Formal Mental Models for Inclusive Privacy and Security

Adam M. Houser
Industrial and Systems Engineering
University at Buffalo
State University of New York
433 Bell Hall, Buffalo, NY 14260
adamhous@buffalo.edu

Matthew L. Bolton, Ph.D.
Assistant Professor
Industrial and Systems Engineering
University at Buffalo
State University of New York
407 Bell Hall, Buffalo, NY 14260
mbolton@buffalo.edu

ABSTRACT

Efforts to bring inclusive privacy and security solutions to disadvantaged populations will require multifaceted approaches. A key aspect of this challenge is understanding the diverse needs of the userbase, as this will help ensure the alignment of proposed solutions with these needs. One potential strategy for addressing this challenge is to rigorously explore the mental models that characterize stakeholders' privacy and security concerns. This paper will suggest a strategy to meet this challenge, drawing on approaches from human factors engineering and formal methods to establish a framework for modeling and exploring user mental models within a security context. Potential areas of exploration using this method will also be discussed.

1. INTRODUCTION

Recent geopolitical events have spurred widespread interest in consumer-grade privacy and security solutions. This has led to a broadened user base that now includes those with a diversity of technical savvy. Gone are the rarefied days of a small poweruser community wielding Pretty Good Privacy (PGP) to secure their communications. Today, the ranks of privacy- and security-conscious users have expanded to include those with a diversity of experience and use cases [15]. Modern security solutions are intended to lower the burden of deployment, enabled now by installing an app or checking a box rather than through arcane system configuration.

However, assumptions are made when developing these solutions for mainstream use, and these assumptions may not hold for a diversity of people or situations. Consider the use of encrypted communication smartphone applications that have recently enjoyed a surge in global popularity. Though perhaps not created for such purposes, some citizen protesters in repressive regimes have used them to coordinate reform movement demonstrations [25]. However, recent news reports have suggested that users installing these applications were flagged as dissidents and targeted for deanonymization attacks by state-level threat actors [16]. If solutions do not adequately align with user needs, then they may be forced to look elsewhere or adopt ill-suited solutions; in more dire circumstances, users could find themselves in danger.

One potential strategy for finding mismatches between user needs and solution capabilities is to explore the mental models of target users, borrowing techniques from formal methods and human factors engineering to discover and characterize misalignments in a rigorous manner. Synergistically integrating these techniques can help analysts find potentially unforeseen interactions between needs and capabilities, an important and flexible approach when considering whether solutions are appropriate for disadvantaged populations.

Symposium on Usable Privacy and Security (SOUPS) 2017, July 12–14, 2017, Santa Clara, California.

2. THEORETICAL PERSPECTIVES

While several disciplines have used the term "mental models" when discussing research findings, the context and details of these uses are sufficiently varied to frustrate direct comparisons. It is therefore important to characterize our use of the term, identify relevant cross-disciplinary work, and integrate our framework with modern theoretical perspectives on mental models.

2.1 Mental models in human factors and cybersecurity

There exists a substantial body of knowledge regarding the structure, development, and use of mental models in human factors engineering. Here, mental models are characterized as internalized representations of the function of a target system, where the term "system" is used to describe any collection of components with the intent of accomplishing one or more goals [17, 27].

Developments in this domain have been used to establish key features and capabilities of mental models, particularly where model form and function *vis-à-vis* the user is important [28]. For example, mental models should be "runnable," or composed in such a way that users can mentally work through their model to explain observed behavior or anticipate future system conditions. However, they may also incompletely capture all system behavior and may be susceptible to degradation through forgetfulness [17]. Users may also develop and refine several different mental models of the same system representing varying levels of abstraction or complexity [24]. Switching between them at will can be useful in a variety of scenarios, including fault diagnosis [19], problem solving [20], and the cultivation of expertise in task performance [7].

Recent work has explored similar concepts with respect to

user privacy and security, usually classifying these "mental models" as folk models (see [26], below). Similar to the mental models of human factors literature, folk models can encompass user expectations about system behavior but tend to lack sufficient detail to be runnable. For example, Camp established a series of folk models potentially useful for communicating computer security risks [6]. Asgharpour, et al., attempted to validate these findings with card-sorting tasks, finding that some models were more tractable than others among users, as well as differences in the models used by experts and novices [1]. Kang, et al., used interviews to investigate how folk models of internet security and functionality relate to user actions online [13]. Raja, et al., explored firewall warning messages leveraging risk communication strategies resultant from findings in [1] and [6].

Wash explored folk models and adherence to security advice, finding that the models with which people identified at least partially impacted the programs users employed, the habits they adopted with software maintenance, and how they behaved online [26]. Kauer, et al., replicated Wash's work in Germany, finding additional mental models among individuals studied there and suggesting that important differences may exist within seemingly similar groups of people around the world [14].

In perhaps one of the few examples of work extending folk models to what human factors engineers would consider runnable mental models, Blythe and Camp implemented select models in state transition diagrams to simulate user actions when faced with certain conditions [3]. However, because this analysis was simulation-based, it did not consider all possible interactions between user mental models and system elements. Such capabilities have been achieved in the human factors community through the use of formal methods. These are discussed next.

2.2 Analysis with formal methods

Formal methods are well-defined mathematical languages and techniques for the modeling, specification, and verification of systems [29]. System models are descriptions that leverage well-supported theoretical formalisms to represent a system's behavior [4]. Specifications are desirable system behaviors captured as propositions expressed in a mathematically precise, unambiguous manner. Upon describing the system model and specifications, verification mathematically proves (often using exhaustive searches) whether specifications do or do not hold in the sytem model. A specification is proved if it is satisfied by the model, while specification violations can return a counterexample. This is an execution trace capturing the model state in which the violation occurred and a list of incremental model states leading to the violation [2].

Researchers have developed promising techniques for examining mental models with formal methods, particularly where mismatches between user mental models and system capabilities can be dangerous for users [5]. Degani, Heymann, and others have found discrepancies between pilot mental models and autopilot [10], autoland [18], and cruise-control systems [9]. Rushby, et al., have explored similar concepts using automated model checking tools [22, 23], and Combéfis, et al., have developed methods for examining whether user mental models include enough information to allow user control over full system capabilities [8].

3. SYNERGISTIC EXPLORATION

These findings suggest novel opportunities for the analysis of inclusive privacy and security solutions by discovering unforeseen mismatches between solution capabilities and user needs. This can help developers find ways to improve solution capabilities and ensure that solutions adequately meet user needs. Furthermore, this approach could be adaptable to a diversity of populations, including those with physical disabilities and situational impairments.

Synergistically integrating these methods into a research framework may provide new insights for solution developers. First, steps should be taken to elicit and describe mental models that characterize the target user base, focusing on aspects of privacy and security within the context of those users' challenges. While extracting accurate mental constructs from users' heads can be challenging, there exist a variety of useful techniques, including card-sorting tasks [1], interviews [26], task observations [11], cognitive walkthroughs [12], and training artifact analysis [21], among others. Upon extraction, these findings can be translated into runnable formal models that will thus be explorable with both manual and automated techniques (see [10], [18], and [22]).

Models capturing system properties related to user privacy and security must also be constructed. Importantly, models used in formal analyses need only be complex enough to accurately capture system behavior at issue, rather than that of the entire system (see [10] and [21]). This helps manage the problem's statespace and properly scope targets for analysis.

Once these steps are complete, formal methods can be used to comparatively explore user needs and system capabilities. The power of this approach lies in the systematic exploration of the entire problem space, a capability driven by the application of modern verification tools. This exhaustive search allows the analyst to discover unanticipated interactions between model components, the existence of which may result in undesirable system behavior. With respect to inclusive privacy and security, these problematic interactions may actually pose a threat to user privacy, security, and usability, or reveal solution features that seem beneficial to developers but may actually result in disadvantageous outcomes for end users.

Consider once more the use of encrypted communication applications, but now with respect to the mental models of different user groups: casual users; protesters in oppressive regimes; and victims of spousal abuse. Each group may differ on anticipated threats, environments for use, and application features, among others, and the features added to the benefit of one group may inconvenience or endanger another. For example, blocking a user from taking screenshots may prevent an abusive spouse from gathering evidence of covert support, but it may also prevent a protester from saving protest coordination information if cellular network access is cut. Well-intentioned developers may not have knowledge of the risks that domestic violence victims face, and so may inadvertently design out an important user group. However, exploring this problem space with formal methods is uniquely suited for discovering these and other unantici-

pated interactions. This technique could be a powerful tool in the design of broadly-inclusive privacy and security solutions.

4. CONCLUSIONS

Efforts to improve the privacy and security of disadvantaged populations must answer to the needs and challenges of these groups. One promising strategy may be the use of formal methods to explore the mental models of these groups and discover unanticipated, potentially dangerous human-system interactions. Drawing on existing work from a variety of fields, this approach can reveal insights that developers may miss and could suggest realistic improvements in solution design.

Limitations to this approach do exist, particularly with respect to the successful elicitation of user mental models and appropriately scoping formal verification model analysis. Future work will focus on developing this framework and further refining these formal analytic techniques.

5. REFERENCES

[1] F. Asgharpour, D. Liu, and L. J. Camp. Mental models of computer security risks. In *Workshop on the Economics of Information Security (WEIS)*, 2007.

[2] C. Baier, J.-P. Katoen, and K. G. Larsen. *Principles of model checking*. MIT press, 2008.

[3] J. Blythe and L. J. Camp. Implementing mental models. In *IEEE Symposium on Security and Privacy Workshops (SPW)*, 2012.

[4] M. L. Bolton, E. J. Bass, and R. I. Siminiceanu. Using formal verification to evaluate human–automation interaction: A review. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 43(3):488–503, 2013.

[5] J. Bredereke and A. Lankenau. A rigorous view of mode confusion. In *International Conference on Computer Safety, Reliability, and Security*, 2002.

[6] L. J. Camp. Mental models of privacy and security. *IEEE Technology and Society Magazine*, 28(3):37–46, 2009.

[7] M. T. Chi, P. J. Feltovich, and R. Glaser. Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2):121–152, 1981.

[8] S. Combéfis, D. Giannakopoulou, and C. Pecheur. State event models for the formal analysis of human–machine interactions. In *Formal Verification and Modeling in Human–Machine Systems: Papers from the AAAI Spring Symposium*, 2014.

[9] A. Degani. *Taming HAL: Designing interfaces beyond 2001*. Springer, 2004.

[10] A. Degani and M. Heymann. Formal verification of human–automation interaction. *Human Factors*, 44(1):28–43, 2002.

[11] J. M. Dutton and W. Starbuck. Finding Charlie's run–time estimator. In *Computer Simulation of Human Behavior*, pages 218–242. Wiley, New York, 1971.

[12] D. N. Ford and J. Sterman. Expert knowledge elicitation to improve mental and formal models. *System Dynamics Review*, 14(4):309–340, 1997.

[13] R. Kang, L. Dabbish, N. Fruchter, and S. Kiesler. "My data just goes everywhere:" User mental models of the internet and implications for privacy and security. In *Proceedings of the Symposium on Usable Privacy and Security (SOUPS)*, 2015.

[14] M. Kauer, S. Günther, D. Storck, and M. Volkamer. A comparison of American and German folk models of home computer security. In *International Conference on Human Aspects of Information Security, Privacy, and Trust*, 2013.

[15] N. Kobie. *How to keep messages secure*, March 2017. `http://www.teenvogue.com/story/how-to-keep-messages-secure`.

[16] J. Murdock. *Hackers with links to Iranian government attempt to 'map' 15 million Telegram users*, August 2016. `http://www.ibtimes.co.uk/hackers-links-iranian-government-attempt-map-15-million-telegram-users-157393`.

[17] D. A. Norman. *Some observations on mental models*. Erlbaum: Hillsdale, NJ, 1983.

[18] M. Oishi, I. Mitchell, A. Bayen, C. Tomlin, and A. Degani. Hybrid verification of an interface for an automatic landing. In *Proceedings of the 41st IEEE Conference on Decision and Control*, 2002.

[19] J. Rasmussen and W. B. Rouse. Human detection and diagnosis of system failures. In *NATO Conference Series on Human Factors, Series 3*, volume 15. Plenum Press, 1980.

[20] W. B. Rouse and N. M. Morris. On looking into the black box: Prospects and limits in the search for mental models. Technical report, Georgia Institute of Technology, May 1985.

[21] J. Rushby. Modeling the human in human factors. In *International Conference on Computer Safety, Reliability, and Security*, 2001.

[22] J. Rushby. Using model checking to help discover mode confusions and other automation surprises. *Reliability Engineering & System Safety*, 75(2):167–177, 2002.

[23] J. Rushby, J. Crow, and E. Palmer. An automated method to detect potential mode confusions. In *Proceedings of the 18th Digital Avionics Systems Conference*, 1999.

[24] P. M. Sanderson. Knowledge acquisition and fault diagnosis: Experiments with PLAULT. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(1):225–242, 1990.

[25] B. Sharafedin. *Barred from streets, Iran's reformists push for votes online*, February 2016. `http://www.reuters.com/article/us-iran-election-campaign-idUSKCN0VW1Z3`.

[26] R. Wash. Folk models of home computer security. In *Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS)*, 2010.

[27] M. D. Williams, J. D. Hollan, and A. L. Stevens. *Human reasoning about a simple physical system*, pages 131–154. Hillsdale, NJ: Lawrence Erlbaum, 1983.

[28] J. R. Wilson and A. Rutherford. Mental models: Theory and application in human factors. *Human Factors*, 31(6):617–634, 1989.

[29] J. M. Wing. A specifier's introduction to formal methods. *IEEE Computer*, 23(9):8–22, 1990.