Exploring Expert and Novice Mental Models of Phishing

Olga A. Zielinska, Allaire K. Welk, Christopher B. Mayhorn, Emerson Murphy-Hill
North Carolina State University

Experience influences actions people take in protecting themselves against phishing. One way to measure experience is through mental models. Mental models are internal representations of a concept or system that develop with experience. By rating pairs of concepts on the strength of their relationship, networks can be created through Pathfinder, showing an in-depth analysis of how information is organized. Researchers had novice and expert computer users rate three sets of terms related to phishing. The terms were divided into three categories: prevention of phishing, trends and characteristics of phishing attacks, and the consequences of phishing. Expert mental models were more complex with more links between concepts. Specifically, experts had sixteen, thirteen, and fifteen links in the networks describing the prevention, trends, and consequences of phishing, respectively; however, novices only had eleven, nine, and nine links in the networks describing prevention, trends, and consequences of phishing, respectively. These preliminary results provide quantifiable network displays of mental models of novices and experts that cannot be seen through interviews. This information could provide a basis for future research on how mental models could be used to determine phishing vulnerability and the effectiveness of phishing training.

INTRODUCTION

A survey done by the Census Bureau in 2013 indicated that 84% of U.S. households owned a computer or laptop and 74% of households had internet access. This leaves many computers in the United States vulnerable to phishing attacks 0. Phishing is defined as a social engineering tactic that is used to trick people into revealing personal information 0. Information could include dates of birth, credit card numbers, and Social Security numbers. There are multiple resources households could use to protect themselves; however, the number of phishing attacks continues to grow. A 2013 report from Kasperksy Lab indicated that there were 37.3 million phishing attacks in 2013, up from 19.9 million attacks in 2012 0, which raises the question: what is causing users to continually fall victim to phishing?

One approach to determining this is to analyze the mental models of a computer user on the topic of phishing. Mental models are internal representations a user has of a concept or system 0. In a mental models study of phishing, Wash found that people form mental models about phishing from stories they have heard from friends or colleagues. The knowledge and perception they have about phishing has been shown to affect the actions they take to protecting themselves. For example, people that believe that hackers are “mischievous teenagers showing off to their friends” have put up firewalls to protect themselves; while people who believed that hackers were criminals often believed that they were not rich or important enough to be targeted and in turn did not need to secure their computers 0.

Mental models grow with interaction with a system or concept and eventually the user will be able to use his or her developed mental models to predict or explain the system or concept 0; 0. Accordingly, as users develop expertise, they have qualitative changes in their mental models. Experts are able to quickly analyze a situation or case and make quick decisions because of their coherent organization of information 0; 00. Mental models develop with feedback given by the system; however, in phishing the feedback is usually not immediate and it may take months before a user realizes they have been phished, and may not know where or when they became vulnerable to the attack 0; 0; 0.

Mental models are difficult to measure; however, several researchers have used think-aloud protocols, interviews, or multivariate statistical techniques in an attempt to describe mental models 0. One way of utilizing multivariate statistical techniques to observe mental models begins by asking participants to rate a pair of concepts. Participants should be instructed to perform these ratings based on the amount of similarity between two concepts or strength of the relationship between the concepts 0. These ratings are rendered into a graphical network representation of the relationship with the concepts represented as nodes, and links that show if a relationship exists between concepts. Pairs that are highly similar or strongly related are shown with a direct link between the two concepts 0; 0.

The networks are mental model representations that can be quantified and compared 0. Networks are often compared with a referent structure or a network completed by subject matter experts 0 0. One study compared experts’ and novices’ mental models of general computing concepts, Window-based applications, and word processing programs 0. Another study has examined the similarity of students’ and instructor’s mental models at the beginning and end of a semester, showing that the mental models of students are more similar to the instructor’s at the end of semester, especially if they have performed well in the class 0. However, up until this point no studies have examined the networks of phishing mental models between computer security novices and computer security experts.

In this exploratory study, researchers aim to analyze the networks of computer security novices and computer security experts in the context of phishing. Researchers hypothesize that the networks of novices and experts will be significantly different. Additionally, researchers hypothesize that expert network models will have more links between concepts than

novices, as expert mental models are more developed and have been shown to have qualitative differences when compared to novices 000.

METHOD

Participants

Thirty-five participants (20 novices and 15 experts) were recruited. Novice participants were undergraduate students enrolled in an introductory psychology class and received research credit for class. There were 12 male and 8 female novice participants with a mean age of 18.75 (*SD*=1.02).

Fifteen expert participants were recruited through convenience and snowball sampling. Two expert participants were recruited through an Industry Day Event, where companies and university researchers come together to collaborate on projects. In turn the two experts, recommended fourteen additional participants that would be interested in participating in the study. There were 13 male and 2 female expert participants with a mean age of 47.07 (*SD* = 6.45). Each expert worked in the IT department of their company and had experience with attempted phishing attacks at their company.

Materials

For the current study, three sets of terms were generated, each relating to phishing. Term selection was done carefully through a review of phishing textbooks, phishing journal articles, and pilot data. The strength and type of relationship between concepts is dependent on the context therefore, context was provided for each set of terms 0;0; 0;0; 0; 0. Specifically, the first set examined terms in relation to the prevention of phishing attacks, the second set examined terms in relation to the trends or characteristics seen in a phishing attack, and the last set of terms examined the context of the consequences of phishing. Terms are listed in Table 1.

A survey collection application, Qualtrics, recorded relatedness ratings. Participants rated the strength of relationships among pairs of concepts on an eleven-point scale. Instructions were given to rate the strength of the relationship between the two terms on a scale of 0 to 10. A rating of 0 would indicate that there is no relationship between the two terms while a rating of 10 would indicate that the terms are strongly related. Participants were instructed to rate the terms in the context listed at the top. The context was one of the three categories: prevention of phishing attacks, characteristics of phishing attacks, and consequences of phishing attacks.

Table 1: List of terms used for relatedness ratings.

Prevention	Trends/Characteristics	Consequences
Updates	Unknown Sender	Financial
Anti-Malware	Known Sender	Proprietary
Training	Personalized Content	Emotional
Red Team	Too Good to Be True	Passwords
Warnings	Bad Spelling/Grammar	Police Involvement
Passwords	Quick Response	Social Credibility
Software	Link	Credit Score
Authentication	Attachment	Loss of Customers
Encryption	Legitimate Appearance	Suicide
Black List	Social Engineering	Virus

Procedure

We instructed participants to complete the survey in a quiet office setting through a computer or laptop with Internet connection. After reading through and agreeing to the consent form, participants provided demographic information. The survey asked participants to complete three sets of relatedness ratings, followed by a few interview questions, and a debriefing form. The three sets included ten terms in the context of the prevention of phishing, the trends/characteristics of phishing, and the consequences of phishing. Each participant completed all three sets of relatedness ratings, one set at a time. The survey randomized the presentation order of the sets of terms to minimize ordering effects.

Data Analysis

Relatedness ratings for each set of terms were input into Pathfinder, a statistical software package that represents pairwise proximities in a network 0; 0. The network was displayed with nodes, represented by the concepts, and links, which showed the relationships between the concepts. This network representation summarized the data and has also been shown to convey information about the relationships that is not seen in the ratings themselves 0;0; 0. Results from the experts were aggregated to form one group. Previous studies have shown that judgments from several subject matter experts are frequently combined to derive a true score and has the advantage of overcoming personal biases 0. A similar aggregation was done with the novice ratings to derive a true score for novices and facilitate comparison between the two groups.

RESULTS

A graphical representation of the network is seen below in Figure 1. Trends and characteristics of phishing are in the left column and the consequences of phishing are in the right column. Novices are displayed in the top row, followed by experts in the second row, and an overlap of experts and novices are in the last row. Novices had much simpler networks with a maximum of three links for a concept, while experts had much more complex networks, with concepts containing up to seven links. The graphical network of prevention terms for novices and networks can be found in a previously published article (Zielinska et al., 2015). Alternate analyses containing networks with concepts in the same location for both novice and experts was conducted. Due to space limitations those graphs were not included, but they are available upon request from the first author.

Table 2, below, describes the networks quantitatively. It lists the number of links in the novice network, the number of links in the expert network, the number of links the novice and expert networks have in common (com), the number of links

in common when correcting for chance (ccom), the similarity (sim) of the networks, the similarity of the networks when corrected for chance (csim), and the probability of obtaining this many or more common links (tprob). Similarity is rated on a scale of 0 to 1. Two networks that show no links of similarity will be represented by a 0, while two identical networks will have a similarity of 1 0.

For example, when examining the prevention terminology, novices had 11 links in the network between the 10 concepts, while experts had 16 links between the concepts. There were 6 links between concepts that novices and experts had in common; however, when correcting for chance, the common links dropped to 2.09. A similarity rating of 0.29 was given to the comparison of novice and expert networks. This is closer to 0 than to 1 indicating that the networks are not very similar. When correcting for chance, the similarity dropped to .11, showing they are even less similar.

Table 2: Number of links per network for novices and experts, number of common links for both networks, and similarity of networks for each set of terms.

	Prevention	Trends	Consequences
Links-Novice	11	9	9
Links-Experts	16	13	15
Com	6	5	6
CCom	2.09	2.4	3
Sim	0.29	0.29	0.33
CSim	0.11	0.16	0.19

DISCUSSION

Results from this study indicate novices and experts have significantly different ways of organizing and conceptualizing information about phishing. Novices have much simpler mental models on how to prevent phishing, the trends and characteristics, and the consequences associated with a phishing attack. This research supports previous findings that mental models develop with experience 0.

Novices and experts had the most links in common when correcting for chance (ccom) and the highest similarity of the three sets of terms. This could be due to the fact that both groups could be reading the same articles and news stories about what the consequences of a phishing attack are and therefore creating similar mental models. Phishing prevention and characteristics are not glamorized as much in the media so novices may not have the same exposure to that information as the experts who work with this on a daily basis.

Additionally, the trends illustrated in mental model networks could have implications for training. For example, the aggregate expert model illustrated “unknown sender” as a central node connected to “social engineering”, “link”, “attachment”, and “bad spelling/grammar”, whereas novices only linked “unknown senders” to “attachment” and “link”. This illustrates that experts likely have a more comprehensive understanding of how unknown senders can relate to a broad array of phishing trends and characteristics. Training programs might aim to replicate this expert model in novices by providing information regarding the interconnectedness of these trends and characteristics related to unknown senders.

Limitations

There are a few limitations to consider with this study. First, the novice group contained strictly undergraduate students. This may not represent the 84% of people in the United States who own a computer or laptop. Undergraduate students may have more computer and Internet experience than populations older than them. Additionally, universities regularly send out emails warning users of possible phishing attacks, exposing them to the topic. A home user may not necessarily have someone educating them about the phishing topic, and therefore the average computer user mental models may be even simpler than the ones found in the study.

Additionally, Pathfinder gives novices credit each time they link the same two concepts as an expert (common links); however, this does not necessarily mean that they connect the same two concepts in the same way 0. For example, social engineering and legitimate appearance in the trends and characteristics group had a link connecting them for both novices and experts. Novices, some of which believed that social engineering is talking to people through social media, may believe that social engineering has a legitimate appearance because it is run by Facebook or Twitter. Experts may believe social engineering has a legitimate appearance because they are attempting to trick the user into believing they are part of a real company so users give up information. The expert may believe that social engineering has little or nothing to do with social media, but the link remains even though the relationship is different. Although Pathfinder provides a visual representation of the connections, it is necessary to find out more information about the relationship to see if they have the same meaning between novices and experts.

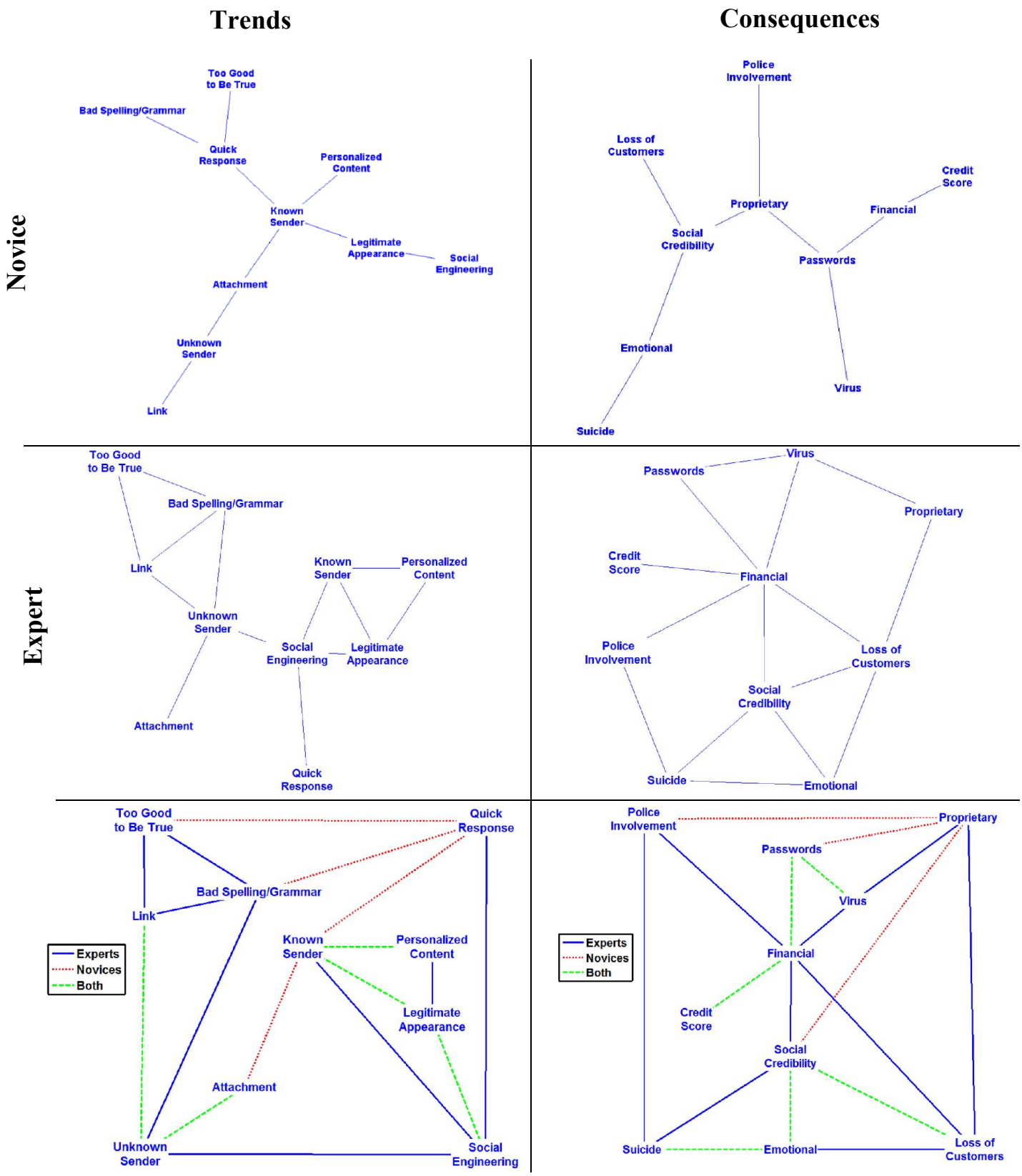

Figure 1 : Graphical networks representing the terms rated related to the concept of phishing. Trends are on the left. Consequences are on the right. The top row contains the networks of the novices; the middle row contains the network of the experts; and the last row contains the overlap of the novice and expert networks. The prevention term networks can be found in Zielinska et al. (2015).

Finally, participants completed 45 relatedness ratings per set of terms and a total of 135 relatedness ratings. Participants may have found these ratings to be tedious, as found in previous studies 0; 0. Although the presentation order of each set was randomized to minimize ordering effects, participants could have been fatigued when completing their second or third set of ratings.

Future Research

Mental models created through Pathfinder have been used to determine levels of expertise and performance 0; 00; 0. This could be applied to phishing knowledge to determine phishing vulnerability. By pairing mental models with a phishing vulnerability assessment, we could predict which model is the most vulnerable. We could also assess primary differences among models, such as which links are present in people that are less vulnerable that are missing in the more vulnerable people. This could tailor training programs to users to make them more effective.

Additionally, mental models have been used to assess learning and training 0; 0. One study compared a teacher's mental model to students' mental models at the beginning of a course and at the end of the course in relation to course performance. At the beginning of the course, students had different mental models from their instructor, but the mental models evolved over the duration of the course and those who performed well in the course had mental models most closely aligned with the teacher's mental model 00. Mental models could be assessed before phishing training to determine a user's initial mental model and then after training to determine if the mental model has changed with the training, and if it is making the person less vulnerable.

ACKNOWLEDGEMENTS

The authors would like to thank the Science of Security Lablet for funding this project. The authors would also like to thank Rebecca McNulty for her assistance in this project.

REFERENCES

Cooke, N.J., (1992) Predicting judgment time from measures of psychological proximity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 18. 640-653.

Cooke, N.J., & McDonald, J.E., (1987) The application of psychological scaling techniques to knowledge-based systems. *International Journal of Man-Machine Studies*. 26, 533-550.

Cooke, N.J., & Schvaneveldt, R.W. (1998) Effects of computer programming experience on network representations of abstract programming concepts. *International Journal of Man-Machine Studies*. 29, 407-427.

Cooke, N.M., Durso, F.T., & Schvaneveldt, R.W. (1986). Recall and measures of memory organization. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 12. 538-549.

Day, E. A., Arthur Jr, W., & Gettman, D. (2001). Knowledge structures and the acquisition of a complex skill. *Journal of Applied Psychology*, 86(5), 1022.

Dorsey, D.W., Campbell, G.E., Foster L.F., Miles, D.E. (1999) Assessing Knowledge Structures: Relations With Experience and Post training Performance, *Human Performance*, 12:1, 31-57.

File, T. and Camille R., (2014) "Computer and Internet Use in the United States: 2013," American Community Survey Reports, ACS-28, U.S. Census Bureau, Washington, DC, 2014.

Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. Journal of educational psychology, 83(1), 88.

Hong, E., & O'Neil, H.F. (1992) Instructional strategies to help learners build relevant mental models in inferential statistics. *Journal of Educational Psychology.* 84(2), 150-159.

Interlink. (1992). Pcknot [Computer software documentation]. Las Cruces, NM: Interlink, Inc.

Kaspersky Lab Report. (2013) 37.3 Million Users Experienced Phishing Attacks in the Last Year. http://www.kaspersky.com/about/news/press/2013/Kaspersky_Lab_report_37_3_million_users_experienced_phishing_attacks_in_the_last_year

King, D.L. (1988). Similar as well as dissimilar contextual stimuli increase rated similarity. *Memory & Cognition.* 16(1), 71-78.

Murphy, G.L., & Medin, D.L., (1985) The role of theories in conceptual coherence. *Psychological Review*. 92, 289-316.

Norman, D.A. (1983) Some observation on mental models. In D.Genter & A.L. Stevens (Eds.) *Mental Models*. (pp.7-14). Hillsdale, NJ: Erlbaum.

Rasmussen, J., & Jensen, A., (1974) Mental procedures in real-life tasks: A case study of electronic troubleshooting. *Ergonomics*. 17, 293-307.

Rowe, A.L., Cooke, N.J., Hall, E.P., & Halgren, T.L., (1996) Toward an online knowledge assessment methodology: Building on the relationship between knowing and doing. *Journal of Experimental Psychology: Applied*. 2. 3-47.

Rowe, A. L., & Cooke, N. J. (1995). Measuring mental models: Choosing the right tools for the job. Human resource development quarterly, 6(3), 243-255.

Rowe, A.L. (1994). Mental models of physical systems: Examining the relationship between knowing and doing. Unpublished doctoral dissertation, Rice University, Houston, TX.

Schvaneveldt, R. (1990) *Pathfinder associative networks: Studies in knowledge organization.* Norwood, NJ: Ablex.

Shavelson, R.J., (1974) Methods for examining representations of a subject-matter structure in a student's memory. *Journal of Research in Science Teaching.* 11, 231-249.

Tversky, A. (1977) Features of similarity. *Psychological Review,* 84, 327-352.

Wash, R. (2010) Folk models of home computer security. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, ACM, pp. 1-16.

Wash, R., & Rader, E. (2011) Influencing mental models of security: a research agenda. In *Proceedings of the 2011 workshop on new security paradigms workshop* (pp. 57-66). ACM.

Zielinska, O. A., Tembe, R., Hong K.W., Ge X., Murphy-Hill E., and Mayhorn C.B.. (2014) One Phish, Two Phish, How to Avoid the Internet Phish Analysis of Training Strategies to Detect Phishing Emails. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Chicago, Illinois, SAGE Publications, 1466-1470.

Zielinska, O., Welk, A., Mayhorn, C. B., & Murphy-Hill, E. (2015) Exploring expert and novice mental models of phishing. In *Proceedings of the 2015 Symposium and Bootcamp on the Science of Security* (p. 22). ACM.