected in the filtering and reordering of search results, which are biased in ways that match not only the original user query, but also her implicit user preferences.

The concept of Recommender System (RS) is a broader term that combines typical features related to personalization and context. They were born as a solution to the huge amount of information the users can find in the Internet. RSs are applications that give advice to the user about items (movies, music, etc.) that are likely of interest to the user, according to their preferences and tastes. The system usually compares the user's profile with some information extracted from the item (content-based recommending), or from other users who have similar preferences (collaborative recommendation) [AT05]. Finally, Hybrid RS combine both content and collaborative methods [Bur02]. Many different RS approaches have been published using methods from machine learning, approximation theory and various heuristics. With respect to the first type of recommendation, the key element is the concept of profile, a description of the type of items that are interesting for the user. A second possibility could be a history of users interactions with the RS (e.g. items viewed, ratings, items purchased, etc.).

Given a new item and the user model, the recommender predicts whether the user would be interested in the item. Some techniques applied to this problem are, for example, decision trees and rule induction, knn methods, linear classifiers or Naïve Bayes classifiers, among others. The main problems with content-based RS are firstly, the difficulty of making accurate recommendations to users with very few ratings, and secondly, overspecialization. Collaborative filtering assumes that people with similar tastes will rate things similarly. These systems produce predictions for a given user to new items, taking into account those rates that similar users have given to these new items. Some examples of techniques adapted and applied to this field are: User-
Based Nearest Neighbour Algorithms, clustering algorithms to get the groups of similar users, item-based nearest neighbour algorithms, Association Rule Mining, probability algorithms (mainly Bayesian networks), among much others. Collaborative RS have their own limitations: the new user problem, new item problem and the sparsity problem (little information in common among users).

Personalization and context modelling remain hot topics in IR research and industry. Important problems are yet to be solved in order to achieve the quality, reliability and maturity required for a widespread deployment of these techniques. Personalization systems often fail to acquire enough or sufficiently accurate knowledge about users, as finding implicit evidence of user needs and interests through their behaviour is not an easy task. Inherent difficulties are involved indeed when attempting to deal with (or even define) aspects related to human cognition and volition. Even when the system assumptions are correct, the adaptive actions can be obtrusive or inappropriate, if not handled properly. Coping with the dynamics of user interests (e.g. persistent vs.

occasional), the different time scales on which they evolve (e.g. slow persistent changes, quick temporal changes), the interrelations among different time windows (e.g. a temporal interest becoming persistent, a long-term preference coming into play, etc.), the multiple sides or user preferences, or the relations between preference and situation, are some of the challenging problems in this area.

Therefore, there are three main pillars supporting this session: Personalization, context and recommendation. Each of them is a research field in itself garnering the attention of a wide community researchers worldwide, as evidenced by the increasing number of events dealing with these topics. The special session that we are proposing aims to create a new and dynamic forum in order to contribute to the discussion of new proposals in these areas, linked to the scope of the FQAS conference.

Scope

This session is intended as a discussion forum where novel proposals around the concepts of personalization, context and recommendation are presented. We welcome contributions related, but are not restricted, to the following topics:

- Modelling and profiling ~~of~~ personal, social and contextual information.
- User profiling, preference elicitation and use.
- Context modelling, identification and use.
- Personalized information access.
- Content-based, collaborative, and hybrid recommender systems.
- Group recommendation.
- Evaluation methodologies and metrics.
- ...

Reviewing Process

All submitted papers will be reviewed by two or three experts in the areas of the session.

Important Dates

- Distribution of the Call for Papers: June 12, 2009??
- Deadline for paper submission: July 12, 2009.
- Notification of acceptance: July 27, 2009.
- Deadline for camera ready submission: August 9, 2009.
- Session date: October 26-28, 2009.

Publicity

A Call for Papers will be prepared and disseminated through the major IR distribution lists and publicized on a web page, as well as by means of personal invitations to relevant researchers in the area.

Submission Format

As the papers accepted in the session will be published in the proceedings of the conference; in a Lecture Notes in Artificial Intelligence volume, submissions shall adhere to the guidelines of the general conference, that is, no more than 12 pages in the LNCS/LNAI format (see http://www.springer.com/computer/lncs?SGWID=0-164-7-72376-0 for more information).

Tentative Organization of the session

Assuming a maximum of 6 accepted papers, the session could be allocated in a morning or an afternoon session slot, with the following tentative schedule:

- Presentation of the session (15').
- Part I. Presentation of three papers (3 * (25'+5') = 1h 30').
- Break (15').
- Part II. Presentation of three papers (3 * (25'+5') = 1h 30').
- Part III. Panel and discussion (30').
- Conclusions and closing (10').

Total time: 4h 10'.

Organizers

Juan M. Fernández-Luna (jmfluna@decsai.ugr.es, Universidad de Granada, Spain) got his Computer Science degree in 1994 at the University of Granada, Spain. In 2001 he got his PhD at the same institution, working on a thesis in which several retrieval models based on Bayesian networks for Information Retrieval where designed. This subject is still his current research area, although moving to structured documents and recommender systems. He has got experience organising international conference as he was the co-chair of the 27th European Conference on Information Retrieval Research and of the I and II International Workshop on Teaching and Learning of Information Retrieval, and the SIGIR'07 Workshop on Information Retrieval and Graphical Models. He has also been guest editor of a special Information and Processing Management issue on Bayesian networks and Information Retrieval, and of two special issues of Information Retrieval.

Juan F. Huete (jhg@decsai.ugr.es, Universidad de Granada, Spain) is assistant professor at the Department of Computer Science and Artificial Intelligence at the University of Granada. He got his PhD in 1995, researching on the uncertainty treatment in Artificial Intelligence under the formalism of Bayesian networks. From 1998, his research interest is Information Retrieval, and later recommender systems, designing several models based on these graphical models. He has been co-editor of a special special Information and Processing Management issue on Bayesian networks and Information Retrieval, co-chair of the I and II International Workshops on Teaching and Learning of Information Retrieval and the SIGIR'07 Workshop on Information Retrieval and Graphical Models.

Pablo Castells (pablo.castells@uam.es, Universidad Autónoma de Madrid, Spain) is assistant professor at the Escuela Politécnica Superior of the Universidad Autónoma de Madrid (UAM). He earned a MSc degree in Mathematics in 1989 and a PhD in Computer Science in 1994 at UAM. His research field lies in the area of Information Retrieval, where his current interests focus on context modelling, personalization, recommender systems, and social networks. He has led or participated in several national and international R&D projects and private contracts in these areas. He serves regularly as a reviewer or scientific committee member for major international journals and

conferences, and was guest editor of a special issue on Multimedia Semantics, Adaptation and Personalization, in the Signal, Image and Video Processing journal. From 2000 to 2004 he was a steering board member of the Spanish HCI Association.

Bibliography

[HSS01] Hanani, U., Shapira, B., and Shoval, P. Information Filtering: Overview of Issues, Research and Systems. User Modelling and User-Adapted Interaction 11(3), 2001, pp. 203-259.

[JW03] G. Jeh and J. Widom. Scaling Personalized Web Search. 12th International WorldWide Web Conference (WWW 2003). Budapest, Hungary, May 2003, pp. 271-279.

[LYM04] F. Liu, C. Yu, and W. Meng. Personalized Web Search For Improving Retrieval Effectiveness. IEEE Transactions on Knowledge and Data Engineering 16(1), January 2004, pp. 28-40.

[MS04] A. Micarelli and F. Sciarrone. Anatomy and Empirical Evaluation of an Adaptive Web-Based Information Filtering System. User Modelling and User-Adapted Interaction 14(2-3), 2004, pp. 159-200.

[Roc71] Rocchio, J., Relevance feedback information retrieval, in Salton, G. (Ed.), The Smart Retrieval System- Experiments in Automatic Document Processing, Prentice-Hall, 1971, pp. 313-323.

[STZ05] Shen, X., Tan, B., and Zhai, C., Context-sensitive information retrieval using implicit feedback, 28th annual international ACM SIGIR, 2005, pp. 43-50.

[BL05] Badros, G. J., and Lawrence, S. R. (2005). Methods and systems for personalized network searching. US Patent Application 20050131866.

[LJB01] Linden G., Jacobi, J., and Benson E. (2001). Collaborative recommendations using item-to-item similarity mappings, US patent 6266649B1.

[SLZ05] Smith B. R., Linden, G. D., and Zada, N. K. (2005). Content personalisation based on actions performed during a current browsing session, US patent 6853983B2.

[ZKF05] Zamir, O. E., Korn, J. L., Fikes, A. B., and Lawrence, S. R. (2005). Personalization of placed content ordering in search results, US Patent Application 20050240580.

[AT05] G. Adomavicius, A. Tuzhilin. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extentions. IEEE Trans. on Knowledge and Data Engineering, 17(6):734-749, 2005.

[BFHS07] J. Ben Schafer, D. Frankowski, J. Herlocker, S. Sen. Collaborative Filtering Recommender Systems. Lecture Notes in Computer Science. 4321, 291 – 324, 2007.

[Bur02] R. Burke. Hybrid recommender systems: survey and experiments. User modelling and User-Adapted Interaction 12:331-370, 2002.