Comparative evaluation of personalization algorithms for content recommendation

Carlos R. C. Alves, Lúcia V. L. Filgueiras

LTS – Departamento de Engenharia de Computação e Sistemas Digitais – Escola Politécnica da Universidade de São Paulo – Av. Luciano Gualberto, 158, trav. 3 – 05508-900

Abstract. Personalization techniques that combine user characteristics, user behavior, and content organization can be used to help users on finding objectively content on the web. The main contribution of this text is the multidisciplinary study that was conducted integrating different areas on human knowledge in order to find the best way to direct content, including some wide research on personalization concepts and applications. This study also presents the development of the Argo software which is formed by a web site, a component that captures and stores information about the user navigation, and three different personalization algorithms. Using navigation data it is possible to generate user profile, which is used to recommend content. Tests were conducted to check efficiency of the personalization algorithms.

1 Introduction

Nowadays there is a huge amount of available information on the Internet. This growth increases difficulty for users on their task of content search. Within a particular web site the content may be organized, following its author's point of view [11]. Quite often this perspective is different from the users', making their objectives even harder to achieve. Personalization is applied in these situations, preserving one of the best characteristics found on content search: the user individual and unique experience.

We have developed the Argo system, in order to answer three questions: what are the factors that characterize a user and his/her behavior, during his/her navigation experience; how these factors interact, defining users profiles; and what is the influence of content organization over these profiles.

In this paper, section 2 presents personalization concept in its different ways, as found on technical literature. Section 3 and 4 describe how this concept may be represented and turned into computational variables. The way to combine these variables is shown in section 5, as the three different personalization algorithms developed by the authors. Section 6 presents the proposed test environment, used to compare and verify efficiency of the algorithms. Section 7 brings evaluation results and section 8 contains conclusions.

R. C. Alves C. and V. L. Filgueiras L. (2005).
Comparative evaluation of personalization algorithms for content recommendation.
In *Proceedings of the 1st International Workshop on Web Personalisation, Recommender Systems and Intelligent User Interfaces*, pages 56-65
DOI: 10.5220/0001421900560065

2 Personalization concept

There are several approaches to face personalization to the web [17][19][20]. On technical literature it is usual to find basically three kinds of definitions. The first one is user based, i.e., personalization is a way to capture behavior patterns and interests from the user, based on his/her navigation experience. Mobasher et al. [15] say it is based on modifying user's experience based on his/her preferences.
Another group defines personalization as content based. This means that the organized content within a web site is the base to direct it to users. Finally there is the hybrid definition that merges user information with content organization, making up an integrated base of knowledge about users attached to a determined context (i.e., the web site).
For this work, it is found that Eirinaki et al. [8] definition to reflect better the author's application of personalization, as any action that can adapt information to the user. It is accomplished combining user behavior and content structure.

3 Personalization application

Based on definitions given on section 2, the authors decided to work in a system that could recommend pre-organized content to the user based on his/her preferences.
For this work Argo system has been developed, that is a software environment able to apply personalization concept on an engineering articles magazine (named Revista Politécnica), which has 138 articles grouped into 108 content categories.
To reach a better understanding of which factors could be considered on this task, a multidisciplinary study was conducted, as shown in figure 1.

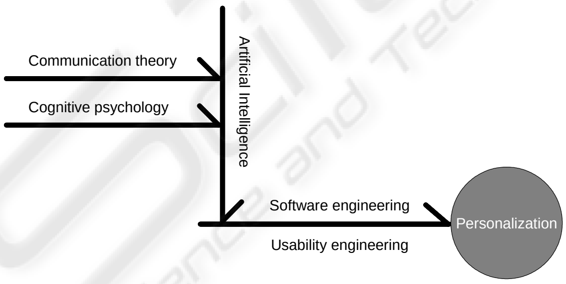

Fig. 1. Interaction representation among knowledge areas studied to apply personalization

Using study results from communication theory and cognitive psychology it is possible to establish factors that define user's profiles, his/her way of thinking and the way that information must be transmitted to communicate more efficiently. These factors can be converted into computational variables using artificial intelligence techniques, as it presents mathematical models for them. Based on software and usability engineering methods, it was possible to develop a system to test these factors, which must represent the user, the content and their interaction.
The user is characterized as being adaptable, i.e., he/she can perform within different tasks and evolve (change) in time. Usually other works found on technical literature

[6][20] generate users profiles based on demographic researches. For this work it is said that a user profile is a representation of his/her behavior due to the task of content search, based on navigation, within a web site. Some related works were used to develop the way to obtain user information [6], such as using a proxy [7], test methods [21], and system architecture [1].

Content is characterized based on its organization, its meaning, and context. For this work, net structure was used, defining parent, child, and jump relations between content nodes. Some works in literature are based on content organization [5][15], and information that was directly used on Argo, as using crossed-references among nodes [18] and giving points due to content relevance, based on its semantic relationship [12].

3.1 User-content integration

It is necessary to use the context where user is into when recommending content. This context is represented by user and organized content metadata. User data is based on his/her identification (such as login or cookie) and his/her metadata are obtained through his/her behavior while navigating on the web site. Content data are the magazine's articles, and metadata is their semantic relationship, represented by content categories. Based on [19], Figure 2 brings a general view of personalization techniques used in Argo (marked as dark boxes).

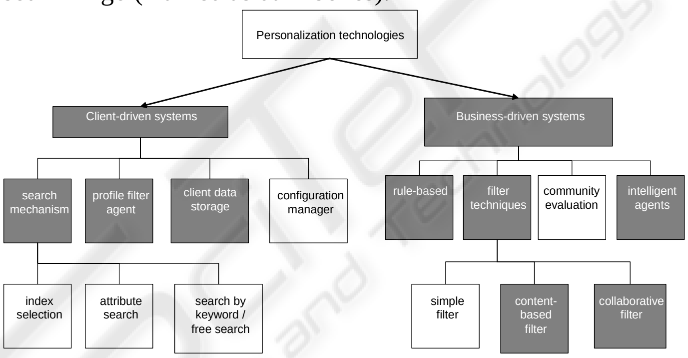

Fig. 2. Used technologies for web personalization, based on [19]

The technologies are divided in two groups: client-driven (user) and business-driven (content). Search mechanism is based on key words that are input in a query by users. Profile filter agents use information about groups of users to drive content based on them. Data storage may contain information such as user behavior during a session, using logs or online processing. Content-based filter behaves the same way profile filter agent does, but it uses information on content organization. Collaborative filtering [13] uses opinions (as marks or concepts) of other users about a specific content node, due to its relevance.

Intelligent agents is the technique that is being more commonly used [2][3][4][13] due to its ability to implement adaptability, i.e., to generate environments that are able to store information about users and combine them with content organization, dynamically.

4 Personalization factors

The implementation of personalization concept is necessary to follow four basic steps, represented on figure 3.

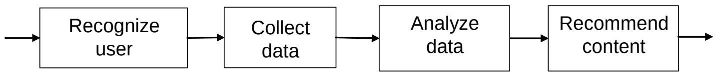

Fig. 3. Steps to implement personalization concept

The user is recognized by cookies or login. Data collection is done extracting navigation information to form users' profiles. Data analysis is conducted, in this work, with three different algorithms combining user profile and content organization. Content recommendation is the result of the process, delivering content to a specific user based on his/her preferences.
To establish data analysis it is necessary to define what piece of information will be collected, from the user experience with the web site. Figure 4 represents the interaction among personalization factors.

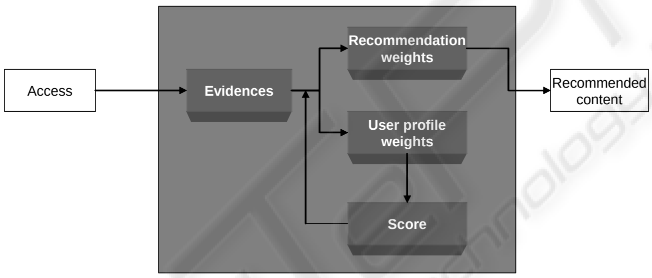

Fig. 4. Information flux through personalization algorithms

Evidences are directly related to information that may be captured during each one of user navigation on the web site. These factors are based on the context in which the user is being held:

- Accessed content: article that has been accessed.
- Accessed meta-content: which category of content has been accessed.

User profile weights create a historic base, i.e., characterize user and his/her preferences according to access done in the web site. These factors are defined based on cognitive aspects of the user.

- Access time (Ta): time spent by the user to read an article. Using fuzzy logic, z is a parameter defined from experiments found in [16], representing a time mark where the extension of the access is cut off.

$$\mu_{Ta} = \begin{cases} \sqrt{\dfrac{Ta}{z}} & \text{, when } Ta < z \\ 1 & \text{, when } Ta \geq z \end{cases} \qquad (1)$$

- Access chain (Cc): access sequence of the articles and categories read by a user, during a session [5][17]. It is used data from other users accesses, based on the longest one. Fuzzy logic was used to represent the sequences to try to predict user's next step. In this work, *c*=4.

$$\mu_{Cc} = \begin{cases} \left(\dfrac{Cc}{c}\right)^2 & \text{, when } Cc < c \\ 1 & \text{, when } Cc \geq c \end{cases} \tag{2}$$

- Source (Or): content node from which the user reached the actual one. It is represented using a probability distribution.
  – Recommendation (*Or(r)*=0.6): access to recommended content by system.
  – Structure (*Or(e)*=0.1): based on the web site menu.
  – Search (*Or(b)*=0.3): content reached from a search result.
- Source order (Oo): which list item that had been accessed. It applies only to recommendation and search. Based on [14], a variance is established according to the relative importance due to the way it was accessed. The factor is represented by using a probability distribution:
  – *Oo(s)*=0.90: for selected and viewed article.
  – *Oo(i)*=0.09: for an ignored article (e.g.: articles there are below the selected one on a list).
  – *Oo(k)*=0.01: for skipped articles (e.g.: when the third one of a list is selected, the first two are considered skipped).

Recommendation weights are used to establish which content node would be presented to the user, based on its organization.

- Content relation (Rc): relation between the actual accessed content and all other content nodes, based on the net structured content. It is represented by a probabilistic distribution:
  – *Rc(∅)*=0.0: no relation.
  – *Rc(p)*=0.1: actual content is parent of the recommended one.
  – *Rc(f)*=0.2: actual content is child of the recommended one.
  – *Rc(i)*=0.35: actual content is sibling of the recommended one.
  – *Rc(j)*=0.35: actual content is jump of the recommended one.
- Content node distance (Dn): net length between actual accessed content and the other content nodes, present on original content structure. Dijkstra algorithm was used to calculate the optimal path between content nodes. To determine the factor, fuzzy logic was used establishing relevance between content and user profile. Parameter *h* was determined by experiences found on related work [12]. In this work *h*=10, to limit length search on the net structure.

$$\mu_{Dn} = \begin{cases} -\dfrac{Dn}{h} + 1 & \text{, when } Dn < h \\ 0 & \text{, when } Dn \geq h \end{cases} \tag{3}$$

A control to differ content recommendation calculation was created. The trigger (Ga) defines a point from which the recommendation weights would be calculated, i.e., the number of accesses of a specific user determines his/her experience on using the web site. From similar experiments presented in literature [11], *Ga*>10 accesses was used.
The score (Pl) represents the result of the algorithms for content recommendation, integrating users profiles and content nodes. This calculation uses a feedback mechanism to consider users history and evolution in time. So it is able to combine factors found during current user navigation and his/her history (profile).

- User profile: a new access (Na) is defined by the navigation chain of the user, during a session.

$$Na = Ta \times Cc \times Or \times Oo \quad \textbf{(4)}$$

  Calculating again, to consider user's history:

$$Pl(t+1) = Pl(t) \times Na \quad \textbf{(5)}$$

- Content recommendation: if the trigger is activated (*Ga*>10), score is calculated to all content nodes on the database, except for the one that has been accessed, considering user's history also.

$$Pl(t+1) = Rc \times Dn \times Pl(t) \quad \textbf{(6)}$$

The following section presents three different ways to calculate these shown equations, i.e., how symbol × may be replaced by artificial intelligence techniques.

5 Personalization algorithm

Three different algorithms were developed to verify which artificial intelligence techniques would be more adequate to personalization solutions; to build a system that may be implemented in different environments, since it is developed with components; to obtain results on a comparative way to check their efficiency. The algorithms combine different factors described on section 4, to form users' profiles.

- Bayes (BY): using conditional probabilities theory from Bayes [9] it was possible to combine the factors by multiplication.
- Certain factors (CF): the values assumed by factors are used as weights of such certain factors [9]. Calculation is done as this equation, considering feedback to adjust its value:

$$CF_{recalc}(CF_1, CF_2) = Pl(t+1) = CF_1 + CF_2(1 - CF_1) = Pl(t) + Pl(t+1)\,(1 - Pl(t)) \quad \textbf{(7)}$$

- And fuzzy (FZ): it is applied to the AND fuzzy operation to combine the variables [4]. It was adapted because in this work the variables may value zero, so it was introduced a conditional term. The following equations were used for user profiling and content recommendation:

$$\begin{aligned}&\text{If } Pl{=}0, \text{ then } Pl(t+1) = Na * Pl(t) + 1 - Na\\ &\text{else, } Pl(t+1) = Na * Pl(t) + 1 - Pl(t)\end{aligned} \quad \textbf{(8)}$$

6 Test environment

Only users within the target audience of the web site were considered, i.e., engineers. Behavior profile is determined by navigation of the users through the site. Users were recruited by e-mail to participate on final tests phase, and data were collected from ten of them. This number is considered valid, according to Nielsen [apud 10].
Each user received ten articles to be ordained according to his/her preferences. The answers were collected five days later.
The three algorithms were run over the same content sent to the users. Their results were evaluated against the lists ordained by users, using scoring criteria that should aim 30% of equivalence between users' lists and algorithm classification.

7 Results and discussions

Results were evaluated on two aspects: statistical analysis, checking values obtained for factors; and efficiency test, to compare the algorithms.

7.1 Statistical analysis

When users navigated through the web site, data was collected to form their profiles. This evaluation was done using four criteria:

- Absolute value of scores.
- Relative value of the difference between maximum and minimal scores.
- Algorithms similarity.
- Ability to classify direct hits (that are volunteer accesses from a user to a specific content node).

As shown in table 1, CF and FZ presented a larger absolute difference between results. This represents the trust on algorithms responses, because the interval of scores is used to differentiate among content nodes of the web site.

Table 1. Score values obtained by the personalization algorithms

Algorithm	Maximum score (Pl_{max})	Minimum score (Pl_{min})	Absolute difference (Pl_{max} - Pl_{min})	Relative difference (1 - Pl_{min} / Pl_{max})
BY	0.009	0.001	0.008	89.0%
CF	1.000	0.184	0.816	81.6%
FZ	0.904	0.030	0.874	96.7%

Analyzed results indicated that BY scores tend to 0, while CF's tend to 1. This is obtained due the calculation structure of these algorithms, based on multiplications.
Another statistical analysis was conducted comparing obtained scores from each algorithm, presented on table 2.

Table 2. Descriptive statistics study over obtained scores by each algorithm

	BY	CF	FZ
Mean	0.006	0.805	0.563
Standard variation	0.002	0.234	0.192
Variance	$4.6 \cdot 10^{-6}$	0.055	0.037
Amplitude	0.008	0.808	0.710

Sample distribution does not correspond to common probabilistic functions (such as normal, t-student, or others). However these data confirm that BY algorithm concentrate its scores in a range smaller than 0.10. CF and FZ present a larger distribution, with closer results. If the variance is high, difference between scores of content nodes is clearer. As CF mean is greater than 0.50, FZ is shown more adequate. This statistical distribution shows the ability of the algorithm to differentiate among content nodes, to recommend to the user.

Algorithm similarity was verified when trying to separate scores within the ten articles sent to users. Obtained results were equivalent for all three algorithms, and it was not possible to put in order the preference list in order to compare it with user sent data. For this reason a new test was conducted, as seen in next section.

Direct hits were masked on final score evaluation, because the use of a feedback mechanism. It was not possible to determine on this proposed testing condition if direct hits increase scores, however it was verified when evaluating individual equations to score calculation.

7.2 Efficiency test

Continuing algorithms evaluation, a new test was conducted by selecting six users that accessed more the web site. Three categories and three articles were sent to each of them, which should be classified by his/her preferences. To compare the lists answered by users and those obtained by the three algorithms, the following points criteria was used: 0, if were all wrong; 1, if the second one is right; 2, for one right (being first or third on the list, once it demonstrates algorithm and user’s position in list for the node is the same); 6, for all right. Table 3 presents these results.

Table 3. Points obtained by test sending articles and categories

User	points (article)	points (category)
10	2	6
11	1	6
5	2	6
8	6	2
2	6	6
9	0	6

Using the same criteria of 30% (points greater than or equal to 2), it was obtained 83% (10 out of 12) of accuracy on these tests. It confirms that the algorithms are efficient for personalization.

8 Conclusions

Considering questions presented on section 1, it had been proven that personalization factors developed in this work represent user profile related to what he/she does, i.e., his/her behavior. Based on experimental results, the proposed way on combining factors is coherent, due to similarities found among the three algorithms. It is possible to verify that content organization, defined by the site's author, interferes on scores. However the users' point of view is also considered when integrating to score calculation factors related to his/her behavior. In summary, it is necessary to join user profile defined by his/her behavior and content organization, that represents the context in which the user is inserted when trying to perform his/her tasks.

The use of a feedback mechanism to calculate scores presents benefits to final results, because user history, i.e., past accesses, acts as a weight on pondering actual access and system database, containing information on user behavior.

Artificial intelligence techniques allow to combine personalization factors that can be measured as numeric results, represented by the scores (Pl). These can be used to differentiate content nodes based on users' preferences.

On algorithms evaluation, FZ has proven to be more adequate to be used on personalization application environment, because it presented greater variance, allowing differentiating among content nodes on recommendation.

On research method, combining software and usability engineering methods was positive. Multidisciplinary research proved to be a determinant tool on analyzing on depth factors that contribute to personalization and their interaction. Using disciplines as communication theory and cognitive psychology enhance knowledge on users and their way of interacting with the environment.

9 Future works

It is necessary to dissociate content structure proposed by the web site author to its presentation to users. It can be done applying human-computer interaction techniques to modify dynamically the interface.

It is possible to enhance tests scope with other algorithms, due the modular structure of this system development. Numeric calculus may be applied to simulate and to obtain partial results. To improve content organization it must be used onthologies and data mining techniques.

Acknowledgements

Authors thank Thomas Ufer, Mauricio de Diana and Alexandre Veiga Gimenes for their collaboration on work development; and to professors Junia A. Silva and Selma S. S. Melnikoff for their valuable reviews.

References

1. Adomavicius, G., Tuzhilin, A. Using data mining methods to build customer profiles. *Computer 34, 2* (2001), 74-82.
2. Albanese, M., Picariello, A., Sansone, C., Sansone, L. Web personalization: Web personalization based on static information and dynamic user behavior. *Proceedings of the 6th annual ACM international workshop on Web information and data management* (2004), 80-87.
3. Ardissono, L., Goy, A., Petrone, G., Segnan, M. A multi-agent infrastructure for developing personalized web-based systems. *ACM Transactions on Internet Technology – TOIT 5, 1* (2005), 47-69.
4. Ardissono, L. et al. Intrigue: personalized recommendation of tourist attractions for desktop and handset devices. *Applied Artificial Intelligence* (2002).
5. Bauer, T., Leake, D. Using document access sequences to recommend customized information. *IEEE Intelligent Systems 17, 6* (2002), 27-33.
6. Cannataro, M., Cuzzocrea, A., Pugliese, A. A probabilistic adaptive hypermedia system. *Proceedings of international conference on information technology: coding and computing* (2001), 411-415.
7. Chen, C.C., Chen, M.C., Sun, Y. PVA: a self-adaptive personal view agent system. *Proceedings of ACM SIGKDD international conference on knowledge discovery and data mining* (2001), 257-262.
8. Eirinaki, M., Charalampos, L., Stratos, P., Vazirgiannis. Web personalization integrating content semantics and navigational patterns. *Proceedings of the 6th annual ACM international workshop on Web information and data management* (2004), 72-79.
9. Genesereth, M.R., Nilsson, N.J. *Logical foundations of artificial intelligence*. Morgan Kauffman Publishers (1987).
10. Hackos, J.T., Redish, J.C. *User and task analysis for interface design*. John Wiley & Sons (1998).
11. Hölscher, C., Strube, G. Web search behavior of internet experts and newbies. *Computer networks 33* (2000), 337-346.
12. Jokela S. et al. The role of structured content in personalized news service. *Proceedings of the 34th Hawaii International conference on systems sciences* (2001).
13. Lin, R., Kraus, S., Tew, J. OSGS-A personalized online store for e-commerce environments. *Information retrieval 7*, Kluwer Academic Publishers (2004), 369-394.
14. Menczer, F., Street, W.N., Monge, A.E. Adaptive assistants for customized e-shopping. *IEEE Intelligent Systems 17, 6* (2002), 12-19.
15. Mobasher B. et al. Integrating web usage and content mining for more effective personalization. *Electronic commerce and web technologies 1875* (2000), 165-176.
16. Padmanabhan, B., Zheng, Z., Kimbrough, S.O. Personalization from incomplete data: what do you don't know can hurt. *ACM SIGKDD international conference on knowledge discovery and data mining* (2001), 154-163.
17. Perkowitz, M., Etzioni, O. Towards adaptive web sites: conceptual framework and case study. *Computer networks 31* (1999), 1245-1258.
18. Ramakrishnan, N. PIPE: Web personalization by partial evaluation. *IEEE internet computing 4, 6* (2000), 21-31.
19. Sae-Tang, S., Esichaikul, V. Web personalization techniques for e-commerce. *Proceedings of the 6th international computer science conference*. Springer-Verlag (2001), 36-44.
20. Smyth, B., Cotter, P. A personalized television listings service. *Communications of the ACM 43, 8* (2000), 107-111.
21. Soltysiak, S.J., Crabtree, I.B. Automatic learning of user profiles – towards the personalization of agent services. *BT Technology Journal 16, 3* (1998), 110-117.