Web-Systems Remote Usability Tests and Their Participant Recruitment

Piotr Chynał(✉) and Janusz Sobecki

Department of Informatics, Wroclaw University of Technology,
Wyb.Wyspianskiego 27, 50-370 Wroclaw, Poland
{piotr.chynal,janusz.sobecki}@pwr.edu.pl

Abstract. In this paper we present a description of a proposed hybrid, remote usability testing method and a comparison of different approaches to participant recruitment for the test conducted according to this usability evaluation method. Moreover this paper contains a description of the implemented hybrid method and its characteristic. One of the main features of this method is that it allows to perform remote online tests. These tests are an alternative to traditional laboratory tests. They don't require a special laboratory space, gathering participants in one place, a moderator or other equipment to perform the tests. However we have to face a challenge – how to recruit participants for remote usability test, which is more complicated because we must motivate our users to participate in such test without having a direct contact with them.

This paper presents a comparison of a few selected methods that we used to encourage users of website HotelGo24.com to take part in usability evaluation test of that site. We present how many users were ready to participate in our study depending on the applied method of encouragement and their reward for participating in the study.

Keywords: Usability evaluation · Remote testing · Participant recruitment methods

1 Introduction

Since the emerging of web-systems usability evaluation field, many various testing methods have been proposed. They have been changing and adopting according to the used technologies and trends in web-system development. Currently there are almost one hundred usability methods that can be used for such evaluation [1] that have been developed in the last four decades. They evolved from the laboratory end-user testing [2], through inspection methods [3], to remote testing [4]. They also utilize very different equipment from video recorders, eye trackers [5] or EEG devices. Having so many various usability testing methods we can combine few of them, during one evaluation, to perform a full usability audit of a given system. However doing so can be time and effort consuming. The solution to this problem is to create one method, which would combine particular elements of other methods – hybrid method.

The main assumptions for the hybrid method are following: it should have an ability to perform complex usability tests much quicker than using other methods and

M. Kurosu (Ed.): Human-Computer Interaction, Part I, HCII 2015, LNCS 9169, pp. 175–183, 2015.
DOI: 10.1007/978-3-319-20901-2_16

ability to gather all sorts of data regarding user's interaction with the web-based system being evaluated. Moreover it should be low cost, without the requirement of moderation, and it should allow to test a large group of users at once. The detailed description of hybrid method assumptions can be found in [4].

One of the main features of this method is that it allows to perform remote online tests. These tests are an alternative to traditional laboratory tests. They don't require a special laboratory space, gathering participants in one place, a moderator or other equipment to perform the tests. Main goal of such evaluation is to test the usability of a given system by users working in their natural environment, so they behave more naturally, like they would normally do while using the given web-based system. Another advantage of such tests is that they do not require to gather all the users at one time, they can take part in the evaluation when they want. Moreover it is possible to test participants from all over the world and not only those who can physically visit our laboratory. Also some studies have been performed that show that the same usability problems have been found using traditional laboratory testing and remote tests [6–8]. In the end, in the hybrid method we have used elements of the following methods:

- Questionnaires - survey regarding experience with the system
- Individual User Testing - tasks for users in the evaluated system, task success rate and task completion time
- Clicktracking - clicks and path on the website for each user
- Unmoderated Testing - frame on the top of the evaluated system with tasks for participants
- Automatic Testing - automatic usability calculation from obtained data
- Remote Testing - online application for testing

After our hybrid method was developed we have faced another serious challenge - how to recruit participants for usability tests to be carried out with the proposed method? This problem has already been addressed by at least several works. In the remote usability tests described in [6] the participants were recruited by email message that was sent to randomly selected employees of a company for which usability tests have been made. In [9] and [10] from the other hand, authors describe utilization of crowdsourcing platforms such as Amazon Mechanical Turk and CrowdFlower for conducting remote usability tests. Crowdsourcing is defined in [11] as 'the act of taking a job traditionally performed by a designated agent and outsourcing it to an undefined, generally large group of people in the form of an open call'.

The following chapters present the description of implementation of hybrid usability evaluation method, the experiment with participants using this method and different ways to encourage users to take part in this experiment, results of this experiment, summary and future works.

2 Hybrid Method Implementation

After designing the hybrid method, the next step was to implement an application that would allow to perform usability tests according to this method. According to the method assumptions this method works as an application that allows to perform remote

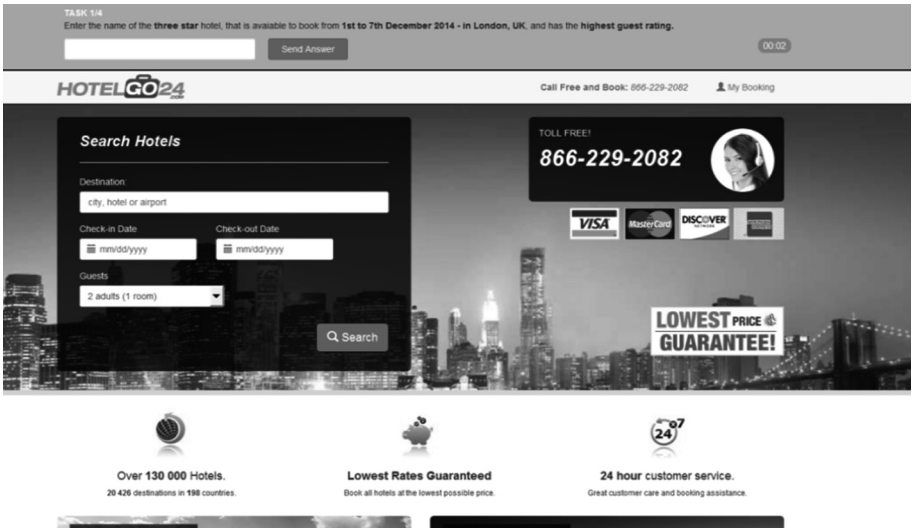


Fig. 1. The usability application is open in a small frame on the top of the screen and displays task for users, questionnaires etc.

unmoderated online tests. With this method participants perform predefined tasks in the system and the application records their progress. At the end the application displays this data and some calculated metrics that allow the evaluator to assess the usability of the evaluated system.

This part of research has been completed with help of the company Fragria Systems as a part of the Grant Plus Scholarship (Human Capital Operational Programme, Priority VIII Regional Human Resources Management, Measure 8.2 Transfer of Knowledge, Sub-measure 8.2.2 Regional Innovation Strategies) from the Faculty of Economy Lower Silesian Marshal Office.[1] After testing various ideas and technologies, finally the system was implemented in JavaScript and PHP languages and works as a script that is added to the evaluated web-system. Working application is presented in Fig. 1.

The usage of this application is simple, the evaluator creates tasks and questionnaires for users, inserts the JavaScript code into each of subpages and after those participants complete the tasks, during which the data is recorded. The process of evaluation with hybrid method is presented on the following schema (Fig. 2).

After the evaluation, the recorded data for each participant can be viewed in the administrator panel (Fig. 3).

This data allows to perform detailed analysis of each participant's performance. We can view the task success rate and completion times, answers to the questionnaire and we are able to check the progress for each task (visited subpages and clicked elements). This way we can look for some usability issues that participant may have encountered, for example, while completing a task, participants spend a lot of time on a particular

[1] http://www.grantplus.dolnyslask.pl/.

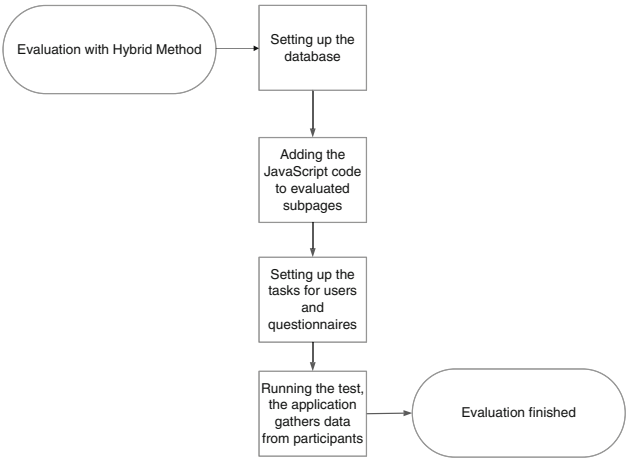
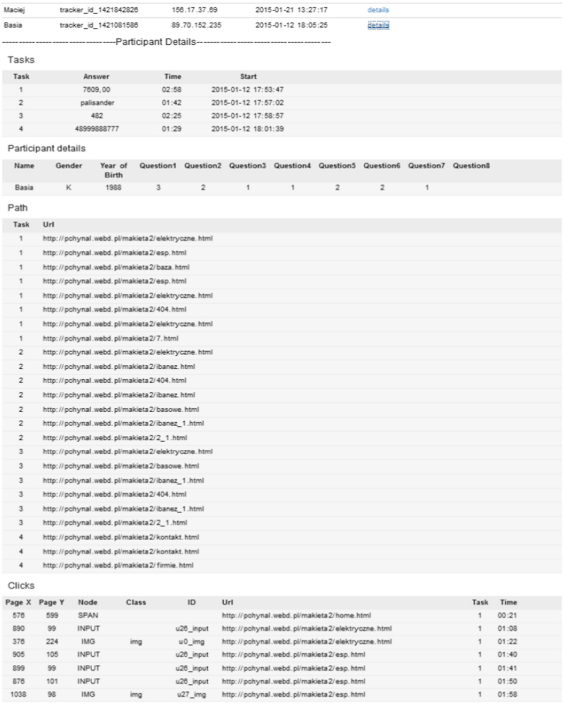


Fig. 2. Evaluation with the hybrid method

subpage, before they clicked on proper element – that means that this element is not visible to users and it is a usability problem regarding design of information presentation.



Maciej tracker_id_1421842826 156.17.37.69 2015-01-21 13:27:17 details
Basia tracker_id_1421081586 89.70.152.235 2015-01-12 18:05:25 details

----------Participant Details----------

Tasks

Task	Answer	Time	Start
1	7609,00	02:58	2015-01-12 17:53:47
2	palisander	01:42	2015-01-12 17:57:02
3	482	02:25	2015-01-12 17:58:57
4	48999888777	01:29	2015-01-12 18:01:39

Participant details

Name	Gender	Year of Birth	Question1	Question2	Question3	Question4	Question5	Question6	Question7	Question8
Basia	K	1988	3	2	1	1	2	2	1	

Path

Task	Url
1	http://pohynal.webd.pl/makieta2/elektryczne.html
1	http://pohynal.webd.pl/makieta2/esp.html
1	http://pohynal.webd.pl/makieta2/baza.html
1	http://pohynal.webd.pl/makieta2/esp.html
1	http://pohynal.webd.pl/makieta2/elektryczne.html
1	http://pohynal.webd.pl/makieta2/404.html
1	http://pohynal.webd.pl/makieta2/elektryczne.html
1	http://pohynal.webd.pl/makieta2/7.html
2	http://pohynal.webd.pl/makieta2/elektryczne.html
2	http://pohynal.webd.pl/makieta2/ibanez.html
2	http://pohynal.webd.pl/makieta2/404.html
2	http://pohynal.webd.pl/makieta2/ibanez.html
2	http://pohynal.webd.pl/makieta2/basowe.html
2	http://pohynal.webd.pl/makieta2/ibanez_1.html
2	http://pohynal.webd.pl/makieta2/2_1.html
3	http://pohynal.webd.pl/makieta2/elektryczne.html
3	http://pohynal.webd.pl/makieta2/basowe.html
3	http://pohynal.webd.pl/makieta2/ibanez_1.html
3	http://pohynal.webd.pl/makieta2/404.html
3	http://pohynal.webd.pl/makieta2/ibanez_1.html
3	http://pohynal.webd.pl/makieta2/2_1.html
4	http://pohynal.webd.pl/makieta2/kontakt.html
4	http://pohynal.webd.pl/makieta2/kontakt.html
4	http://pohynal.webd.pl/makieta2/firmie.html

Clicks

Page X	Page Y	Node	Class	ID	Url	Task	Time
576	599	SPAN			http://pohynal.webd.pl/makieta2/home.html	1	00:21
890	99	INPUT		u26_input	http://pohynal.webd.pl/makieta2/elektryczne.html	1	01:08
376	224	IMG	img	u0_img	http://pohynal.webd.pl/makieta2/elektryczne.html	1	01:22
905	105	INPUT		u26_input	http://pohynal.webd.pl/makieta2/esp.html	1	01:40
899	99	INPUT		u26_input	http://pohynal.webd.pl/makieta2/esp.html	1	01:41
876	101	INPUT		u26_input	http://pohynal.webd.pl/makieta2/esp.html	1	01:50
1038	98	IMG	img	u27_img	http://pohynal.webd.pl/makieta2/esp.html	1	01:58

Fig. 3. Data gathered during the test – task completion times with answers to those tasks, questionnaire scores, path for each task and click for each task.

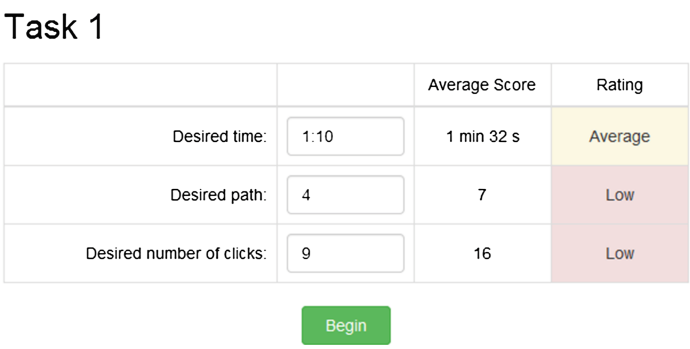


Fig. 4. Automatic usability assessment based on optimal parameters and data recorded from participants.

Moreover the evaluator can view the usability of the website that was calculated based on implemented rules that assess usability taking into account provided optimal parameters and data gathered from the participants, for each task, as shown in Fig. 4.

3 The Experiment

The experiment was part of the evaluation of the hybrid method. We wanted to check how it compares to other methods by performing usability evaluation of a website www.hotelgo24.com. It is a typical hotel booking website that allows its users to search and book hotel rooms all around the world. For the purpose of the evaluation with the hybrid method we have created four tasks for the participants:

- Enter the name of the three star hotel, which is available to book from 1st to 7th April 2015 - in London, UK, and has the highest guest rating.
- Check if you can indicate special room requests or preferences, such as connecting rooms, bed type, smoking type, or early check-in?
- Enter the address of the cheapest hotel that is available to book from 3rd to 8th May 2015 - in New York, USA.
- Enter the telephone number to HotelGo24.

The next step was to recruit participants for the tests. Our target was to recruit 100 users to participate in our test. We wanted to check the effectiveness of various recruitment methods. We have divided them into two groups – banners on the evaluated website and advertisements on other websites. For the first group the effectiveness would be calculated by comparing the total visits on the website with the number of people that participated in the evaluation, during the time that advertisement was on the website. For the second group we would set a goal of 20 participants and we

measured how long it will take to get such amount of people from each source. The methods that we have used are following:

a. Banners on our website that were encouraging to take part in the evaluation:
   (i) Take part in usability evaluation of HotelGo24.com and help us to improve our website
   (ii) Take part in usability evaluation of HotelGo24.com and earn 5mBTC
   (iii) Take part in usability evaluation of HotelGo24.com and help us to improve our website, for each participant we will donate 1mBTC to charity[2]
b. Advertisement on the BitCoin forum[3] – same as banner on our website, 5mBTC as a reward
c. Advertisement on HotelGo24.com profile page and hotel/travel related groups on Google+ - same information as banner on our website, 5mBTC as a reward
d. Advertisement among friends on Facebook – encouragement to help a friend with the research and to share the post on own Facebook Timeline.

Duration of visibility of banners on the website was 10 days. We used BitCoin (BTC) currency because it is the easiest way to make payments on the Internet. 5mBTC was worth around $1 during the evaluation. More on BTC can be found in [12]. For those methods that included payment we have used verification with vouchers – after the evaluation was completed, participants were presented with generated voucher that they were supposed to enter on HotelGo24.com Google+ profile, along with the Bit-Coin account number (in case of charity only voucher).

We also wanted to place an advertisement on charity website – encouraging users and Facebook fans of this website to take part in the evaluation, for each participant 1mBTC would be donated to that organization. However after sending e-mail to about 15 different charity organizations that accept BTC with a proposal of such cooperation, none of those organizations expressed willingness or even replied to our e-mail. There are also other methods that we have considered such as payment for the recommendation of other users – after the evaluation participants enter the voucher that they have received from the recommending user, who gets paid for each of those participants. We will try to check this method in the future research. Moreover there are some obvious ways to recruit participants such as asking students to take part in the evaluation or simply asking friends and family, but they results may not be entirely credible as they will have some predefined attitude towards the evaluation (friends will look more favorably on the product and students probably the opposite). Another possible ways of participants' recruitment is application of platforms such as Amazon Mechanical Turk, CrowdFlower MTurk or other crowdsourcing platform.

[2] http://thewaterproject.org/.

[3] https://bitcointalk.org/.

4 Results of the Experiment

Till the moment of writing of this article, we were able to evaluate 73 patricians. The results of the experiment in terms of effectiveness of hybrid method will be published in future works. Here we present results from participant recruitment methods that we have used (Tables 1 and 2).

First table presents how many people completed the evaluation depending on the encouragement method on the displayed banner. Percentage of participants was the highest for the method with reward for each participant. For other two methods percentage was about the same, people were not tempted to support charity organization. However even though users were encouraged by payment in BTC, only few more than without any reward were inspired to complete the evaluation. It might be caused by a fact that BitCoin awareness is still small and people do not use this currency. However for the second table, we can clearly see that people who were encouraged by payment and were familiar with BTC were eager to participate in the evaluation. As for the ads on Google+, the response was much slower, probably also because of people being not familiar with the BTC. Maybe a good idea would be to promote such ad on some groups related to BitCoin instead of groups related to traveling and hotels. From Facebook friends it was relatively easy to find 20 people willing to help, without any financial benefits. However as mentioned earlier, evaluation by such people might not be fully objective, as friends have some predetermined attitude towards the person posting for help.

Table 1. Results for encouragement with banners on the HotelGo24.com website

Banner	No reward	5mBTC	1mBTC to charity
Number of unique visits on the website during evaluation	278	309	284
Number of participants of the evaluation (visited link)	11	25	14
Number of successful participants of the evaluation (completed)	3	7	3
Percentage of successful participants	1.08 %	2.26 %	1.06 %
Cost	0	35mBTC	3mBTC

Table 2. Results for the advertisements on other websites

Advertisement	B-forum	C-Google+	D-Facebook
Time for 20 people to complete the evaluation	2 days	7 days	9 days
Range	1238 views of the ad	Around 1100 People might view the ad	Around 700 people might view the ad
Cost	100mBTC	100mBTC	0

5 Summary and Future Work

To sum up, our experiment has shown that encouraging participants to take part in the remote usability tests is not an easy task. If we try to evaluate a website with a statistically significant number of users, without giving them any reward for completing the evaluation, it would take a very long time to do so. On the other hand, rewarding the participants makes them much more eager to take part in such evaluation as we found out from the advertisement on BitCoin forum. It was the fastest way to gather 20 participants. Bit Coins prove to be an easy way to reward users for their participation, however many people that visited HotelGo24.com website did not know what BTC is and they were afraid that it is some kind of hoax. Experiment results show also that people are not eager to perform some tasks without any reward, even if they work would correspond to some donation to charity organization.

Moreover the direct advertisements on some websites, where people are looking for ways to earn money are definitely more efficient then trying to encourage the hard users of our website. This however has some drawbacks, because the test participants might not be the target group of our website and might never use it without this reward. The best way is to try to encourage users of our website and simultaneously hire some other people to complete the evaluation, to get the best diversity of user profiles and most valuable usability testing results.

Regarding future work, we would like to test other methods of participant encouragement, for example by rewarding user for recommending the evaluation to other participants, application of crowdsourcing platforms or other forms of rewarding the participants, i.e. mobile phone top-up, which could encourage especially young people to participate in our tests. We would also like to try rewarding users differently than using BTC. Moreover the proposed hybrid method is going to be thoroughly evaluated and further developed.

Acknowledgements. The work was supported by The European Commission under the 7th Framework Programme, Coordination and Support Action, Grant Agreement Number 316097 [ENGINE].

References

1. Rajeshkumar, S., Omar, R., Mahmud, M.: Taxonomies of user experience (UX) evaluation methods. In: 2013 International Conference on Research and Innovation in Information Systems (ICRIIS), pp. 533–538, 27–28 Nov 2013
2. Rubin, J., Chisnell, D.: Handbook of Usability Testing: How to Plan, Design and Conduct Effective Tests. Wiley, New York (2008)
3. Nielsen, J.: Usability inspection methods. In: Conference Companion on Human Factors in Computing Systems, pp. 413–414. ACM, Apr 1994
4. Chynał, P.: Hybrid approach to web based systems usability evaluation. In: Nguyen, N.T., Attachoo, B., Trawiński, B., Somboonviwat, K. (eds.) ACIIDS 2014, Part I. LNCS, vol. 8397, pp. 384–391. Springer, Heidelberg (2014)

5. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., Van de Weijer, J.: Eye Tracking: a Comprehensive Guide to Methods And Measures. Oxford University Press, Oxford (2011)
6. Tullis, T., Fleischman, S., McNulty, M., Cianchette, C., Bergel, M.: An empirical comparison of lab and remote usability testing of web sites. In: Usability Professionals Association Conference, July 2002
7. Brush B., Ames M., Davis J.: A Comparison of Synchronous Remote and Local Usability Studies for an Expert Interface. In CHI 2004 Extended Abstracts on Human Factors in Computing Systems, ACM, New York, USA, pp. 1179–1182 (2004)
8. Oztoprak A., Erbug C.: Field versus Laboratory Usability Testing: a First Comparison. Technical report, Department of Industrail Design - Middle East Technical University, Faculty of Architercture, Inonu Bulvari, 06531 Ankara, Turkey (2008)
9. Liu, D., Bias, R.G., Lease, M., Kuipers, R.: Crowdsourcing for usability testing. Proc. Am. Soc. Inf. Sci. Technol. **49**(1), 1–10 (2012)
10. Kittur, A., Chi, E.H., Suh, B.: Crowdsourcing user studies with Mechanical Turk. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 453–456. ACM, Apr 2008
11. Howe, J.: Crowdsourcing: A Definition. http://crowdsourcing.typepad.com/cs//06/crowdsourcing_a.html. Accessed Jan 2015
12. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system. Consulted **1**(2012), 28 (2008)