Studying Users' Willingness to Use a Formally Verified Password Manager

Carolina Carreira

INESC-ID and IST, University of Lisbon, Portugal

Abstract. Password Managers (PMs) help users manage their passwords safely but many users do not trust them. To mitigate users' doubts, formal verification can be used. Formal verification can guarantee the absence of errors and make PMs more reliable. Nonetheless, the impact it has on the adoption of formally verified software is unknown. In previous work, we performed a preliminary user study which suggests that formal verification increases users' willingness to use PMs. However, a large-scale study is required to confirm our findings. As such we designed and plan to deploy a large-scale study to confirm our previous work and gather further insight on users' perceptions of formal verification in PMs.

Keywords: Usable Security · Formal Verification · Password Manager.

1 Introduction

While text passwords are one of the most used security mechanisms, users fail to use them effectively and safely [10,12,8,14]. To combat this, experts recommend the use of Password Managers (PMs) to help users generate and manage their passwords. However, their adoption is low as users do not trust PMs [13]. Formal verification can provide strong assurances, making software more reliable. Previous uses of formal verification in password security include the creation of certified password composition policy (PCP) enforcement software [3] and the use of Coq to model PCPs [6].

A formally verified PM that guarantees properties (e.g. on password generation [5,4,1]) was built in the context of the PassCert project[1]. Even though formal verification could help increase users' trust, we do not know the impact it effectively has on users. Therefore, we designed the first user-studies on users' perceptions of formal verification. Preliminary results from a first study suggest that formal verification has a positive impact on users' willingness to use PMs [1]. A second, larger-scale, study is now being designed.

Our main goals are to gather insights on users' perceptions of formal verification in PMs and to assess if formal verification has an impact on their willingness to use PMs.

[1] The PassCert project aims to build a formally verified PM and to investigate ways to effectively convey to users the formally verified properties. Project URL: `https://passcert-project.github.io`

2 Current Work

Our current work gathers conclusions from two studies on PMs and formal verification. This section briefly describes these two studies.

2.1 First Study

In the first small-scale study with 15 users, we compare a baseline PM (without formal verification) with a PM that includes visual aids (icons) to highlight formally verified features and brief explanations about them. Our goal was to gather preliminary insights on users' overall perception of formal verification in PMs. The emerging themes from the interviews were: (a) Users associated formal verification with security; (b) The use of formal verification may have increased some users' trust; (c) Users may be more willing to use a formally verified PM [1].

2.2 Second Study

To confirm the preliminary results obtained in the first study, we designed a large-scale study focused on the impact that formal verification has on PMs' users. Specifically, in this second study, we aim to answer:

RQ1. How does formal verification impact users' willingness to use PMs?
RQ2. What features would users like to see formally verified in a PM?
RQ3. Do users value the guarantees that formal verification can provide in PMs?

The design process of this study includes the use of techniques more adequate for large-scale surveys such as Likert scales and closed survey questions. To immerse users in the topic of a formally verified PMs we use vignette scenarios, which describe a protagonist faced with a realistic situation pertaining to the construct under consideration [7]. This study will be deployed in Prolific[2].

To answer RQ1 we present a scenario where we explain what is a PM and we ask users what factors impact their willingness to use a PM. Among these, we include formal verification. If users state that formal verification would affect their willingness to use a PM, we ask why. With this question, we hope to understand why users value (or not) the use of formal verification in PMs. The insights gathered here may provide relevant information about how users perceive formal verification in software and may be applied to other domains (e.g., Decentralized Finance Protocols [2]).

To answer RQ2 and RQ3 we begin by gathering all the common features of a PM (e.g. Password Generator and Clipboard clearing). For each of these, we present scenarios that represent the impact that formal verification can have on each feature. For example, for the Password Generator, the scenario is: "*Imagine that you are creating a new account on a website (e.g. twitter, facebook). To*

[2] Prolific is a crowd-sourcing platform that enables large scale user studies by connecting research and users `https://prolific.com`

increase security, you ask the Password Manager to generate a password with 7 characters and with at least 2 numbers. However, the password generated does not include any numbers." After each scenario, we ask users if that scenario would make them stop using a PM with a 5-point Likert agreement scale.

To minimize the introduction of biases, when designing these scenarios we: (a) remove mentions of formal verification; (b) randomize the order of the scenario descriptions; (c) remove jargon (e.g. "memory" and "encrypted"). By presenting the advantages of formal verification and excluding the term "formal verification" we aim to: (i) mitigate the Hawthorne effect[3] by hiding that the study is about formal verification; (ii) and better understand what specific advantages of formal verification users find important in PMs.

Another important concern is the sample of participants taking part in the study. To help characterize it we ask demographic questions (e.g., age, gender, and ethnicity) and questions specific to our study, including questions about users' previous experience with PMs.

Users' perceptions when using a product are influenced by their assumptions about it (e.g. previous experience or recommendations from friends can shape users trust in a website [11]). As we are studying the impact of formal verification, it is thus important to understand if users are familiar with the concept. With this goal in mind we ask questions about users' familiarity with the term "formal verification" and ask them to define it.

3 Conclusion and Impact

Investigating users' views of formal verification is largely unexplored. We hope to fill a gap in knowledge with the first large-scale user study on users' views of formal verification in PMs. Our work will provide insights on users' motivations and may be used to increase the adoption of PMs.

Correctly identifying where formal verification is valued by users will help understanding the priorities for future implementations of formally verified features in PMs. These insights may lead to: (i) the formally verification of features not yet formally verified; (ii) and, a higher adoption of PMs by matching the users' preferences with the software that is offered to them.

We also anticipate that our findings can be applied to other domains where formal verification is used. Learning about users' current perceptions of formal verification will enable us to address identified issues and misconceptions [2]. Moreover, our methodology can easily be replicated in other domains by adequately adapting the scenarios mentioned in Section 2.2.

[3] Hawthorne effect consists in users being inclined to agree with researchers [9]

Acknowledgments I thank João F. Ferreira, Alexandra Mendes, Nicholas Christin, and Sarah Pearman for their valuable and constructive support. This work was partially funded by the PassCert project, a CMU Portugal Exploratory Project funded by Fundação para a Ciência e Tecnologia (FCT), with reference CMU/TIC/0006/2019 and supported by national funds through FCT under project UIDB/50021/2020.

References

1. Carreira, C., Ferreira, J.F., Mendes, A.: Towards improving the usability of password managers. INFORUM (2021)
2. Carreira, C., Ferreira, J.F., Mendes, A., Christin, N.: Exploring usable security to improve the impact of formal verification: A research agenda. First Workshop on Applicable Formal Methods (co-located with Formal Methods 2021). (2021)
3. Ferreira, J.F., Johnson, S., Mendes, A., Brooke, P.: Certified password quality: A case study using Coq and Linux pluggable authentication modules. In: 13th International Conference on Integrated Formal Methods (2017)
4. Grilo, M., Campos, J., Ferreira, J.F., Mendes, A., Almeida, J.B.: Verified password generation from password composition policies. In: 17th International Conference on Integrated Formal Methods (2022)
5. Grilo, M., Ferreira, J.F., Almeida, J.B.: Towards formal verification of password generation algorithms used in password managers. arXiv:2106.03626 (2021)
6. Johnson, S., Ferreira, J.F., Mendes, A., Cordry, J.: Skeptic: Automatic, justified and privacy-preserving password composition policy selection. In: 15th ACM Asia Conference on Computer and Communications Security (2020)
7. Lavrakas, P.J.: Encyclopedia of survey research methods. Sage publications (2008)
8. Lyastani, S.G., Schilling, M., Fahl, S., Backes, M., Bugiel, S.: Better managed than memorized? studying the impact of managers on password strength and reuse. In: 27th USENIX Security Symposium (USENIX Security 18). pp. 203–220 (2018)
9. Merrett, F.: Reflections on the hawthorne effect. Educational Psychology **26**(1), 143–146 (2006). https://doi.org/10.1080/01443410500341080
10. Pearman, S., Zhang, S.A., Bauer, L., Christin, N., Cranor, L.F.: Why people (don't) use password managers effectively. In: Fifteenth Symposium On Usable Privacy and Security. USENIX Association, Santa Clara, CA. pp. 319–338 (2019)
11. Seckler, M., Heinz, S., Forde, S., Tuch, A.N., Opwis, K.: Trust and distrust on the web: User experiences and website characteristics. Computers in human behavior **45**, 39–50 (2015)
12. Shay, R., Komanduri, S., Durity, A.L., Huh, P., Mazurek, M.L., Segreti, S.M., Ur, B., Bauer, L., Christin, N., Cranor, L.F.: Can long passwords be secure and usable? In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 2927–2936 (2014). https://doi.org/10.1145/2556288.2557377
13. Silver, D., Jana, S., Boneh, D., Chen, E., Jackson, C.: Password managers: Attacks and defenses. In: 23rd USENIX Security Symposium. pp. 449–464 (2014)
14. Ur, B., Noma, F., Bees, J., Segreti, S.M., Shay, R., Bauer, L., Christin, N., Cranor, L.F.: "I Added '!' at the End to Make It Secure": Observing password creation in the lab. In: Eleventh Symposium On Usable Privacy and Security. pp. 123–140 (2015)