HAL
archives-ouvertes.fr

Collecting Users Profiles for Web Applications

Amin Rasooli, Peter Forbrig, Fattaneh Tagivareh

▶ **To cite this version:**

Amin Rasooli, Peter Forbrig, Fattaneh Tagivareh. Collecting Users Profiles for Web Applications. Marco Winckler; Peter Forbrig; Regina Bernhaupt. 4th International Conference on Human-Centered Software Engineering (HCSE), Oct 2012, Toulouse, France. Springer, Lecture Notes in Computer Science, LNCS-7623, pp.275-282, 2012, Human-Centered Software Engineering. <10.1007/978-3-642-34347-6_18>. <hal-01556822>

HAL Id: hal-01556822
https://hal.inria.fr/hal-01556822

Submitted on 5 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution| 4.0 International License

Collecting Users Profiles for Web Applications

Amin Rasooli[1,2], Peter Forbrig [1] and Fattaneh Tagivareh [2]
[1] University of Rostock, Computer Science Department, Albert-Einstein-Str. 21,
18051 Rostock, Germany
{amin.rasooli, peter.forbrig}@uni-rostock.de
[2] School of Electrical and Computer engineering, University of Tehran, Tehran, Iran
{rasooly, ftaghiyar}@ut.ac.ir

Abstract. Currently providers are trying to personalize their websites according to user profiles. With respect to the wide variety and great volume of websites, providers look for a design that is more attractive than that of competitors. They look for a unique solution. In this uniqueness, any point such as design, user-friendliness, and content offered to the customer plays a key role in its success. The main objective of this study is to provide profiles of different kinds of users. Later on, this information can be used to design appropriate websites. This kind of information can be explored from social networks. We obtained a dataset of 500 users and we have clustered this dataset to 12 clusters, and then applied Collaborative Filtering on user data to improve the results. The paper will present the corresponding results and provide an interesting overview of different profiles of users in different parts of the world.

Keywords: User Profiles, Web Application, Personalization, Web Design.

1. Introduction

Most of the time, standard solutions provided by web design techniques have limited personalization ability. Owners would like to have their website usable, aesthetic and attractive according to their personal feeling. Any point such as design and user friendliness, content, services offered to the customer and etc. can play a significant role to make a website succeeds. The first impression the user gets of a website is through its appearance. She/he may like the website at his/her first view and finally be loyal to it while churn probability will be minimized. Many modern techniques have been presented in data personalization in order to collect behavioral information of the users; for example: suspension time in user interface, ratio of the observed links to the total number of links and user interaction on any accessible data [1], [2]. These new inputs can be employed to understand the hidden interests of users in order to choose the best personalized data for them. We started by introducing the goals of our work and now we present our experiment in Section 2. Finally, we will summarize the main contributions of our work in Section 3. Because of lack of space we focus on our experiment. A discussion of related work can be found in [3].

2. Experiment

Based on the analysis of popular web sites different designs were collected. These designs have been clustered in 30 different types. These types were presented to users of social networks and their votes were collected according to the following rules.
Three choices were displayed to a user. After a user selected his/her favorite design a new page provided more choices.

The system had announced to the users that the first choice they observe is not necessarily an optimized one and they have to declare their vote without looking at the presented order. 58% of the users have voted to the first choice, 26% to the second one and remaining 16% to the third choice. The software contains a learning system in such a way that it updates the design rating considering the positive vote and the user's cluster. Therefore, when the user gives his/her positive vote to a design, the ratio of votes given to a specific design toward the total votes was added to its rating. This value divided by remaining designs in a group will also be subtracted from the total number of designs. Similarly, designs will tend to be optimized which will result in a powerful learning system.

The total number of the participants was 481 with 840 valid votes for the existing design templates, while 664 additional votes were added to first level votes taking into account the details.

3. Results

3.1. Color Preferences

Colors with darker tones such as blue and brown were more attractive to men, while designs having a color spectra such as pink, purple, red and yellow, in comparison with other colors, were a prevalent trend among women's choices. Designs with the desired tone for men were 48% of total designs, 67% of which were voted as the first choice and the remaining 33% were voted as second choices. Designs with the desired tone for women were 30% of total designs, 82% of which were selected as the first choice while the other 18% were voted as secondary choices. The sampling set was composed of 268 male and 192 female participants. Designs were also tagged such that they were sorted in four groups from the darkest to the lightest tone. It was observed that women show more interest in lighter colors, whereas men show more interest in darker colors. In Fig. 1, the horizontal axis of the diagram shows the average dark to light approximate tone scaled from 1 to 4 and the vertical axis of the diagram indicates the redundancy of each group.

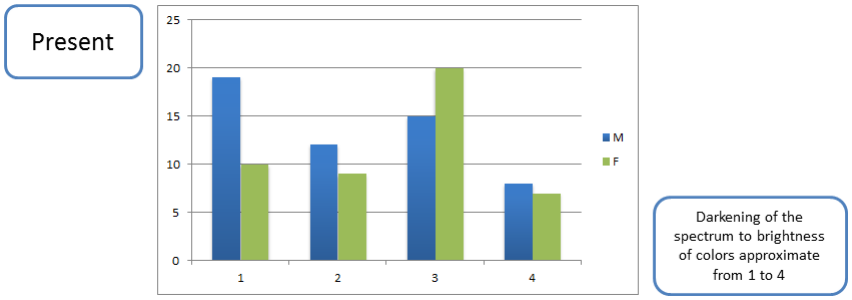

Fig. 1. Sex towards color

Since most of the participants who have taken part are from Iran and some European countries, the results in the Middle East and 1st level European countries (more prosperous) have provided the most accurate results. Fig. 2 compares chosen countries and favorite kinds of colors of the participants.

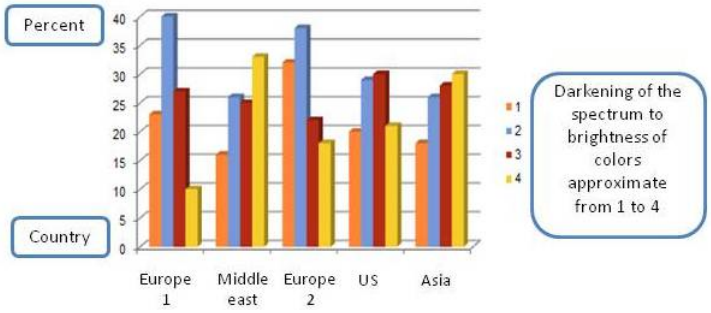

Fig. 2. Color toward User's Location

As shown in Fig. 2, lighter color designs have attained higher ratings in Asia and the Middle East compared to other tones; meanwhile European countries have experienced a rather reversed regime. Asian countries, especially those in Middle East, have chosen light colors. In Fig. 2, the horizontal axis of the diagram shows the average dark to light approximate tone scaled from 1 to 4 while the vertical axis of the diagram indicates the percent of designs chosen by individuals in each group.

3.2. Time spent in website with relation to user's education

According to Fig. 3 those who have a higher education spend more time on a website. Many participants with a higher education have spent 40-45 sec time using the website. In this diagram, the horizontal axis shows the education level while each color is indicative of the average time spent on the website. The vertical axis shows the percent of designs chosen by individuals in each group. The average time spent on web pages as reported by the literature was higher than the time we inferred based on our research.

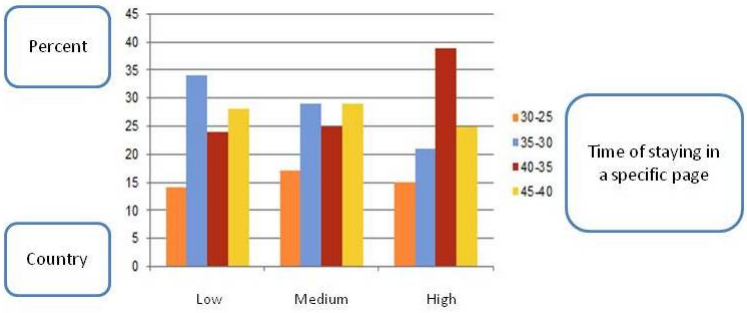

Fig. 3. Spent time in website towards user's education

3.3. The role of location

Location of the participants has been divided into 5 regions and the characteristics have been shown in such a way that they could be comparable. These diagrams are used to compare different locations towards some specifications which have revealed successful results. This diagram is indicative of the average time spent by individuals on websites. Thereby, time has been displayed and location of each region has been specified according to the average time. People who live in the Middle East and USA spent the shortest time on websites, while those who live in EU and South America spent the longest time on web pages.

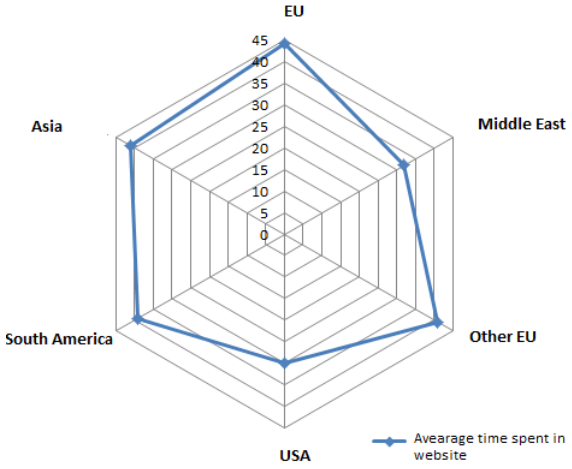

Fig. 4. Location towards average time spent in websites

Fig. 5 illustrates the number of links in web pages since the average number of links towards each region has been specified on the diagram. The number of links affects user's concentration and accuracy of his/her choice. Whenever the number of links is greater, it would be more difficult for the user to choose his/her favorite link.

There could be a solution by combining links with help information. People, who live in 2nd level European countries, are more interested in greater number of links.

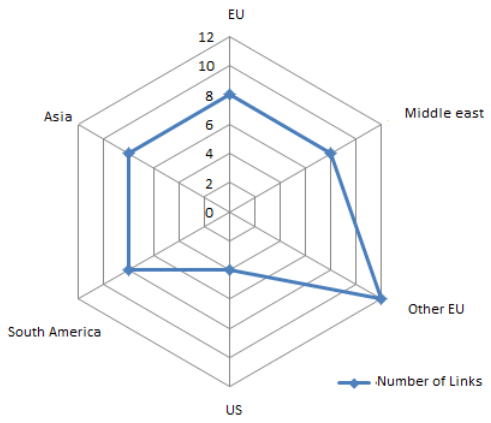

Fig. 5. Location towards number of links

Considering the interest of people from various regions in the world in higher or darker tone colors, Fig. 6 reveals the relation between the color variety and the location of the participants. In this diagram, the amount of color variety has been determined for each region. People from North America show less interest in high variety of colors.

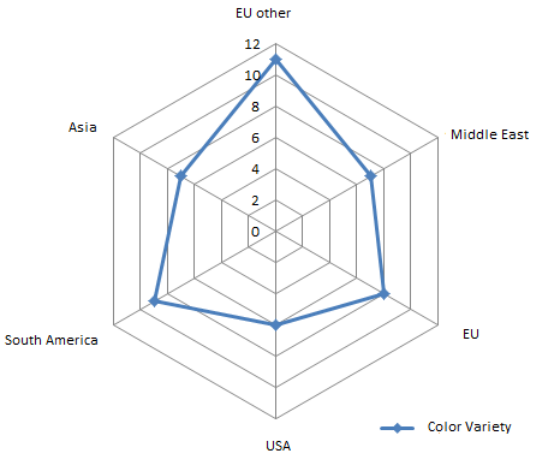

Fig. 6. Location towards color variety

3.4. Location towards page complexity

Fig. 7 focuses on the interest in complex pages in various regions. In this diagram, the average tagged amount of complexity has been specified for each region. Advanced European countries in addition to Asian countries show great interest in more complex pages.

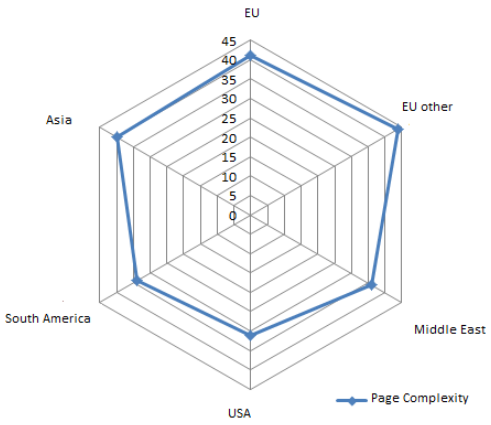

Fig. 7. Location towards page complexity

3.5. Age towards number of links

In this diagram, the horizontal axis has 4 values: 1 for the age group presented on the X axis, 2 and 3 for the age group of presented on the Y axis (between 18 and 31 years old which has been divided into two sets), and 4 for the age group of presented on the Z. Loads of the diagram are the number of links and the vertical axis represents the redundancy.

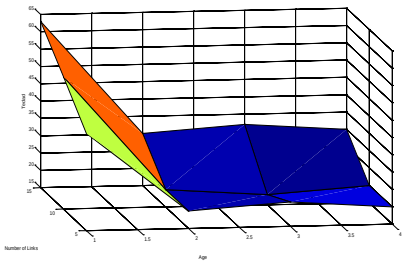

Fig. 8. Age towards number of links

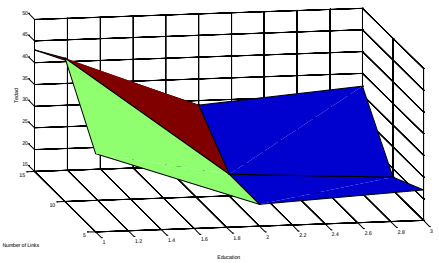

Fig. 9. Education towards number of links

As can be observed, users of lower ages have chosen the greatest number of links. The trend in this diagram is similar to the previous one and it is possible to interpret relations between age and education. One of the horizontal axes is education in which value “1” denotes lower education levels and value “4” denotes higher education ones; while the other one is the number of links. The vertical axis represents the redundancy. It can be seen that lower education has been accompanied by the greatest number of links.

3. Summary

The paper presented some aspects that were analyzed based on experiments with users that have been attracted from social networks. Thze results can be summarized in the following way.

Users with lower education level together with users of lower income are interested to be placed the overview columns at the right. It can therefore be concluded that they have been more related to the Persian websites whose exploring experience was limited to simple issues in Iran governmental websites e.g. websites of gas card, post, Housings and etc. which usually have their column at right. One conclusion made from these observations was that exploring experience of a user has direct impacts on his/her favorite styles.

People who have education level less than bachelor, prefer less than 8 links. Users having an education level lower than master show the least attention to page components. They are also interested in greater number of links.

Users who have low income level like to visit a website with light colored and busy header pictures. On the other hand, users with high income level are keen to visit websites with warm colored and nature header pictures. Individuals whose average time spent on various web pages is more than normal, are interested in third type of websites which do not have columns or busy design.

Users older than 31 years choose the header picture from busy ones. They also use the darkest colors to design the website and prefer plenty of pictures in it. With respect to the limitations such as participation of users in all ranges under study, some of the specifications have adequate resources while others lack such resources. For example, majority of the sampling population are Iranian students who will affect education level and financial situation as well as other specifications

One other drawback experienced was related to the difficulty to bring people of other countries to agreement. The least number of participants were from Africa and Australia while the number of participants from America was not enough as well. It seems that the income criteria have not been answered properly. The major limitation of the research was in social networks in which users hardly trust newly released applications due to incidence of virus worms. Therefore, it will be rather difficult to satisfy those users.

Achievements of this research can be further employed in the design of dynamic websites which requires much time and cost; to conform page patterns regarding the tags from page classification in which the users have shown interest and to develop results obtained from observing users' behavior toward other practical areas.

Considering greater number of characteristics with more details in addition to more users can certainly impose a positive effect on the model and tagging the designs, meanwhile it will make results more accurate. Furthermore, if user's exploring behavior is recorded in longer period of time, the results will be of higher accuracy. A number of parameters have been considered here for the voting system and page ranking. For example, receiving feedback from the user affects its class cluster designs in such a way that adds a coefficient for the positive votes and

subtracts it for the negative ones. Modification of these coefficients can lead to higher performance. Coefficient mentioned for distinction of specifications for each part of the website can also be altered. Changing the normalizing parameters can have positive effects on the results.

Some characteristics which have been collected from questionnaires may lack enough accuracy. Besides, some people may not allow their demographic information to be published through social networks due to security issues.

Applying techniques to make them agree to do so, (for example through trust let techniques) can have significantly influence the accuracy of data.

References

1. B. Mobasher , R. Cooley , J. Srivastava, "Automatic personalization based on Web usage mining", Communications of the ACM, v.43 n.8, p.142-151, (2000).
2. S. Schiaffino, and A. Amandi "User–interface agent interaction: personalization issues." International Journal of Human-Computer Studies 60(1): 129-148, (2004).
3 .A. Rasooli, F. Taghiyareh, P. Forbrig, “Categorize Websites based on design issues”, Human-Computer Interaction, HCII 2011, LNCS 6764, pp. 510--519. Springer, Heidelberg (2011).
4. K., J. Huang, et al, "A survey of E-commerce recommender systems. Service Systems and Service Management", International Conference on In Service Systems and Service Management, 2007 International Conference on, pp. 1-5, (2007).
5. S. Casteleyn, O. De Troyer, and S. Brockmans, "Design Time Support for Adaptive Behaviour in Web Sites." In Proceedings of the 18th ACM Symposium on Applied Computing, pages 1222 - 1228, ACM, ISBN 1-58113-624-2, (2003).
6. Leonidis, A., Antona, M., Stephanidis, C. "Rapid Prototyping of Adaptable User Interfaces". International Journal of Human Computer Interaction, (2011).
7. Stephanidis, C. The concept of Unified User Interfaces. In C. Stephanidis (Ed.),” User Interfaces for All - Concepts, Methods, and Tools.” Mahwah, NJ: Lawrence Erlbaum Associates, pp. 371-388 ,ISBN 0-8058-2967-9, (2001).
8. Stephanidis, C., Paramythis, A., Sfyrakis, M., & Savidis, A. “A Case Study in Unified User Interface Development: The AVANTI Web Browser”. In C. Stephanidis (Ed.), User Interfaces for All - Concepts, Methods, and Tools (pp. 525-568). Mahwah, NJ: Lawrence Erlbaum Associates, (2001).
9. Stephanidis, C., Paramythis, A., Zarikas, V., Savidis, A. “The PALIO Framework for Adaptive Information Services.” In A. Seffah & H. Javahery (Eds.), Multiple User Interfaces: Cross-Platform Applications and Context-Aware, (2004).
10. Antona, M., Savidis, A., & Stephanidis, C. “A Process–Oriented Interactive Design Environment for Automatic User Interface Adaptation”. International Journal of Human Computer Interaction, 20 (2), 79-116, (2006)..
11. C. Petit-Rozé, and E. Grislin-Le Strugeon, "MAPIS, a multi-agent system for information personalization." Information and software technology, ISSN 0950-5849, Vol. 48, No. 2 , pags. 107-120, (2006).
12. P. Kazienko, and M. Adamski ,"AdROSA—Adaptive personalization of web advertising." Information Sciences Wrocław, Poland 177(11): 2269-2295, (2007).
13. J. Seo, K. Lee, "Development of Website Design Personalization Service Using Design Recommender System”, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.108.3769, Access time: October 2010, (2003).