Studies in usable security - advantages and drawbacks of the proposed solutions

Christina Krutzenbichler

Abstract— This paper proposes an overview of types of studies, that were used in the field of usable security. Three papers have been chosen and the different experiments have been compared respective their advantages and disadvantages. The first one examines the user tolerance of security delays, using a new web-based document viewer. The second explores usable security performing a laboratory study that compares the effectiveness of passive and active phishing warnings. The third study collects eye tracking data to examine the user's awareness to browser security. In the comparison seven characteristics of the studies have been used. These are the recruitment process of the subjects, the pilot studies that have been made, the overall costs of the study, the average time of the execution of the study, the environment the study was held in, the different phases of the studies and the results.

Index Terms—usable security, methodology, study, human computer interaction

1 Introduction

Nowadays most people use the internet and so usable security is getting more important. Also since you can not only use the internet from your desktop computer, but also from your smartphone, laptop, tablet and other devices, the security is an important factor. Most of the people do not know much about security or do not even care and that is why you have to improve usable security [2] . The studies help to improve security, because you analyse the behaviour of people using e.g. the internet and how they interact with the different security breaches that are presented there. The following sections of this paper will give a review of different methods of studies on usable security and their excecution found in the literature.

2 Studies in Usable Security

2.1 Field Study

In my research, I only found one field study, that really fit into my subject-matter. In this field study [2],the researchers analyzed the user tolerance of security delays. Hence, they invented a new web-based document viewer called "SuperViewer" and the test persons were told that they were beta-testing this reader. For this they had to count how many times, a specific word in the document appears. In this paper the researchers used Amazon Mechanical Turk [1] to recruit the test persons for the study.

Four different versions of the "SuperViewer" were created. The first one was named *Control* and here there was only a progress bar called "Loading" when starting the program and you had to view the pages in order. The second one called *Loading*, which is identical to the *Control* Condition except when switching pages, there also was a progress bar called "Loading". *Security* was the third condition and it was the same as the Loading version with one exception. The label on the progress bar changed from "Loading" to "Performing Security Scan". The last version is SecPrimed, which was identical to the third condition, however here, the subjects were informed of the danger of viruses before launching the viewer. The scientists decided then, to pilot their experiment. Overall, they had to do three pilot studies, until they had their ideal outcome. Here, they decided on the payment, that the participants were going to receive and they switched from a Java programmed SuperViewer to one that was programmed in Flash, because it worked better. After finishing their first phase of the study, with the conditions I explained above, they decided to do another phase. Here they created four more versions of the "SuperViewer". The first one named *Adjusting*, was basically identical to the Loading condition, however the progress bar label changed from "Loading" to "Adjusting document width". Secondly, there is AdjPrimed, which is the same as the "Adjusting" condition, but they added priming information, when the viewer was opened. The third condition was AdjSecure, which also was identical to the "Adjusting" condition, but they added the same security information, as in "SecPrimed". The final condition was Downloading, being the same as "Adjusting", however they did change the label on the progress bar to a non-security information. At the end of the two phases, the subjects were offered extra money to complete a exit survey on their opinions of the viewer. Here, they analysed for the different conditions, who of the participants read all, some or none of the text and how many were right with their answer.

For this study, there was a lot of quantitative data to analyse. The researchers looked at each of the conditions and for the total time it took the participant to complete, the time he spend per page, the unique pages, the total pages and the cheaters they had.

2.2 Laboratory Study

Egelmann et. al, L.Faith Cranor and J. Hong [3] explore usable security performing a laboratory study that compares the effectiveness of passive and active phishing warnings and analyze if, why and how they will fail users. The structure of the study was designed to be a between-subjects study. This means that every participant only did one of the four models, using either Firefox 2.0 or the Internet Explorer 7.0:

- Firefox warning
- active IE warning
- passive IE warning
- no warning

The recruitment took place all over Pittsburgh, for the outcome to be generalizable. In this study as well they did not tell the test persons that they did a study about usable security, but they told them they did a study about online shopping. Now, they created a online survey, so they could filter out some persons. The criteria for participating in the study, were for example the email provider or the browser version, which the subject currently used. For the study they decided to spoof Amazon and Ebay and the participants had to purchase one item on each of the plattforms. After a purchase, they would receive a phishing email and a legitimate email. At the laboratory, the test person would be recorded, so they had to think aloud and the experimenter would be sitting behing them. They had to purchase paperclips from both of the websites. After the purchases were done, they received an online

- *Chrisitna Krutzenbichler is studying Media Informatics at the University of Munich, Germany, E-mail: c.krutzenbichler@campus.lmu.de*
- *This research paper was written for the Media Informatics Proseminar on "Medieninformatik", 2015.*

exit survey. Here every test person recieved a payment of 35$ and they would be repayed for the items they purchased. The factors, that were analysed in this study were the phising susceptibility, the warning comprehension, the attitudes and beliefs, the motivation and warning behaviors and the environmental stimuli.

2.3 Laboratory Study

A study collecting eye tracking data to examine user's awareness to browser security was covered by T. Whalen and K.Inkpen [4] . The recruitment here was rather simple, they recruited test persons from their local university, using a prescreening questionnaire, in which they wanted to find out, if the subjects were experienced web users. Overall they had sixteen participants, to whom they paid 10$ to complete the study.

The main part of the study was split into two phases, which every participant had to complete. The purpose of the first phase was to examine the users normal browsing, without him knowing about the security analysis. Here, they had to complete five different tasks, which either were secure or non-secure. The secure ones were to use web-based email, log into a bank website using a lab account and purchase a cable from a small obscure retailer using a research account credit card. Read a news article and perform a simple google search were the non-secure tasks.

After completing the tasks, they were presented with a questionnaire, focusing on the three security tasks. For the second phase, there were only three tasks for the praticipant and they were told to specifically look if the websites were secure or not. The first one was to access an online banking account from a website none of them had ever used. Secondly, they had to log on to Microsofts Hotmail Web Service to get emails. The last task was to just buy a book from a known online bookstore, paying with the research lab's credit card. At the end of this phase, there also was a questionnaire, which asked them if they checked for security and what indicators they used for making the decision if a website is secure or not. Analysing the collected data, the researchers here realized, that they were not able to use the data from phase one, since they failed to reproduce normal browsing behaviour. For phase two, they could compare some results, the overall usage of web security information, the usage of browser-specific security cues and the recognition of secure connections.

3 Comparison of Studies

The three studies [4] [2] [3] , which I just described, are all examining web-based securtiy. The first one [2] is a field study and the other two [4] [3] are laboratory studies.

3.1 Recruitment

The biggest advantage of [2] is that in a field study, the recruitment process is different. Here you can reach out to far more persons, even globally, and the variety between them is bigger. Compared to [3] which is a laboratory study, the approach is similar. They are trying to reach people with different backgrounds. Nevertheless they where limited to the city of Pittsburgh. The severest disadvantage relating to the recruitment was in [4] . Here they only had the universties faculty and staff and a few students. Therefore they can not refer to the results in the way, that they would refer to the general public.

3.2 Pilot Study

Since [4] was only a preliminary study, they did not have a pilot study. To their disadvantage, they had experimental difficulties and so they were not able to use the eyetracking data from phase one. Using a pilot study, it prevented [2] from loss of data. In their pilot study, they found out, that some test-persons were not able to use the *SuperViewer* written in Java, so they rewrote it in Flash. Whereas in [3] ,they implemented a online screening survey to rule out persons that did not fit the criteria, which worked really well for them.

3.3 Study

The overall costs of the study, is the first attribute to be compared. [3] was with distance the most expensive study. Since they had to pay each participant 35$ for time compensation and in addition they had to repay each one for their online purchases. Overall the estimated costs only for the test persons are around 3000$. Compared to that, the [2] and [4] were cheap. For [2] the researchers only payed 109$ for the test persons and in [4] only 160$. But for all that you can not forget about the costs of the equipment. You have an advantage here, by using a field study like [2] ,because they only had to provide a software but not any hardware. For the other two studies [4] [3] the lab had to be equipped with the material, which also is a big cost factor.

Here, the second characteristic to compare, is the environment, in which the study took place. In [4] the drawback was, that the participants had to be in the research lab and use the research lab's acoounts for the online shopping. Compared to [3] where the test persons were told to use their own credit cards. Still, the researchers there, told the participants, that they were recorded and someone from the staff would always be sitting behing them. The best environment for the study was created by [2] ,due to the fact that it was a field study and the probands were at home while participating in the study. Since at home, people act more like they usally do.

My third characteristic that I am going to compare is the time it took to perform the study. Since [4] and [3] were laboratory studies, every test person had to be invited to the lab. Having to observe every subject while doing the experiement is one of the drawbacks since it takes a lot of time. [2] has the tremendous advantage, that it is a field study and the subjects were recruited over Amazon Mechanical Turk [1]. It took the researchers there a lot less time, to recruit and to perform the experiment.

One important thing, all of the three studies [4] [3] [2] have in common, is that non of the participants were told, that the study was about security. Despite that, the main structure of the studies was different. In [4] the participant would go through phase one and after completing this phase fill out a survey. Then before starting with phase two, he would be told to specifically look for indicators of securtiy breaches. Antoher survey had to be filled out after completing the phase. This has the advantage, that you can compare the way the test person acts, when not knowing and when knowing to look for security breaches. The structure of the other two studies [3] [2] were similar, they only had one phase every test subject had to complete, but they offered more different ways in which you could complete the study. [2] also had a second phase, but choose to recruit different probands for it. When doing this, you have the disadvantage, that you compare other people in phase two then in phase one. Those people mostly react in a different way. This structure is called a between-subjects study.

My last characteristic of the study, is how the data was analysed. In [4] there was a big disadvantage. The researchers realized, that they were not able to reproduce normal browsing behaviour during their experiment and therefore could not use the collected data. The other two, [2] and [3], on the other hand, analysed their data using different proven methods. To their advantage they could both conclude, that not many of the participants paid attention to the security warnings.

4 Conclusion

Each of the proposed studies has advantages and disadvantages. In my opinion it depends on the situation which one is the best to use. [2] would be helpful for a research team, if they need a lot of quantitative data. Here you get a really good overview of how tolerant people are with the used security. The solution in [4] would be advantageous for a preliminary study. Since you only have a few participants and the study does not take too much time. [3] can be useful for overlooking how the subject really acts. In this study, you record the participant and you watch him interact with your system, so the quality of the collected data is a lot more specific.

Further research of usable security is important as it provides a secure environment for the internet users. Furthermore the existing studies should be improved in the future. In all three papers the authors proposed suggestions for improving their solutions.

References

[1] Amazon.com.

[2] S. Egelman, A. Acquisti, D. Molnar, C. Herley, N. Christin, and S. Krishnamurthi. Please continue to hold an empirical study on user tolerance of security delays.

[3] S. Egelman, L. F. Cranor, and J. Hong. You've been warned: An empirical study of the effectiveness of web browser phishing warnings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1065–1074, New York, NY, USA, 2008. ACM.

[4] T. Whalen and K. M. Inkpen. Gathering evidence: Use of visual security cues in web browsers. In *Proceedings of Graphics Interface 2005*, GI '05, pages 137–144, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2005. Canadian Human-Computer Communications Society.