Personalization of Parliamentary Document Retrieval using different User Profiles

Eduardo Vicente-López, Luis M. de Campos, Juan M. Fernández-Luna, and Juan F. Huete

Departamento de Ciencias de la Computación e Inteligencia Artificial, E.T.S.I.I.T., CITIC-UGR, Universidad de Granada, 18071-Granada, Spain
{evicente,lci,jmfluna,jhg}@decsai.ugr.es

Abstract. Owing to the information overload we are faced with nowadays, personalization approaches are becoming almost a must, in order to provide relevant information for users. These personalization techniques retrieve results closer to the user interests and preferences, by using the information stored in the user profile. We have carried out a comparative study between six different user profile representation approaches, based on the content of the documents of the Andalusian Parliament, obtaining quite good personalized performance results and some interesting conclusions about the goodnesses of these content-based approaches.
Keywords. User Profiles, Personalization, Information Retrieval, Parliamentary Documents

1 Introduction

Over the last few years, the amount of digital information is rising exponentially [10], so its access is everyday more difficult. The use of Information Retrieval Systems (IRS) has become essential to find relevant information within this huge bunch of data. The use of such systems in e-Government will help to deliver the needed information to citizens, representing a particular, and important, application of IR techniques.

This paper is framed within our collaboration with the regional Parliament of Andalusia (Spain). Particularly, we have built an IRS [6] to enhance the access of the citizens to the Records of Parliamentary Proceedings, called *Seda* (`http://irutai2.ugr.es/SEDA`), taking the most of the internal structure of such documents and founding our search engine on XML retrieval. Among the proceedings, the textual transcriptions of the working Committees can be found, considering policy issues, conducting inquiries and producing reports on a range of matters. Each Committee is devoted to a wider topic of interest as agriculture, education or economy (the number of committees and the covered topics varies between terms of office – for example, nowadays there are 11 different committees). Each of these records (or documents) contains an average of 5.6 initiatives, which present a detailed discussion of the members of the Parliament about a specific issue. In turn, each of these initiatives is tagged with one or more subjects extracted from the EUROVOC thesaurus[1], being manually assigned by parliamentary documentalists as the best representation of its content.

The citizens can search for a piece of information by submitting a query to the system. Although in the last years our system has been providing quite good results, it offers the same output for a given query, independently of the user, since it only considers the query keywords as the representation of the user information needs. This issue is well-known as the *'one size fits all'* paradigm.

If we join the continuous increase of data, with the users tendency to formulate short and ambiguous queries [19], a new approach is required, in which the user context, and not only the query, is considered as an important part within the retrieval process. Personalization [2, 3, 22] is this possible solution, and hot arising research area, whose main objective is to retrieve results closer to the users, in order to better satisfy their specific information needs.

Any personalization process has three main different stages: 1) to acquire and represent the user interests and preferences in the *user profile*, 2) to exploit the best as possible the user profile information within the retrieval process, and 3) to evaluate the whole personalization process. We may consider some additional issues, such as privacy in the personal data collection and management process [13], or different ways to present the personalized results [1], with the intention of presenting this information to the user in the most easy and intuitive way.

It seems quite obvious that the personalization process expected performance depends on the quality of the user profile information. For this reason, in this article we have focused on the analysis of different ways to build user profiles. Concretely, we shall focus on content-based user profiles which are frequently used in *contextual evaluation* environments, such as [18]. Additionally, this kind of profiles could be ideal for the introduction of personalization in privacy constrained environments. We have concretely faced this problem with the Andalusian Parliament, where the members of the Parliament do not allow any personal data recollection of themselves nor the citizens. In this way, the parliamentary IRS could integrate personalization techniques to improve its retrieval performance and user satisfaction, only giving the user the possibility to choose with which of the simulated profiles he/she is more alike.

To build these user profiles, we have to analyse the content of the Records. We take the advantage of a pre-classified collection, i.e. each document belongs to one committee representing a different area of interest or category, in which the future users could be interested in. As a first approach, we develop a user profile based on the EUROVOC thesaurus subjects, manually assigned by the Andalusian Parliament documentalists to each initiative discussed in a committee session. Secondly, we build a user profile only based on terms (independently of where they appear in the document), and thirdly we have configured a hybrid user profile composed of both terms and subjects. Finally, we evaluate the use of each alternative in order to find the most appropriate in terms of retrieval performance.

[1] http://eurovoc.europa.eu/

While these content-based simulated user profiles could be considered as lacking *'reality'*, since they do not represent real users, they are a valid approach [9, 18] for possible users interested in some areas of interest[2]. If we join the recent rise of personalized systems, together with the fact that their evaluation through user studies is rather complicated (due to the large required resources, such as, access to real users, time, money or even the needed infrastructure for their implementation), we consider particularly important to test and improve the quality of content-based user profiles.

The rest of the article is organized as follows: Section 2 presents a review of the state-of-the-art in user profiles. Section 3 shows how the profiles are built, based on subjects, terms and a combination of both of them at the same time. How these profiles are used in conjunction with the user query, the experimental design and the evaluation is described in Section 4. Finally, the last section of the paper shows the conclusions and proposals for further research.

2 User Profiles Literature

The quality of personalized results will highly depend on the user profile quality and how it is exploited in the retrieval process. Hence, the user profile building process is one of the most important steps to obtain good personalized results, but at the same time very difficult, since user interests and preferences are difficult to be captured and they also change over time [14, 17].

The three most important steps in the user modeling-user profile building process, according to [11] are the following : 1) acquisition of user information, 2) user profile representation, and 3) user profile update. We will focus on the second step, since this article main goal is to make a comparative study between different user profile representations performance.

The three main user profile representation approaches are: a set of weighted keywords, semantic networks, and a set of weighted concepts:

a) Weighted keywords: it is the most common user profile representation. They may be automatically learned from user visited documents or directly given by the user. The keyword weights show the importance of each keyword within the profile. Examples of this approach are [20], where they build three different user profiles based on relevance feedback and implicit information, user browsing history, and a modified collaborative filtering. Other examples are [16], where they learn the user profiles from the user visited web pages, based on the well-know *tf*idf* approach, and [8], where they build user profiles formed by a vector of keywords for each user area of interest.

b) Semantic networks: in order to handle the keyword user profile polysemy problem, a weighted semantic network is included, in which each node represents

[2] In this situation the user might also opt for the inclusion of several terms in the query describing the committee content, terms that could be difficult to select for a citizen, appearing also a query drift problem. Or, otherwise, opting for filtering out the documents which do not belong to the committee, but in this case there might be relevant results which are not shown to the user (around 25% in our studies).

Table 1. Examples of the three proposed user profiles, for the 'agriculture and livestock' area of interest (unstemmed and translated into English).

sProf	$s = \{$ *0.216*"agriculture aid" 0.127*"agricultural policy" 0.098*"agricultural production" 0.098*"oily" 0.095*"food industry" 0.091*"fishing" 0.083*"oil" 0.075*"huelva province" ...* $\}$
tProf	$t = \{$ `0.007*agriculture 0.007*sector 0.004*fishing 0.004*agrarian 0.004*production 0.003*aid 0.003*farmer 0.002*product 0.002*rural 0.001*oil ...` $\}$
stProf	$s_1 =$ *0.216*"agriculture aid"* $t_{s1} = \{$`0.007*aid 0.006*sector 0.006*agriculture 0.005*farmer ...`$\}$ $s_2 =$ *0.127*"agricultural policy"* $t_{s2} = \{$`0.009*agriculture 0.007*agrarian 0.006*production ...` $\}$ $\vdots$ $\vdots$

a concept. For example, in [15] a filtering interface is created to personalize the results from the Altavista search engine. Another semantic network example is [18], where a personalized search system with ontology based user profiles is presented. These user profiles are built assigning scores to user interests, implicitly derived from concepts of the ODP ontology. Since the user interests are dynamic, a propagation algorithm is used to keep these interests updated.

c) Weighted concepts: they are similar to the semantic networks, since they also have conceptual nodes and relations between them, but in this case, the nodes are represented by abstract topics of interest for the user instead of terms. But, at the same time, they are also similar to the weighted keyword user profiles, since they are usually represented as vectors of weighted concepts. Nonetheless, in the last few years is common to use a hierarchical representation of concepts, usually derived from a taxonomy, thesaurus, or a reference ontology, instead of using concepts with no structure, allowing a much richer representation. An example of this approach is [21], where using concepts from the ODP ontology first three levels, they build user profiles based on the user browsing history. Another example is [4], where they show three different ways to use ODP: first, as a semantic support to find relations between concepts; second, identifying some ODP structure parts relevant to the user; and third, the user directly choose the ODP concepts he/she is interested in. After that, they study how to exploit these three user profiles, with personalization techniques based on query modification and re-ranking.

3 User Profiles Building Process

Due to the frequent important restriction concerning collecting user personal information, and additionally to the difficulty to have accurate and updated user profiles, we have decided to build simulated user profiles based on content. Particularly we focus on the information available in the transcriptions of the working Committees, where much of the work of the Parliament takes place. Thus, assuming that those topics in a given committee might represent the interests (preferences) of the citizens, we analyse its content to learn the profile. In this paper we will explore three different types of user profiles, see Table 1 for an example:

- *sProf*: This first approach, based on the initiative subjects, can be considered as a *weighted concept* profile, since these subjects represent abstract topics of interest for the user but not terms. They are represented as vectors of weighted concepts, without any structure. Concepts profiles main assets are their robustness to vocabulary variations and a less requirement of user feedback. These characteristics and the fact that the subjects are manually selected by experts in the document collection, as the best content representation for the parliamentary initiatives, were the reasons which made us to start with this approach to learn the user profiles.
- *tProf*: The second profile approach, based on the collection terms, can be considered as a *weighted keyword* profile, since the terms themselves are the items which represent the user interests. These profiles are the easiest to build, but they need to have many terms to accurately define a user interest. These profiles are also less understandable for users than those based on concepts, since their interests are much easily mapped with concepts than with isolated terms. But at the same time, the terms let a more fine-grained representation of the collection content.
- *stProf*: The third profile approach, based on subjects and terms, is an hybrid approach among the weighted concept and weighted keyword profiles, keeping concept abstraction but enriched by the terms fine-grained contribution. To build this profile we learn the most representative terms for each collection subject. Thus, this new profile now contains two levels: the first, with the subjects which represent the profile, and the second formed by the terms which represent each first level subject.

We now show the way we select the elements of each type of profile. Let X represent either a subject in the case of *sProf* or a term in the case of *tProf*, and let Y represent a profile. Then we define $f^{+}(X,Y)$ as the frequency of X in documents belonging to any area(s) of interest which form the profile Y; $f^{+}(Y)$ is the number of elements (either subjects for *sProf* or terms for *tProf*) within Y; $f^{-}(X,Y)$ and $f^{-}(Y)$ are respectively the frequency of X and the number of elements in documents outside the profile Y. For the *stProf* profiles, X represents a term and Y represents a subject, $f^{+}(X,Y)$ being in this case the frequency of X within initiatives classified by the subject Y and $f^{+}(Y)$ the total number of terms within these initiatives; $f^{-}(X,Y)$ and $f^{-}(Y)$ have in this case the obvious meaning. We then define the relevance of X with respect to Y, $R(X,Y)$ as

$$R(X,Y) = \frac{f^{+}(X,Y)}{f^{+}(Y)} - \frac{f^{-}(X,Y)}{f^{-}(Y)}$$

i.e., the normalized frequency of X within Y minus the normalized frequency of X outside Y. If the final value is $R(X,Y) \leq 0$, it means that X is more frequent outside than within Y, so it is not representative of Y and we do not consider it. However, if the final value is $R(X,Y) > 0$, this means that X represents Y at a certain degree, so we keep it. All the retained elements are sorted in decreasing order of relevance to form the profile. In the case of the *stProf* profile we first calculate the list of subjects and next the list of terms associated to each subject.

Table 2. Final *sProf* and *tProf* user profiles using $exp[Subj|Terms] = 5$ and $maxNorm = 0.66$.

sProf	*0.66*"agriculture aid" 0.388*"agricultural policy" 0.299*"agricultural production" 0.299*"oily" 0.290*"food industry"*
tProf	`0.66*agriculture 0.647*sector 0.401*fishing 0.399*agrarian 0.398*production`

4 Evaluation Framework and Results

This section shows the components of the used evaluation framework, how we have used the previous user profiles, and the obtained results and conclusions.

The evaluation framework is composed by the following components: a document collection formed by 658 Committee Sessions from the sixth and seventh Andalusian Parliament terms of office, marked up in XML (containing 432,575 retrievable structural units); an heterogeneous set of 23 queries formulated by real users of the document collection; the search engine is Garnata [5]; the relevance assessments were obtained from a carried out user study, which involved 31 users, with a total of 126 evaluation triplets (user, query, profile), i.e., the relevance assessments provided by a given user, evaluating a given query under a given profile (considering each user chose the user profile closer to his/her interests - none of the user profile representations discussed in this article was provided to the user, but a brief general description of its expected content); the NDCG evaluation metric (Normalized Discounted Cumulative Gain) [12], with some special considerations due to the structured nature of the documents; and the personalization techniques are *NQE, HRR, SRR, IRR, NQE+m, HRR+m, SRR+m, IRR+m, CAS and CAS-or*, which represent a highly heterogeneous set of personalization techniques. You can see [7] for a more detailed explanation about any of these evaluation components.

Using the user profiles. We are going to explain how the profiles have been used in the experimentation.

1) *sProf and tProf:* The use of subject-based and term-based profiles is quite simple. It basically involves taking the top-n relevant subjects (*expSubj*) or terms (*expTerms*), with $n = 5, 10, 20, 40$. Once we have these first *expSubj* or *expTerms*, we normalize (proportionally) their weights in such a way that the maximum value (*maxNorm*) is: 0.33, 0.66, 0.99. The combination of *expSubj* or *expTerms* with *maxNorm*, gives us a total number of 12 different weighted subject or term sets, to provide to each personalization technique. Check Table 2 to see an example of these final user profiles from Table 1.

2) *stProf:* Its use is somewhat more complicated. In principle, the process should be to get the first *expSubj* profile subjects (again $expSubj = 5, 10, 20, 40$), and for each of these subjects to get the first *expTerms* terms, with values $expTerms = 1, 5, 10$. Each term weight will be multiplied by its corresponding subject weight. Thus the terms, which will be the ones finally used by the personalization techniques, will already incorporate in their weights the influence of their subjects.

Table 3. Final *stProf* user profile using $expSubj = 2$, $expTerms = 3$ (to make it more clear and short), and $maxNorm = 0.66$.

add	0.66*agriculture 0.421*aid 0.372*sector 0.237*agrarian 0.219*production
max	0.66*aid 0.583*sector 0.548*agriculture 0.371*agrarian 0.344*production
addFill	0.66*agriculture 0.421*aid 0.372*sector 0.307*farmer 0.237*agrarian 0.219*production
maxFill	0.66*aid 0.583*sector 0.548*agriculture 0.482*farmer 0.371*agrarian 0.344*production

But we find a problem in the previous process: when joining the different terms associated to different subjects, some of these terms are repeated (several subjects have terms in common, as `agriculture` in the example of Table 1). Since having repeated terms with different weights makes no sense, we consider the following approaches to fix the weights of these terms:

a) Add weights (add): collapse the repeated terms into one with a weight equals to the addition of the individuals weights.

b) Maximum between weights (max): we keep the repeated term with highest weight, removing the others.

c) Add weights, filling terms (addFill): same as *add*, but each time a term is deleted from a subject, the next one in the list is included until having *expTerms* terms for each subject.

d) Maximum between weights, filling terms (maxFill): same as *addFill*, but using maximum instead of sum.

The first two approaches involve that we do not always obtain the same number of terms for the personalization techniques, as it happens with the last two approaches. It should be noted that, in the last two approaches the filling process starts from the last *expSubj* subject, since we want more information from the most profile representative subject, i.e., the first *expSubj* subject. At the end of this process, the final terms will be also normalized with a maximum normalization value (*maxNorm*), with values: 0.33, 0.66, 0.99. The combination of the *expSubj*, *expTerms* and *maxNorm* gives us a total number of 36 different weighted term sets to provide to each personalization technique. Check Table 3 to see an example of these user profiles from Table 1.

4.1 Results

This section shows the results of the different experiments carried out, under the previous evaluation framework and considering the six proposed user profiles.

In Table 4, each cell represents the maximum NDCG value among the (12 or 36) different profile configurations, for each possible user profile and personalization technique. These values are the averaged NDCG values from the carried out user study 126 evaluation triplets.

Firstly, we can see that personalization helps the user to find relevant information, since in all cases we obtain significant improvements with respect to the non-personalized IRS performance (NDCG = 0.388). Depending on the profile and the personalization technique these improvements range from 50% to 80%.

The two main conclusions drawn from Table 4 are: 1) the best personalization technique is clearly *HRR+m*, and 2) the best user profile approach is *tProf* except in two cases, in which *stProf_maxFill* profile is better. Considering the last conclusion, we may assume that most of the times the best user profile approach to use is the simpler *tProf*, instead of the more complicated *stProf_maxFill*.

With respect to the used number of subjects or terms and the normalization value, the first row in Table 5 shows which profile configuration maximizes performance. We can see that the best user profile configuration is very homogeneous, independently of the user profile approach. This best user profile configuration is formed by $exp[Subj|Terms] = 40$, $expTerms = 10$ (in *stProf* profiles), and $maxNorm = 0.99$. Thus, it seems that the best user profiles are those including a rather high number of subjects and/or terms with high associated weights.

Also, the last two rows of Table 5 show the average and standard deviation for all the proposed user profile approaches. We observe that the highest average value is achieved by the *tProf* approach, with a low deviation value. Meanwhile, the lowest deviation value is achieved by the *sProf* approach, but with a much lower average value than *tProf*. Considering the four *stProf* approaches, we may observe a gradual decrease and increase in the average and deviation values, respectively, following the order of these profiles in the table. This situation indicates that within these user profiles, the further to the right in the table, they achieve more disparate personalization results (higher and lower), so more attention need to be paid to the selection of the right user profile configuration. The fact of having the maximum experimental evaluation performance with *stProf_maxFill* approach confirm this last conclusion.

Considering all the results, could it be concluded that we stand up for the *tProf* profile? Not necessarily, from a user perspective and considering not very small profiles, a *stProf* profile is much easier to understand than a *tProf* profile, since abstract concepts contain more semantics than isolated terms. It is also true that the *stProf* profile with two levels (concepts and terms) could be exploited by a given personalization technique to improve its performance, e.g., easily selecting parts of the user profile which suit more to the query (particularly helpful for heterogeneous profiles). Thus, depending on the application and the used personalization technique, a trade-off decision between pure performance or more expressiveness of the user profile must be taken. Additionally, from a cost

Table 4. NDCG maximum values from the 12-36 possible 'user profile-personalization technique' configurations. Original (non-personalized) NDCG value: 0.388. '*' character shows the best user profile approach for each personalization technique, and '+' character shows the best personalization technique for a given user profile approach.

	NQE	**HRR**	**SRR**	**IRR**	**NQE+m**	**HRR+m**	**SRR+m**	**IRR+m**	**CAS**	**CAS-or**
sProf	0.588	0.603	0.577	0.572	0.632	0.645^{+}	0.552	0.552	0.552	0.578
tProf	0.634*	0.652*	0.625*	0.620*	0.678	0.696^{+}	0.597*	0.597*	0.675*	0.668*
stProf_add	0.610	0.626	0.605	0.603	0.673	0.685^{+}	0.580	0.580	0.659	0.662
stProf_max	0.601	0.624	0.603	0.600	0.681	0.694^{+}	0.584	0.584	0.660	0.662
stProf_addFill	0.626	0.634	0.615	0.611	0.674	0.693^{+}	0.585	0.585	0.660	0.659
stProf_maxFill	0.612	0.633	0.606	0.602	0.683*	0.701^{+*}	0.587	0.587	0.658	0.660

Table 5. Best user profile configuration (*exp[Subj—Terms]-[expTerms]-maxNorm*), average and deviation, for each user profile approach.

	sProf	tProf	stProf_add	stProf_max	stProf_addFill	stProf_maxFill
Prof. conf. (max)	40-0.99	40-0.99	40-10-0.99	40-10-0.99	40-10-0.99	40-10-0.99
NDCG average	0.543	**0.602**	0.575	0.572	0.571	0.565
NDCG deviation	**0.032**	0.034	0.039	0.042	0.043	0.045

perspective these user profiles are not very demanding, since they only change with the inclusion of new documents, which does not happen very often.

5 Conclusions

In this work, we have presented 6 different user profile representation approaches based on content. Firstly, we focused on the subjects from a thesaurus, which are manually assigned to the initiatives in the documents by documentalists. These subjects are considered as concepts for the user profile based on *subjects*. Then, we did not take into account any other information than simply the document terms, to build the user profile based on *terms*. And finally, we proposed a hybrid approach between the two previous approaches (with four variations), having a two level user profile representation, where the first level is represented by *subjects* and the second level by the *terms* representing these subjects.

We have performed evaluation experiments including ten different personalization techniques and a wide range of user profile configurations, for all the proposed user profile approaches. We have obtained very good results, which in the best case reach up to 80.67% of improvement, with respect to the original non-personalized model. Additionally, we have demonstrated that most of the times the use of a simple user profile based on terms is enough to get good personalized results. Anyway, having a user profile with some structure and abstract concepts may help both, users to better understand their own profiles, and also some personalization techniques which may exploit this richer representation.

As future work, we would like to develop some personalization techniques to exploit the hierarchy of the proposed user profiles based on subjects and terms, and to use them to include personalization in privacy constrained environments.

Acknowledgements. Paper supported by the Spanish Andalusian "Consejería de Innovación, Ciencia y Empresa" and the "Ministerio de Ciencia e Innovación" under the projects P09-TIC-4526 and TIN2011-28538-C02-02, respectively.

References

1. J. Ahn, P. Brusilovsky, D. He, J. Grady, and Q. Li. Personalized Web Exploration with Task Models. Proceedings of the 17th International Conference on World Wide Web (WWW'08), 1-10, 2008.
2. J. Allan. Challenges in information retrieval and language modelling. In Report of a workshop held at the Center for Intelligent Information Retrieval, University of Massachusetts, Amherst, 2002.

3. N.J. Belkin. Some(what) Grand Challenges for Information Retrieval. SIGIR Forum 42(1): 47-54, 2008.
4. S. Calegari, and G. Pasi. Ontology-Based Information Behaviour to Improve Web Search. Future Internet 2(4): 533-558, 2010.
5. L.M. de Campos, J.M. Fernández-Luna, J.F. Huete, and A.E. Romero. Garnata: An information retrieval system for structured documents based on probabilistic graphical models. Proceedings of the 11th Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, 1024-1031, 2006.
6. L.M. de Campos, J.M. Fernández-Luna, J.F. Huete, C.J. Martin-Dancausa, C. Tur-Vigil and A. Tagua. An Integrated System for Managing the Andalusian Parliament's Digital Library. Program: Electronic Library and Information Systems 43(2): 121-139. 2009.
7. L.M. de Campos, J.M. Fernández-Luna, J.F. Huete, and E. Vicente-López. Using personalization to improve XML retrieval. IEEE Transactions on Knowledge and Data Engineering, to appear. DOI: 10.1109/TKDE.2013.75.
8. L. Chen, and K. Sycara. A Personal Agent for Browsing and Searching. Proceedings of the 2nd International Conference on Autonomous Agents, 132-139, 1998.
9. M. Daoud, L. Tamine, and M. Boughanem. Learning user interests for session-based personalized search. In ACM Information Interaction in conteXt (IIiX'08), 57-64, 2008.
10. J.F. Gantz. The Diverse and Exploding Digital Universe. IDC white paper (2): 1-16, 2008.
11. S. Gauch, M. Speretta, A. Chandramouli, and A. Micarelli. User profiles for personalized information access. The Adaptive Web: Methods and Strategies of Web Personalization, LNCS 4321, 54-89, 2007.
12. K. Jarvelin, and J. Kekalainen. Cumulative gain-based evaluation of IR techniques. ACM Transactions on Information Systems 20(4): 422-446, 2002.
13. A. Kobsa. Privacy-Enhanced Web Personalization. The Adaptive Web: Methods and Strategies of Web Personalization, LNCS 4321, 628-670, 2007.
14. E.L.C. Law, T. Klobučar, and M. Pipan. User effect in evaluating personalized information retrieval systems. LNCS, 4227, 257-271, 2006.
15. A. Micarelli, and F. Sciarrone. Anatomy and Empirical Evaluation of an Adaptive Web-Based Information Filtering System. User Modeling and User-Adapted Interaction 14(2-3): 159-200, 2004.
16. A. Moukas. Amalthaea: Information Discovery And Filtering Using A Multiagent Evolving Ecosystem. Applied Artificial Intelligence 11(5): 437-457, 1997.
17. G. Pasi. Issues in Personalizing Information Retrieval. IEEE Intelligent Informatics Bulletin 11(1): 3-7, 2010.
18. A. Sieg, B. Mobasher, and R. Burke. Web search personalization with ontological user profiles. Proceedings of the 16th ACM International Conference on Information and Knowledge Management, 525-534, 2007.
19. A. Spink, S. Ozmutlu, H.C. Ozmutlu, and B.J. Jansen. US versus european web searching trends. Proceedings of the 25th ACM SIGIR conference on research and development in information retrieval, 32-38, 2002.
20. K. Sugiyama, K. Hatano, and M. Yoshikawa. Adaptive web search based on user profile constructed without any effort from users. Proceedings of the 13th International Conference on World Wide Web, 675-684, 2004.
21. J. Trajkova, and S. Gauch. Improving Ontology-Based User Profiles. Proceedings of the RIAO 2004 Conference, 380-390, 2004.
22. G. Uchyigit, and M.Y. Ma. Personalization Techniques and Recommender Systems. World Scientific, 2008.