Exploring Usable Security to Improve the Impact of Formal Verification: A Research Agenda

Carolina Carreira
INESC-ID and IST,
University of Lisbon, Portugal

João F. Ferreira
INESC-ID and IST,
University of Lisbon, Portugal

Alexandra Mendes
HASLab, INESC TEC and
Universidade da Beira Interior, Portugal

Nicolas Christin
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

As software becomes more complex and assumes an even greater role in our lives, formal verification is set to become the gold standard in securing software systems into the future, since it can guarantee the absence of errors and entire classes of attack. Recent advances in formal verification are being used to secure everything from unmanned drones to the internet.

At the same time, the usable security research community has made huge progress in improving the usability of security products and end-users comprehension of security issues. However, there have been no human-centered studies focused on the impact of formal verification on the use and adoption of formally verified software products. We propose a research agenda to fill this gap and to contribute with the first collection of studies on people's mental models on formal verification and associated security and privacy guarantees and threats. The proposed research has the potential to increase the adoption of more secure products and it can be directly used by the security and formal methods communities to create more effective and secure software tools.

1 Introduction

Formal verification can be used to secure software systems by guaranteeing the absence of errors and entire classes of attack — recent advances are being used to secure everything from unmanned drones to the internet [5]. However, formally guaranteeing security properties does not guarantee that end-users will trust the verified systems. Since the seminal paper by Whitten and Tygar [26], usability problems — defined to include human factors, such as mental models — have been identified as a major factor in users disregarding security mechanisms. For example, despite some experts recommending password managers, the adoption of these tools is still low, partly because users distrust them [18].

While progress has been made in improving the usability of security products and end-users comprehension of security issues, there have been no human-centered studies focused on the impact of formal verification on the use and adoption of formally verified software products. Given the recent and increasing industrial adoption of formal verification, the time is ripe to perform these studies. The outcomes of the proposed research can increase the adoption of formal verification to secure software systems and change users' behaviors when it comes to adopting security technologies.

In this research agenda, we expose this gap in knowledge and propose several future research goals in three parts. First, we discuss the current perception that users have of formally verified products, secondly how to shape these perceptions, and, thirdly, research methodologies. We conclude our presentation by exposing specific examples of two important domains that have a massive user base (passwords) or are projected to grow immensely (cryptocurrency systems and DeFi protocols), and that may benefit from the proposed research.

Submitted to:
AppFM 2021

2 Background and Current Problems

Security normally adds complexity to a system and interferes with the user's primary goals. Two-factor authentication is a good example: it adds complexity to the authentication process and is strongly recommended for bank applications, but some users will be unsatisfied if such procedures are required for unimportant accounts [22]. Whitten and Tygar [26] analyzed email encryption as a usability problem in their seminal paper written in 1999 and helped establish the area commonly known as "usable security". In the last decade, there has been a lot of work on usable security, across many domains [10, 16, 18].

On the other hand, formal verification has been used to guarantee security properties of many software systems. Well-known examples include correctness and security verification of OpenSSL HMAC [1] and HMAC-DRBG [28], verification of TLS and other components of HTTPS [2], and high-assurance software for unmanned aerial vehicles [5].

Despite all these developments, there have been no human-centered studies focused on the impact of formal verification on the use and adoption of formally verified software products. There is a wide array of factors that influence users' adoption, retention, and usage of software. These factors can be a source of problems or undesired user behaviors. They include, among others: 1) **Perceptions**, being in regards to perceived privacy, usability or satisfaction (e.g., Pearman et al. [18] cited perceptions of increased security as a reason for the adoption of password managers); 2) People commonly construct implicit **mental maps** to understand complex systems when the systems' functionality goes beyond their technical knowledge [13], and such tacit knowledge influences human behavior, even unknowingly; 3) A products' **usability** can impact **adoption** rates (e.g. a reason for low adoption of security software in the past has been usability problems [18]); 4) Lack of **knowledge** about a product can provide a false sense of security or discourage users from using it: for example, Gao et al. [6] found that bitcoin users' lack of knowledge about the technology presented an entry barrier for new users; 5) **Motivation** is also a crucial factor in encouraging users to overcome the initial overhead of using new software [22].

It is known that using a well-designed security mechanism is still more effort than not using it at all, and users will always be tempted to cut corners [20]. Hence, it is the developers' and designers' job to convince users to use their products. We thus believe that human-centered studies regarding end-users mental models, cognitive biases, common misconceptions, and general perceptions of formal verification are necessary to improve the adoption and impact of formally verified products.

3 Research Agenda

In this section, we discuss three main future research paths for human-centered studies on formal verification. First, we propose studying the current perception users have of formally verified products, secondly how to shape this perception, and thirdly relevant research methodologies. For each path, we present the main research questions to be addressed.

3.1 Understanding the impact of formal verification

As users implicitly construct mental models of systems [13], their behavior is affected, even unknowingly. We claim that performing a thorough study of mental models on formal verification, similar to studies on mental models about other technologies [16], is necessary and important to understand users' misconceptions and their grasp of what it means for a product to have formally verified features.

Formal verification of software can guarantee the absence of errors and entire classes of attack. This is a major selling point of formally verified software and can affect users' choices when deciding what

software to use. As such, it is important to understand users' perceptions of formal verification and whether it has an impact on their adoption of software products. Additionally, users may value the use of formal verification in some types of products, but not in others. For example, in a previous study, secure access to financial accounts was valued above other types of online accounts [18]. It is thus important to survey in what type of software products formal verification has a greater impact on users.

Misconceptions about underlying concepts and processes are common and can have serious effects on trust and adoption of products. For example, Pearman et al. [18] studied users' perceptions of password managers and found that lack of knowledge about how they worked posed a barrier to their effective use and adoption. In a different study, Mai et al. [16] studied users' mental models of cryptocurrency systems and found a misconception where many participants presumed their transactions could not be tracked due to the encryption used; this is not true and gave users a false sense of privacy. If users hold misconceptions about the role of formal verification in a system, these need to be identified and addressed. It is also relevant to know if users find that formal verification makes products safer.

Learning more about users' perceptions of formal verification is crucial for formal methods practitioners, as with this information they can better understand the impact their products and developments have on the end-user. With this knowledge, it is easier to convey the **usefulness** of formal verification.

The main research questions to be addressed in this line of work are:

Main research questions:
- How do users perceive formal verification in software products?
- What are users' misconceptions about formally verified products?
- What is the impact of formal verification on the adoption of software products?
- In what type of software products does formal verification have a larger impact on users?

3.2 Shaping users' perception of formal verification

After learning about users' current perceptions of formal verification concepts, we can address identified issues and misconceptions. As stated before, usability problems have been proved to be a barrier to user adoption of software products and users have a hard time understanding security concepts or how to be secure. For example, despite experts recommending password managers, some users report unawareness of their existence as a reason to not use them, and even when users are aware of their existence, some are still reluctant in using them due to a lack of understanding of their security properties [18]. This lack of knowledge can have a significant impact: as mentioned above, Mai et al. [16] found that a poor understanding of the cryptocurrency system being used exposes users to privacy risks.

We claim that it is important to **convey information** to users to make them aware of the existence of formal verification and the advantages it can provide. Knowing its advantages can motivate users to choose formally verified products. This information can be general (e.g. educating users about what formal verification is) or domain-specific (e.g. helping users understand what specific features are formally verified and what it means for them).

From the perspective of the user interface, several approaches can be taken to convey important information. Examples include the use of support/help tools like tooltips, tutorials, informative pop-ups, icons, wiki pages, and frequently asked questions pages [11, 22]. For example, in the PassCert project[1], we are exploring all of the above to convey to users what properties are guaranteed and what value it brings to them. An important point is to be careful to not overstate the formal guarantees and

[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`

to ensure that users understand that a formally verified product might still be vulnerable. From a more general perspective, there is an opportunity to raise awareness through advocacy in schools and the media. Educating users, by explaining formal verification usage and associated security and privacy guarantees, or by describing the dangers of unsafe software, could serve as motivation for the adoption and long-term retention of users in formally verified software. The study of what are the best ways to effectively convey formal verification concepts to users and how this information impacts their adoption is crucial and valuable to the formal methods community.

Regarding implicit knowledge maps, the construction of adequate mental models about formal verification should be enabled by the suggestions described above, but there are other generic usability techniques that should not be disregarded, such as the implementation of clear navigation systems and other interface design choices (e.g. following the Eight Golden Rules by Shneiderman et al. [22]). A survey of specific usability principles that should be applied to formally verified products is missing. Usability can be used to enable users' understanding of formal methods topics and ensure the end-product is usable, this allied with a well-implemented interface can influence users' perceptions on formal verification.

The main research questions to be addressed in this line of work are:

Main research questions:
- How can we effectively convey formal verification concepts to users?
- Does users' understanding of formal verification concepts have an impact on the adoption of formally verified software?
- What usability principles should be followed to ensure adequate mental models of formal verification?

3.3 User testing methodologies

We now turn our attention to **how** we can acquire information on users' mental models, perceptions, and understanding of formal verification. For this, existing techniques can be used, such as: 1) **Questionnaires**: several studies regarding users' perceptions of software products have used this method with success [10,18,19]. Questionnaires can also be combined with other user testing tools such as interviews and usability tests [11,22]. When scaled up to encompass a large number of users using online services, questionnaires are a strong quantitative research tool [11]. 2) **Interviews**: these can provide detailed and qualitative information. They can be structured or non-structured and are also commonly used [6,10,18]. One-on-one interviews offer more qualitative results that can serve as the basis for understanding users' expectations, vocabulary, goals, and perceptions [11]. Both techniques can and should be adapted to particular products or testing goals and their outputs could be used both to better understand users' perceptions of formal verification and as feedback about a possible formally verified product. A suggestion we give is to survey these methodologies to understand how to better apply them and understand the impact of formally verified software. Biases that might affect results need to be considered, such as the Hawthorne effect where users may be inclined to agree with the researchers [17].

As stated in the previous section, it is also relevant to learn about the usability of any formally verified product. An iterative design process, with several rounds of user testing, is recommended [22]. Within this process several methods can be used, for example:

- *By experts* – heuristic evaluations, cognitive walkthroughs where experts compare the interface with a set of heuristics rules and simulate users walking through the interface;
- With users *during development* – early user studies can be done with prototypes, by asking users to perform tasks while *thinking aloud* and iterate over that feedback; other tests include competitive

testing to compare different interface versions. Nowadays, remote usability testing has gained popularity as it can be done to a large number of users through online communities including Amazon Mechanical Turk [22].
- With users during *active use* – interviews and focus-group discussions can be productive because the interviewer can pursue specific issues of concern to help in better understanding the users' perspectives [22]. Software should also provide developers with a continuous data logging of user performance and supply information about patterns of use.

These are a few examples of evaluation methods and techniques that can be used. Typical usability tests measure how well the user is able to perform certain tasks in the interface, and its ease of use. Here, we need user studies that go further and provide information about users' perception, understanding, and retention when considering formal methods aspects. The **best methods to learn users' perception of formal verification and how it can affect usability are still unknown** and we claim that there is a need for building solid and replicable user research methodologies that meet our specific goals, provide useful data, and that can be applied to different domains where formal verification is used.

The main research questions to be addressed in this line of work are:

Main research questions:
- How can we effectively test users' understanding of formal verification?
- What tools can be used to learn about the usability of formally verified products?

4 Promising Next Steps

Next steps for this research include the application of the previously described generic research questions to concrete domains. We believe that any domain where formal methods are currently being applied can be considered. However, to maximize impact, we propose two important specific domains: one that has a massive user base (password security) and another that is projected to grow immensely in the near future (cryptocurrency systems and DeFi protocols). In this section, we describe these two domains and, for each, we provide background research, motivation, and research suggestions. It is important to note that these are just two examples from a large set of possible future research domains.

4.1 Password security

Text passwords are one of the most used security mechanisms. However, users often struggle to remember passwords [9, 21, 27]. This results in users frequently reusing passwords across multiple accounts [15]. Also, entering long or complex passwords on mobile devices leads users to use weak passwords [25]. Experts recommend password managers with random-generation features to help users employ strong and unique passwords. However, studies have shown that password managers have low adoption rates, especially among non-experts [10]. Many users do not trust password managers [23] and several usability problems have been found [18].

Following previous work on formal methods applied to password security [4,12], a verified password manager that assures properties on data storage and password generation is being built in the context of PassCert [7], where we are currently studying the effectiveness of status symbols and tooltips that explicitly indicate that certain actions or properties are formally assured. One of the end goals is to propose ideas to integrate non-obtrusive information and documentation on formal verification in password managers. We propose an expansion of this work by: 1) studying in detail users' mental models and common misconceptions on both formal verification and password managers; 2) studying the impact of formal

verification on user acquisition/adoption in this domain; 3) applying relevant usability tools (e.g. status icons) to improve communication on the assurances provided by formal verification; 4) performing large-scale studies that can provide statistical significant actionable findings that can increase the adoption of important tools such as password managers; and 5) surveying user research methodologies to achieve the previously mentioned qualitative and quantitative research.

4.2 Decentralized finance

Of the many recent developments on formal verification of smart contracts [24], we highlight KEVM, a K semantics of the Ethereum Virtual Machine [8]. Runtime Verification Inc is currently using K to verify the Maker Multi-Collateral Dai Protocol[2]. Experts noted that formal verification eliminated many of the 'low-hanging fruit' vulnerabilities, and recommended continued use. This suggests that DeFi companies will increasingly use formal verification, providing extra motivation for the research we propose. A DeFi system enables financial services through decentralized peer-to-peer networks. Advantages include decentralization and transparency [3]. DeFi protocols have seen staggering growth recently, with 2020 hitting the milestone figure of over $12 billion in total locked assets.

Poor usability and lack of knowledge are major contributors to cryptocurrency security failures [14]. Mai et al. [16] studied users' mental models of cryptocurrency systems and found flaws and inconsistencies that expose users to security and privacy risks. Gao et al. [6] examined adopters and non-adopters' perceptions of bitcoin, and found that both lacked knowledge about it. Non-adopters also questioned the security of these systems [19].

Regarding this domain, as a next step, we propose to consider the Maker Protocol and Runtime Verification Inc's formalization. Since the Maker Protocol is complex, there are two main aspects on which we propose to focus: first, investigation of usability principles that should be followed to ensure that financial tools based on the Maker Protocol can help build adequate mental models; second, survey effective ways to convey that the protocol is formally verified and measure if (and how) this can lead to wider adoption of cryptocurrency-based products. We propose an exploration and assessment of different alternatives to inform the users about what properties are formally verified (e.g. by adding cues and visualizations). Moreover, we propose to integrate specialized information viewers that show (and explain) the classes of problems that were proved impossible to occur, thus helping users to make informed decisions.

5 Conclusion

We propose a research agenda that can create new knowledge on how users perceive critical security issues and how they perceive the value of formal verification in software security. Moreover, this knowledge can enable effective communication on the assurances provided by formal verification. As a result, its outcomes can be directly used by the security and formal methods communities to create more effective and secure software tools in most areas where formal methods are applied.

Acknowledgments. We thank the anonymous reviewers for their valuable and constructive comments. 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.

[2] https://security.makerdao.com/audit-reports#runtime-verification-specification

References

[1] Lennart Beringer, Adam Petcher, Q Ye Katherine & Andrew W Appel (2015): *Verified Correctness and Security of OpenSSL HMAC*. In: *24th USENIX Security Symposium (USENIX Security 15)*, pp. 207–221, doi:10.1145/3133956.3133974.

[2] Karthikeyan Bhargavan, Barry Bond, Antoine Delignat-Lavaud, Cédric Fournet, Chris Hawblitzel, Catalin Hritcu, Samin Ishtiaq, Markulf Kohlweiss, Rustan Leino, Jay Lorch et al. (2017): *Everest: Towards a verified, drop-in replacement of HTTPS*. In: *2nd Summit on Advances in Programming Languages (SNAPL 2017)*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, pp. 1:1–1:12.

[3] Yan Chen & Cristiano Bellavitis (2020): *Blockchain disruption and decentralized finance: The rise of decentralized business models*. *Journal of Business Venturing Insights* 13, p. e00151, doi:10.1016/j.jbvi.2019.e00151.

[4] João F. Ferreira, Saul Johnson, Alexandra Mendes & Phillip Brooke (2017): *Certified Password Quality: A Case Study Using Coq and Linux Pluggable Authentication Modules*. In: *13th International Conference on Integrated Formal Methods*, Springer, pp. 407–421, doi:10.1007/978-3-319-66845-1_27.

[5] Kathleen Fisher, John Launchbury & Raymond Richards (2017): *The HACMS program: using formal methods to eliminate exploitable bugs*. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375(2104), p. 20150401, doi:10.1098/rsta.2015.0401.

[6] Xianyi Gao, Gradeigh D Clark & Janne Lindqvist (2015): *Of two minds, multiple addresses, and one history: Characterizing opinions, knowledge, and perceptions of bitcoin across groups*. *arXiv preprint arXiv:1503.02377*, doi:10.2139/ssrn.2575796.

[7] Miguel Grilo, João F. Ferreira & José Bacelar Almeida (2021): *Towards Formal Verification of Password Generation Algorithms used in Password Managers*. *arXiv preprint arXiv:2106.03626*. Paper supporting talk given at INForum 2021 (`https://inforum.org.pt`).

[8] Everett Hildenbrandt, Manasvi Saxena, Nishant Rodrigues, Xiaoran Zhu, Philip Daian, Dwight Guth, Brandon Moore, Daejun Park, Yi Zhang, Andrei Stefanescu et al. (2018): *Kevm: A complete formal semantics of the ethereum virtual machine*. In: *2018 IEEE 31st Computer Security Foundations Symposium (CSF)*, IEEE, pp. 204–217, doi:10.1109/CSF.2018.00022.

[9] Philip G Inglesant & M Angela Sasse (2010): *The true cost of unusable password policies: password use in the wild*. In: *Proceedings of the sigchi conference on human factors in computing systems*, pp. 383–392, doi:10.1145/1753326.1753384.

[10] Iulia Ion, Rob Reeder & Sunny Consolvo (2015): *...no one can hack my mind: Comparing Expert and Non-Expert Security Practices*. In: *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, pp. 327–346.

[11] Aynne Valencia Jenifer Tidwell, Charles Brewer (2020): *Designing interfaces: Patterns for effective interaction design*. " O'Reilly Media, Inc.".

[12] Saul Johnson, João F. Ferreira, Alexandra Mendes & Julien Cordry (2020): *Skeptic: Automatic, justified and privacy-preserving password composition policy selection*. In: *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security*, pp. 101–115, doi:10.1145/3320269.3384762.

[13] Anne R Kearney & Stephen Kaplan (1997): *Toward a methodology for the measurement of knowledge structures of ordinary people: the conceptual content cognitive map (3CM)*. *Environment and behavior* 29(5), pp. 579–617, doi:10.1177/0013916597295001.

[14] Katharina Krombholz, Aljosha Judmayer, Matthias Gusenbauer & Edgar Weippl (2016): *The other side of the coin: User experiences with bitcoin security and privacy*. In: *International conference on financial cryptography and data security*, Springer, pp. 555–580, doi:10.1007/978-3-662-54970-4_33.

[15] Sanam Ghorbani Lyastani, Michael Schilling, Sascha Fahl, Michael Backes & Sven Bugiel (2018): *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.

[16] Alexandra Mai, Katharina Pfeffer, Matthias Gusenbauer, Edgar Weippl & Katharina Krombholz (2020): *User Mental Models of Cryptocurrency Systems-A Grounded Theory Approach*. In: *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*, pp. 341–358.

[17] Frank Merrett (2006): *Reflections on the Hawthorne effect*. *Educational Psychology* 26(1), pp. 143–146, doi:10.1080/01443410500341080.

[18] S. Pearman, S. A. Zhang, L. Bauer, N. Christin & L. F. Cranor (2019): *Why people (don't) use password managers effectively*. In: *Fifteenth Symposium On Usable Privacy and Security (SOUPS 2019). USENIX Association, Santa Clara, CA*, pp. 319–338.

[19] Wanda Presthus & Nicholas Owen O'Malley (2017): *Motivations and barriers for end-user adoption of bitcoin as digital currency*. *Procedia Computer Science* 121, pp. 89–97, doi:10.1016/j.procs.2017.11.013.

[20] M Angela Sasse & Ivan Flechais (2005): *Usable security: Why do we need it? How do we get it?* In: *Security and Usability: Designing secure systems that people can use*, O'Reilly, pp. 13–30.

[21] Richard Shay, Saranga Komanduri, Adam L Durity, Phillip Huh, Michelle L Mazurek, Sean M Segreti, Blase Ur, Lujo Bauer, Nicolas Christin & Lorrie Faith Cranor (2014): *Can long passwords be secure and usable?* In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2927–2936, doi:10.1145/2556288.2557377.

[22] B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmqvist & N. Diakopoulos (2016): *Designing the user interface: strategies for effective human-computer interaction*. Pearson.

[23] David Silver, Suman Jana, Dan Boneh, Eric Chen & Collin Jackson (2014): *Password managers: Attacks and defenses*. In: *23rd USENIX Security Symposium (USENIX Security 14)*, pp. 449–464.

[24] Palina Tolmach, Yi Li, Shang-Wei Lin, Yang Liu & Zengxiang Li (2021): *A survey of smart contract formal specification and verification*. *ACM Computing Surveys (CSUR)* 54(7), pp. 1–38, doi:10.1145/3464421.

[25] Blase Ur, Fumiko Noma, Jonathan Bees, Sean M Segreti, Richard Shay, Lujo Bauer, Nicolas Christin & Lorrie Faith Cranor (2015): *"I Added '!' at the End to Make It Secure": Observing Password Creation in the Lab*. In: *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, pp. 123–140.

[26] A. Whitten & J. D. Tygar (1999): *Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0.* In: *USENIX Security Symposium*, 348, pp. 169–184.

[27] Jeff Yan, Alan Blackwell, Ross Anderson & Alasdair Grant (2004): *Password memorability and security: Empirical results*. *IEEE Security & privacy* 2(5), pp. 25–31, doi:10.1109/msp.2004.81.

[28] Katherine Q Ye, Matthew Green, Naphat Sanguansin, Lennart Beringer, Adam Petcher & Andrew W Appel (2017): *Verified correctness and security of mbedTLS HMAC-DRBG*. In: *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pp. 2007–2020, doi:10.1145/3133956.3133974.