Contextualized Security Interventions in Password Transmission Scenarios

Melanie Volkamer and Steffen Bartsch and Michaela Kauer
CASED, Technische Universität Darmstadt, Germany
email: name.surname@cased.de

Abstract

Usable security user studies as well as the number of successful attacks to end users' data and devices show that today's security interventions like the green URL bar and self-signed certificate warnings do not protect end users effectively for many reasons. To improve the situation, we proposed the Framework fOr Contextualized security Interventions (FOCI). While this framework provides general guidelines how to develop contextualized security interventions, this is the first paper in which this framework is applied to actually develop adequate security intervention strategies and intervention content. We focus on a subset of security- and privacy-critical scenarios in the context of web applications – namely those in which users visit web pages containing a password filed. If either the communication is not confidential and authenticated or the service behind the web page is not trustworthy, entering a password can have consequences like financial loss and privacy leakage in particular for users reusing their passwords for several different web pages. Therefore, it is important to provide effective security interventions for these scenarios.

Keywords

Security intervention, human aspects, contextualized, https, secure password transmission, intervention strategy, threats, consequences, risks.

1. Introduction

Many user studies and statistics concerning successful attacks against end users show that neither current passive interventions (like the green URL bar in case of an extended SSL certificate) nor (active warnings like those in cased of self-signed certificates) do effectively protect their users (Dhamija *et al.* 2006, Sunshine *et al.* 2009). Researchers have identified several reasons. The main reason is that existing interventions do not adequately take into account that security is not their primary task but e.g. transferring money or buying books (West 2008, Sasse *et al.* 2001) and that the user's mental model and knowledge of Internet security is incomplete (Bravo-Lillo *et al.* 2011b). For instance, many users believe that they personally are not of interest for an attacker (Sasse *et al.* 2012). Studies also show that people tend to base the decision whether to use or not to use a web page on the design of the web page and not on (passive) security interventions (Fogg *et al.* 2001, Schechter *et al.* 2006). In general one can say that existing passive security interventions are not noticed by most of the users. Active security interventions interrupt users from their primary task and can therefore not be overlooked. However, active warnings are not much more effective for many reasons; as for instances Sunshine *et al.* (2009) show.

One problem is that the communicated information about the situation or reason for this intervention is currently on a very technical level. Thus, users are not able to deduce the risk and their personal consequences of ignoring this intervention from the provided information (Bravo-Lillo *et al.* 2011a and Kauer *et at.* 2012). In combination with the fact that browsers show the same security interventions in high and in low risk situations, users learned from many low risk situations in their daily lives that nothing "bad" happens if they ignore these interventions (Sotirakopoulos *et al.* 2011, Sunshine *et al*. 2009). Correspondingly, it is not surprising that - due to habituation effects - users will also ignore such interventions in high risk situations.

In order to improve the situation, Bartsch and Volkamer (2012) proposed the Framework fOr Contextualized security Interventions (FOCI). The main idea is that the appearance of an intervention (intervene or not), the time, type and position of the intervention (all this is defined in the intervention strategy), and the content of an intervention is primarily influenced by the user's context. The context is defined by the user's personal characteristics and the situation (what do we know about the web page, the connection to the web server, and the operator of the web page or web service). The authors left it for future work to operationalize this very general framework with respect to the different existing security-critical contexts.

In this paper, we focus on security-critical scenarios in the context of web pages with password fields (e.g. password requests over http or over https with self-signed certificates) and describe the proposed algorithm to determine whether to intervene and how (passive or active, as well as position). The reason to start with these scenarios is, on the one hand, that many of the common phishing attacks are covered with these scenarios and, on the other hand, the fact that many users reuse passwords between different services. Consequently, being able to "phish" one password often enables the attacker to get access to several web services. The proposed interventions strategy combines and extends existing approaches for security interventions; namely the multi-page warning approach proposed by Sunshine *et al.* (2009) and extended by Seikel (2012) and the idea from Maurer *et al.* (2011a and b) to display warnings only if the user starts entering sensitive data and place the warning right where the data is entered.

2. Framework fOr Contextualized security Interventions

The main idea of the framework proposed by Bartsch and Volkamer (2012) is that the intervention strategy including the appearance of an intervention (intervene or not), the time, type and position of the intervention (all this is defined in the intervention strategy), as well as the content of an intervention is primarily influenced by the concrete *situation* and *personal characteristics* (see Figure 1).

Situation indicators include *security-related indicators* which enable the framework to measure the security in a particular situation based on indicators like the trustworthiness of the operator of the web service, the protection of the communication path to the web service, the requested data (e.g. passwords, credit cards or other sensible data), the type of web page (information pages versus online banking), the user's intention (whether to provide this data or not), and whether the

user visited the same URL already in past. Other situation indicators are so called *user-influencing indicators*. Examples are professional design or containing logos from trustworthy institutions but also the type of web page and the requested data. **Personal characteristics** include the user's demographics, user's mental model and knowledge of Internet and computer/mobile security as well as his risk profile. The risk profile defines the readiness to assume a risk in different situations.

The framework requires an interdisciplinary collaboration from security and psychology researchers in order to deduce adequate contextualized intervention strategies and content. Thereby, FOCI includes results from Sunshine *et al.* (2009) and Cranor (2008). Sunshine *et al.* (2009) already indicate that the effectiveness of warnings can be improved, if the content and the intensity are adjusted to the specific context, while context in their paper mainly includes the type of web page the user visits. Cranor proposed (2008) a human-in-the-loop security framework, which describes which factors influence users, namely "personal variables" such as the user's demographics, "intentions" as the user's attitudes and believe, and "capabilities" for the user's knowledge or cognitive and physical skills. Correspondingly, also according to Cranor (2008) different warnings are necessary.

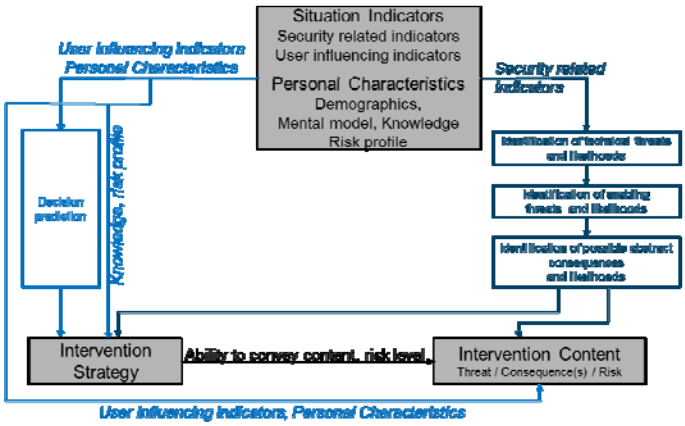

Figure 1: Framework for contextualized Security Interventions

FOCI should support developers in developing an *algorithm* to deduce for individual situations and users the appropriated intervention strategy and the appropriate content if an intervention is displayed which allows conveying content. Such an algorithm can either be integrated in future Browsers or available as an add-on.

As input for the algorithm both disciplines need to *identify and weigh relevant situation indicators and personal characteristics*, respectively. Based on these indicators the psychologists develop an algorithm *to predict the user's decision* without any further intervention (the idea is not to replace existing passive interventions, like the green URL bar). The idea is to only confront the user with additional passive or active interventions if the user would make an unintentional risky decision without. Security experts need to develop algorithms to deduce the

overall *risk level, possible abstract consequences and their likelihoods* based on the identified security related indicators. Bartsch and Volkamer (2012) recommend to conduct user studies, literature reviews and to consult experts both to identify and weighing situation indicators and personal characteristics. In addition, for each of the situation indicators, researchers need to define how these indicators can be deduced. It is recommended that ideally the user is only involved in the setup phase of such a tool by answering some questions concerning his personal characteristics and web pages he has accounts for. Note, already in this stage such a tool would check whether the provided URLs/domains are critical with respect to password transmissions. Therefore, after the setup phase, it can be assumed that on these web pages either no problem exists or the user accepts this risk.

Correspondingly, situation indicators are ideally automatically deduced when visiting a web page. In Bartsch and Volkamer (2012), it is recommended to conduct a cost benefit analysis when thinking of involving the user in determining situation indicators, e.g. by asking him what type of web page he tries to access or whether he plans to login on this page. Correspondingly, it is recommended to first evaluate all the situation indicators which can automatically be determined and all information from the setup phase. Then it needs to be decided whether the intervention strategy and content can be based on this information or whether additional information is required to support the user adequately. As the number of possible combinations of situation indicators and personal characteristics is large, it is recommended to start with concrete scenarios. In this paper we operationalized a first set of concrete scenarios, that is, visiting web pages with password requests (password fields appear on the web page). We focus on the strategy rather than on the content.

3. Adapted approaches for proposed strategy

For the operationalization of our framework we integrate, combine, adapt, and refine the results from the following papers: In 2009, Sunshine *et al*. introduced the concept of multi-page warnings. The idea is, once the browser detects a security risk, a first dialog asks the user which type of web page he wants to visit. Depending on his choice, the user is either shown a warning or not. With this multi-page warning approach, the authors wanted to achieve, that the warning is only shown if the user stated that he wants to reach a web page with a higher criticality (which was either “Bank or other financial institution” or “Online store or other e-commerce website”). One main result of the study is that multi-page warnings are a promising approach for future frameworks. Seikel, (2012) extended this work by systematically identifying a set of seven web page categories for the dialog. The author further proposes to provide the user the option to store this assignment of the corresponding URL. In addition, it is proposed that, if the user selects the option ‘information web pages’ but, then he, later on, tries to login on this page, he is warned again.

The second promising approach in improving the effectiveness of browser security warnings, in particular in the context of phishing, was proposed by Maurer *et al.* in (2011a and b). The main idea is to only show the warning when the user starts entering sensitive data and at the precise spot of the browser where the data is entered. Note, their plug-in disables the original SSL browser warnings. Thus the

user is able to visit web pages without being warned even if the browser detects any problems with SSL. Showing these so called “semi-blocking” warnings is a very interesting approach, while it underlies the same problem as the passive security indicators: Namely, if the ‘look and feel’ of the web page appeals to the user, it is more likely that he will continue entering sensitive data, despite the warning as shown by Gutmann (2011).

4. Situation indicators and personal characteristics

In this section, situation indicators and relevant personal characteristics are identified. Furthermore, it is explained how these indicators can be deduced.

Security related indicators. In order to identify the possible consequences and the risk in a particular situation, we need to identify first the relevant security related indicators for the considered password transmission scenarios. We systematically reviewed important security, Web, and HCI conferences for papers on Web security measures and identified the following security indicators:

- contains password field: Yes / No?
- https: Yes / No?
- valid certificate: Yes / No?
- self-signed: Yes / No?
- CA known: Yes / No?
- trustworthiness rating from external services for this web page such as Web of Trust, McAfee SiteAdvisor, and google safe browsing.

In order to deduce possible (enabling) threats and consequences as well as the risk level the following information is required according to the literature:

- User’s intention to login
- Type of web page: Seikel (2012) systematically deduced the following seven types which we use for this paper: Information Site, Shopping, Online Banking, Social Network, Email, Data Exchange, and Others.

While the first list of security indicators can be automatically deduced, the last two indicators are more difficult to elaborate on. In worse case the user would be bothered with two corresponding questions whenever visiting a web page containing a password field and failing for any of the other security related indicators. Obviously, these costs are too high as users would be very likely to uninstall the tool. However, the tool can take the list of web pages listed in the setup phase into account and assuming that he only logs in on these pages. Furthermore, it can learn from previous actions (including creating new accounts) or decisions of a user on particular web pages, i.e. store this information in a history. Thus, once the user has answered the question about user’s intention and type of web page for a particular URL or domain, this is stored and is used as input next time the same web page is visited. This reduces the number of interactions a lot. Therefore, we consider the

- ‘URL/domain – login – web page type‘ history

as further security related indicator.

Furthermore, *natural language processing* (NLP) techniques are proposed to be integrated to deduce the type of web page. Note, such a solution would in some cases

come to its limits: For instance ‘www.google.com’ provides many different services, with different criticalities and different consequences. For example, on the one hand Google can only be used as a search engine, and on the other hand it provides email services and a social network.

User influencing indicators. We systematically reviewed important security, Web, and HCI conferences for papers on user influencing factors. The main user influencing factor (if no warning is shown) is the design of the page (Fogg *et al.* 2001, Schechter *et al.* 2006). Furthermore, (Fogg *et al.* 2001) showed that user consider the type of web page. However, the type has not necessarily an influence in the same direction; i.e. some users are more concerned and more careful on e-banking pages as others are less because they for instance believe the bank takes care of their security. Similar to the type of web page it also has an influence whether the user knows the company. Correspondingly, the user influencing indicators are:

- Design of the web page (similarity to known pages, general design)
- Type of web page
- User knows the company/service

In order to get the impression about the design and thereby the trustworthiness of this particular user it would be necessary to ask the user. As the costs are obviously too high to ask him for each web page potentially causing threats and containing a password field, this needs to be automatically deduced (as far as possible and even if this will not exactly match to the user’s impression). We propose to analyse whether the web page looks similar to one of the pages the user has an account at (known from the setup phase or the history) and take this similarity aspect as user influencing indicator. In addition, we propose to extend web accessibility evaluation tools to use them to automatically deduce the general design quality of the web page.

The challenges and possible solutions to deduce the web page type, we discussed already with the security related indicators and refer the reader there.

Which companies are known by the individual user can only be answered by the user itself. Of course one can consider here again those pages he has an account at (from the setup phase and history). In addition, one could think of the TOP 100 visited web pages but it is not known whether these are also known to most of the popularity. Correspondingly, we propose to only take companies from the setup phase and the history into account to decide whether someone knows this company/service or not.

Personal characteristics. The effect each influencing factor has (more likely to ignore intervention or not) depends on the personal characteristics. According to the literature these are age, education, ownership, Internet/security knowledge, Internet usage (which applications, since when, and how often), and number of web-enabled devices; as well as situation specific once: namely the perceived risk, expected benefit, and the rating of expected risk. This is obviously a rather complicated field. Thus, we take only scenario specific characteristics into account which are:

- Risk ranking of different web page types (as part of their mental model)
- Knowledge about passwords

We propose the following approach to collect this information: users are ask in the setup phase to sort different web page types according to the risk they perceive if

someone else has access to their account on this web page. In addition, users are asked to participate in a small quiz in order to distinguish between experts and laypersons; while experts are those who know about the consequences if passwords are transmitted unencrypted or to an unauthorized service; and who use different and where recommended secure passwords. This will result in a very small group of experts however for the strategy whether to interfere or not it should be sufficient to distinguish these two groups. If we later also consider the type of intervention then we might need to distinguish between more groups. Note, the development of such a quiz as well as the question whether more groups are adequate is left for future work.

5. Weighing security related indicators and decision prediction

Threats, Consequences, and Risk. Assuming, all the security related indicators can be determined (automatically or by asking), then according to FOCI, the next step is to deduce possible (enabling) threats and consequences as well as objective risk levels. Note, we consider here only those URLs that are not contained in the setup list and not in the history file. Those are treated differently: The tool only intervenes if one of the security-relevant indicators changes.

We propose in Table 1 **risk** levels for different results for different security-critical indicators while assuming we know the user's intention is to login (L – Low, M – Medium, H – High, and X is the risk level according to the other services). If more than one of the five indicators about the connection to the server and the server itself fail then this results in the highest risk assigned to any of these indicators. Note, this table is based on our own opinion while it is recommended in future to ask several experts to fill out this table and compute the average risk level per entry.

	http	Invalid certificate	Self-signed certificate	CA unknown	Other services
Info	L	M	L	M	X
Shopping	M	H	M	H	X
Banking	H	H	H	H	H
OSN	M	H	M	M	X
E-mail	M	H	H	H	X
Data	M	H	M	M	X
Others	M	M	M	M	X

Table 1: Risk level

The possible (enabling) threat if logging in is for all cases identity theft at this account and if the same password is used for other accounts at these accounts as well. Possible **consequences** are: attacker gets access to private data (e.g. shopping history, salary, photos, and email history), losing money, and different types of nuisance. Note, we will investigate more in the concrete consequences when developing proposals for the content of interventions in future.

Decision prediction. The algorithm needs to take into account the personal characteristics and the user influencing indicators. From the literature it is not possible to predict the decision based on corresponding input data. It is rather likely that experts would verify that https in place when logging in on an e-banking or e-shopping page independent from the design and whether they know the company or not. However, as there is no paper that clearly states this, we recommend conducting a corresponding user study for clarification as future work. For this paper, we assume that everyone would log in on any web page based on the existing passive warnings and if not active warning appears. Thus, it is necessary to support the user with additional security interventions.

6. Intervention strategy

General strategy. While all users need support in detecting potentially critical password transmissions, laypersons also need information about the situation and consequences if deciding to login despite the warning. Thus, we propose to use passive interventions for experts (as this seems to be sufficient in many situations) and active once for laypersons. Note, as such we would also educate users. As the consequence, the number of active interventions can be reduced over time as laypersons learn more and more about the consequences in different situations.

Algorithm description: we propose the following steps:

- Security check (password field and any other security indicator)
    - If at least one causes a problem continue
- Check history/setup for intention and type
    - If in history compare whether security decreased
        - If increased warn
        - If not reaction according to setting in history
    - If not continue
- Get web page type from NLP check with corresponding probability
- Get personal characteristics from setup

Afterwards it needs to be decided whether it is necessary to ask the user about his intention and the type of webpage. For experts (CASE 1.E) we propose to use only a passive indicator namely a red background colour for the password field. However, once the user clicks on the password fields and starts entering a password a warning according to the proposal of Maurer *et al.* in (2011a and b) will appear. Note, the content of this active intervention will be different from Maurer et al.'s proposal. For laypersons (CASE 1.L), the tool proceeds in the same manner as for (CASE 1.E) if the result of the NLP check is 'information page' with a high probability.

If the NLP check results for a layperson in one of the other five web page types with a high probability (CASE 2), the user will see a dialog mentioning that there might be a security risk and asking whether the user plans to log in. If he answers yes (CASE 2.YES), the tool displays a warning informing about the concrete risk and consequences when logging in. Note, the consequence will depend on the type of web page. If the user answers no (CASE 2.NO), the background colour of the password field is set to red like in (CASE 1.E) and the tool is set into a `read only'

mode (which is not visible to the user) according to Seidel (2012). In this status the user is able to surf at this domain and read and search for information without being disturbed. But when he tries to login, a warning pop-up is displayed at the precise spot of the browser where the data is entered, again, according to the proposal of Maurer *et al.* in (2011a and b). The warning reminds the user of the dialog displayed earlier and inform the user about the risks and consequences if he decides to login. Afterwards he can still decide whether to login. With this third stage we also address the problem of false positives because the second warning would appear very rarely and so, the effect of habituation is avoided.

If the NLP check is not able to deduce the type of web page (CASE 3), lay persons are asked first whether they intent to log in similar to (CASE 2). If the answer is no (CASE 3.NO) the tool will continue as in (CASE 2.N). If the answer is yes (CASE 3.YES), the user is asked in a second dialog on which web page type he tries to login. Next the tool displays the same type of warning as in (CASED 2.YES).

Note, within all dialogs and warnings the user can store the answer in the history for this domain/URL.

7. Conclusion and Future Work

We started operationalizing the Framework fOr Contextualized security Interventions. We focused on password requesting scenarios. According to the framework, we identified security relevant indicators, user influencing indicators and relevant personal characteristics. We further described how these indicators can be deduced from the context and in the setup phase. Afterwards, we proposed an algorithm to identify threats, consequences and the risk as well as a simplified algorithm to predict the user's decision without intervening. Finally, the intervention strategy was proposed.

As future research, we will use FOCI to deduce systematically content for the security interventions. Here, we will take existing literature into account, e.g. Kauer *et al.* (2012) who observed that the wording of warnings should address the personal risk of the user and possible concrete consequences; and Raja *et al.* (2011) who showed that physical mental models helps improving the efficiency and understandability of computer warnings. In a next step, the tool should also provide recommendations how to proceed; e.g. in the situation that users already have an account on this page, need to get access to the data, or even might notice that they use the same password for another security critical web service. In addition, we will have a closer look on those web pages on which the user cannot login on the first page but only on subpages, e.g. after having made selections what to buy. Note, currently, the tool would only be activated once the page with the password field is loaded which makes it according to Gutmann (2011) harder for the tool to convince the user that there is a problem when logging in. Finally, we will integrate learning mechanisms, i.e. that the strategy and content is adopted over time as we expect people to learn how to behave more securely. Afterwards, the proposed algorithm will be implemented as add-on for Firefox and then tested in a lab study and later in a field study.

8. References

Bartsch, S. and Volkamer, M. (2012). *"Towards the Systematic Development of Contextualized Security Interventions"*. In Designing Interactive Secure Systems, BCS HCI 2012, BCS eWiC repository.

Bravo-Lillo, C., Cranor, L.F., Downs, J., Komanduri, S. (2011a).*"Bridging the gap in computer security warnings: A mental model approach*". IEEE Security and Privacy, pages 18-26.

Bravo-Lillo, C., Cranor, L. F., Downs, J., Komanduri, S., Sleeper, M. (2011b). *"Improving computer security dialogs"*. In Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction – Volume Part IV, INTERACT'11, pages 18-35. Springer.

Cranor, L. F. (2008). *"A framework for reasoning about the human in the loop"*. In Proceedings of the 1st Conference on Usability, Psychology, and Security, pages 1-15.

Dhamija, R. , Tygar, J. D. , Hearst M. (2006). *"Why phishing works"*. In: Proceedings of the SIGCHI conference on Human Factors in computing systems, pages 581-590. ACM.

Fogg, B. J. , Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J. , Rangnekar, A., Shon, J., Swani P., Treinen, M. (2001).*„What makes web sites credible?: a report on a large quantitative study"*. In Proceedings of SIGCHI conference, pages 61-68. ACM.

Gutmann, P. (2011). *"Security and usability fundamentals"*. http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/de//pubs/archive/32872.pdf (Accessed 02/ 2013).

Kauer, M., Pfeiffer, T., Volkamer, M., Theuerling, H., Bruder, R. (2012). *„It is not about the design - it is about the content! Making warnings more efficient by communicating risks appropriately"*. GI Sicherheit 2012, pages 187-198.

Maurer, M.-E. , De Luca, A, Hussmann, H. (2011a). *"Data type based security alert dialogs"*. In: Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems, CHI EA '11, pages 2359-2364. ACM.

Maurer, M.-E., De Luca, A., Kempe, S. (2011b). *"Using data type based security alert dialogs to raise online security awareness"*. In: Proceedings of the Seventh Symposium on Usable Privacy and Security, SOUPS '11, pages 2:1-2:13. ACM.

Raja, F., Hawkey, K., Hsu, S., Wang, K.-L., Beznosov, K. (2011). *"Promoting a physical security mental model for personal firewall warnings"*. In annual conference extended abstracts on Human factors in computing systems, pages 1585-1590. ACM.

Sasse, M. A. , Brostoff, S. . Weirich, D. (2001). *"Transforming the 'Weakest Link' - a Human/Computer Interaction Approach to Usable and Effective Security"*. In Technology Journal, Vol. 19, No. 3. pages. 122-131.

Sasse, M. A., Krol, K., Moroz, M.(2012) *"Don't work. Can't work? Why it's time to rethink security warnings*", 7th Intern. Conf. on Risks & Security of Internet & Systems, pages 1-8.

Schechter, S. E., Dhamija, R., Ozment, A., Fischer, I. (2007). *„Emperor's new security indicators: An evaluation of website authentication and the effect of role playing on usability studies"*. In IEEE Symposium on Security and Privacy.

Seikel, C. (2012), *"Categorization of websites according to the risk during usage"* Bachelorthesis, Technische Universität Darmstadt.

Sotirakopoulos, A., Hawkey, K., Beznosov, K. (2011*)."On the challenges in usable security lab studies: Lessons learned from replicating a study on SSL warnings"*. In: the Seventh Symposium on Usable Privacy and Security, ACM.

Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., Cranor, L. F. (2009). *"Crying wolf: an empirical study of SSL warning effectiveness"*. In: Proceedings of the 18th conference on USENIX security symposium, pages 399-416.

West R. (2008). *"The psychology of security"*. *Commuication of ACM* 51, 4; pages 34-40.