

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

How to Communicate Recommendations? Evaluation of an Adaptive Annotation Technique

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/92410> since 2018-03-08T14:29:00Z

Publisher:

Springer

Published version:

DOI:10.1007/11555261_97

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

How to communicate recommendations? Evaluation of an adaptive annotation technique

Federica Cena^{1,2}, Cristina Gena², Sonia Modeo^{1,2}

¹ CSP Innovazione nelle ICT s.c. a r.l., Torino, Italy
{modeo,cena}@csp.it

² Dip. di Informatica, Università di Torino, Torino, Italy
cgena@di.unito.it

Abstract. In this paper we present an evaluation of an adaptive annotation technique (the use of icons to help the user in the selection of the most relevant suggested item), using the Grounded Theory methodology. The goal of the evaluation was to find out the best icon in order to communicate the system recommendations in the most effective way.

1. Introduction

Information overload is one of the most serious problems that user suffers in every day life, especially during Internet navigation. Recommender systems help the user to choose within this large amount of information. This could be especially helpful in a mobile context where the interaction can be very challenging for the user due to: device limitations (display, battery, connection, I/O devices, etc.), and difficulties of interaction (movement, noisy or badly illuminated environment, etc.). These difficulties requires not only the personalization of the user interaction with the system [10][2], but also the communication of the personalization in the most effective way, as the user must be always aware of the personalization features of the system.

We faced these problems during the development of UbiquiTO, a multi-device mobile tourist guide that provides recommendations adapted to the individual user, her location, her device and the current context conditions (for more details see [1]). Usually, in adaptive systems, recommended items are emphasized with adaptive annotation, a navigation support that consists in attaching various visual cues (e.g., enlarging font, changing color, adding icons, etc.) to the suggested items in order to help the user to select the most relevant one [3]. Empirical studies of adaptive annotation in educational context demonstrate that it could reduce navigation overhead and improve learning activity [4]. Adaptive annotation is also largely exploited in recommender systems (see, for example, [5]), even if there is a lack of empirical studies on the correct way to annotate items (one of them is [6]).

2. Testing and Results

We decided to communicate the UbiquiTO recommendations by means of the adaptive annotation technique. In particular, we decided to add meaningful icons to every suggested item. In order to choose the best kind of icon, we presented the icons commonly used in adaptive systems to real users: (Fig. 1): *traffic light circles*, which are mainly used in adaptive learning systems [4]; *stars*, which are mainly used in recommender systems to express a quantitative assessment and two other less exploited icons (*circles* and *asterisks*) sharing a similar meaning; *emoticons*, which have been used both in recommender and in adaptive learning systems; *full-half full-empty colour*, which are often added to enforce the meaning of adaptive icons [6]. We also presented to the subjects a seventh choice: a percentage associated to each item since our previous prototype used this option. Notice that we used well-known and common icons to facilitate user's recognition rather than recall [7]. Thus, we did not consider creative icons from commercial systems in order to increase the usability of our interface: icons have to be easily recognizable and understandable. The user does not have to think too much working out from memory what an icon is about.

In particular, the goal of the evaluation were to find out: *i*) if the users understand the way the system communicates the strength of recommendations; *ii*) which is the most appropriate symbol for representing the system recommendations; *iii*) which meaning the users attribute to each group of symbols.

In order to evaluate point *i*) we provided the subjects with a set of PDA screenshots with the symbols associated to a non-significant text ("foo"), equal in every screenshot. Each screenshot contained the 3 symbols belonging to the same group (Fig. 2), ordered in a randomized way. We asked the subjects to choose "the item you think the system is recommending as the best one for you". For instance, in the emoticon group, the smiling face suggests the best choice, while in case of stars the three stars are the right choice, and so on. To evaluate point *ii*), we asked the subjects which group of symbols best represents the system personalization and why. Finally, to test point *iii*) we asked the subjects to associate a meaning to every symbol.

We involved in the evaluation 34 subjects, 20 males and 14 females, 25-34 aged, with different types of background and occupation, and being familiar with computer, Internet and new devices (PDA, smartphone, Digital Terrestrial Television). Regarding point *i*), subjects correctly understood the way the system communicates the strength of recommendations: in fact, most of the times they chose the symbols we associated to the best recommendations (69%). They rarely chose symbols referring to the recommendations with medium strength (15%) and worst strength (16%) and the differences were significant ($\chi^2(66)=129.83$, $p<0.001$). The type of symbol (*stars*, *balls*, *emoticons*, etc.) does not influence in a significant way the choices of the best items ($\chi^2(10)=13.92$). Concerning the evaluation of point *ii*), the group of symbols best representing recommendations were the *stars* (chosen by 18 subjects), followed by *emoticons* (6 choices), *traffic light circles* (5 choices), *full-half full-empty circles* (2 choices), *asterisks* (1 choice), *percentage* (1 choice), and *circles* (1) and the differences were significant ($\chi^2(6)=54.5$, $p<0.001$). Finally, to evaluate the meaning the subjects associate to the symbols (point *iii*), we related these results to the subjects' explanations about the symbols best representing the system personalization, to find

possible correlations. Since we gained both quantitative and qualitative results, we decided to apply the Grounded Theory methodology [11], where collected data may be qualitative or quantitative or a combination of both types, since an interplay between qualitative and quantitative methods is advocated. Moreover, as a recent study pointed out [8], statistical analyses are often false, misleading and too narrow, while insights and qualitative studies do not suffer from these problems as they strictly rely to the users' observed behavior and reactions. Even if our quantitative analysis reported significant results, the actual preferences of the users could be different.

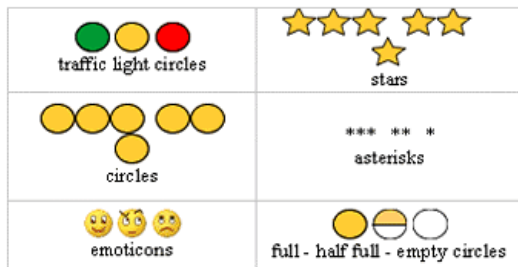


Fig. 1 - The evaluated icons groups.

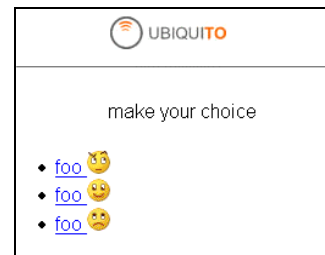


Fig. 2 – A PDA screenshot.

Three phases of data analysis are involved in the Grounded Theory. In the **open coding** phase we identified the concepts the subjects associate to the symbols and their properties (*traffic light circles* = user actions, *stars* = qualitative assessment, *circles* = quantitative assessment, *asterisks* = context-dependent meaning, *emoticons* = system emotion and opinion, *full-half full-empty circles* = quantity, *percentage* = numeric estimate). In the **axial coding** phase we related categories to their subcategories, linking categories at the level of properties and dimensions. For instance, we noticed that *circles* and *stars* share the same category (assessment) by considering different points of views (respectively quantitative and qualitative). In the **selective coding** phase we phrased the theory: the core category that best represents the communication of personalized suggestions is the "non-verbal communication", since the non-verbal hints express those emotional states necessary to make the system-user communication as personalized as possible. In our context we have to disambiguate that the system suggestions are not absolutely the best ones, but they are the best ones for a specific user, and emoticons are able to express feelings that make the interaction more human-like.

3. Discussions

The analysis carried on showed that the symbols that best represent emotional states are the emoticons: they represent opinions and emotions the system aims at communicating to the users. As it can be noticed, the last finding is in contrast to the quantitative evaluation results, which showed that stars are the best symbols for recommendations. Stars usually express a qualitative evaluation associated to an expert opinion or to a general assessment (e.g., 3 stars associated to a restaurant mean that the restau-

rant received good rates by a somewhat guide, not that it is good for me). When users explained their choices, they associated stars to a general qualitative judgment, while they related emoticons to feelings expressed by the system itself. In addition, the usual context where stars are used (e.g., newspapers, books, etc.) is different from a personal interaction, whereas emoticons are able to express positive, neutral or negative feelings. Even if some web sites, such as Amazon, use stars to communicate personalization, our subjects seem to have some difficulties in associating them to a personalization concept. We think that the reported quantitative results oversimplified the discussion about users choices, thus we gave more weight to qualitative results emerged from subjects explanations.

Therefore, we have used emoticons in UbiquiTO, as they seem to be able to communicate personalized messages. Moreover, in a small device with limited screen capabilities, the exploitation of emoticons allows us to use a single icon instead of a group of icons, thus reducing the amount of required space. Last but not least, our results are similar to the “The Media Equation” theory [9], which claims that as humans treat computer socially, the main goal of interface design has to be the replication of human-human communication.

References

1. I. Amendola, F. Cena, L. Console, A. Crevola, C. Gena, A. Goy, S. Modeo, M. Perrero, I. Torre, and A. Toso. UbiquiTO: a Multi-Device Adaptive Guide. In *the Proceedings of conference MobileHCI'04*, Sept. 13–16, Glasgow, Scotland (2004) 409-114.
2. D. Billsus, C. Brunk, C. Evans, B. Gladish, and M. Pazzani. Adaptive Interfaces for Ubiquitous Web Access. *Comm. of the ACM* 45(5), 2003, pp. 34-38.
3. P. Brusilovsky. Adaptive Hypermedia. *UMUAI 11* (1-2): 2001, pp 87-110
4. P. Brusilovsky, S. Sosnovsky, and O. Shcherbinina. QuizGuide: Increasing the Educational Value of Individualized Self-Assessment Quizzes with Adaptive Navigation Support. In the *Proceedings of World Conference on E-Learning*, E-Learn 2004, Washington, DC, USA, November 1-5, 2004, pp. 1806-1813
5. GroupLens Research Projects, <http://www.grouplens.org>, 2005.
6. S. M. McNee, K.S. Lam, C. Guetzlaff, J. A. Konstan, and J. Riedl. Confidence Displays and Training in Recommender Systems. In the *Proceedings of INTERACT '03 IFIP TC13 International Conference on Human-Computer Interaction*, 2003, pp. 176-183.
7. J. Nielsen. Usability Engineering. Boston, MA. Academic Press, 1993.
8. J. Nielsen. Risks of Qualitative Studies. <http://www.useit.com/alertbox/20040301.html>, 2004.
9. B. Reeves and C. Nass. The Media Equation: How People Treat Computers, Televisions, and New Media as Real People and Places. Cambridge University Press, New York, 1996.
10. B. Smyth and P. Cotter. The Plight of the Navigator: Solving the Navigation Problem for Wireless Portals. In the *Proceedings of the 2nd International Conference on Adaptive Hypermedia and Adaptive Web Systems*. Malaga, Spain, 2002, pp. 328-337.
11. A. L. Strauss and J. M. Corbin. Basics of qualitative research: techniques and procedures for developing grounded theory. SAGE, Thousand Oaks, 1998.