Poster: Computer security information in stories, news articles, and education documents

Katie Hoban, Emilee Rader, Rick Wash, Kami Vaniea
Department of Media and Information
Michigan State University
{hobankat, emilee, wash, vaniea}@msu.edu

ABSTRACT

Despite the large amount of computer security information available to them, end users are often thought of as the weakest link in computer security. The information they have access to comes in many formats: news articles, news broadcasts, education documents, books, stories, and many more. However, inefficient or inconsistent communication between content providers and content consumers may result in knowledge gaps for the consumers. The quality and attributes of the computer security information available to users impact their ability to learn from them. To better understand this state of affairs, we collected news articles, education documents, and a survey about stories end-users had heard about computer security. We then analyzed the trends present across all three datasets. We found that there are serious mismatches between these datasets concerning the topics of hackers, viruses + malware, and phishing + spam; discrepancies that may impact the communication of computer security information to end-users.

1. INTRODUCTION

End users are often considered the weakest link in computer security systems. Reasons for this include memory limitations [1], the cost of learning, among others. However, end users cannot be completely removed from the loop of computer security [2]. They are still expected to make security decisions, and to do so, they rely on outside information: namely, the computer security information intended to educate end-users. One of the ways we can better understand computer security education is by examining the materials end-users consume. To this end, we collected 1062 computer security news articles, 518 computer security education documents, and 301 stories that people had heard about computer security.

Our work is important because of its focus not only on end-users, but also the information they consume and relate regarding computer security. The trends across the multiple studies we have done so far provide a holistic context for behavioral information security. We examined trends across three distinct sources of computer security information, all of which impact the security beliefs and practices of end users. By examining these stories, news articles, and education documents, our study provides an in-depth look at both what end users are consuming, as well as what they relate to their peers.

2. METHODOLOGY

To understand the variety of information that end users receive about computer security, we collected data from three sources of security information: stories people tell each other, newspaper articles about security, and education documents produced by security experts, intended for use by non-experts.

Story survey: We distributed a survey in 2011 to 301 undergraduate students. The survey asked the students to relate a story they had heard about computer security, the moral they took away from that story, and whether or not they changed their behavior in reaction to that story. These stories were then coded into 6 topics they addressed [3].

Newspaper articles: In 2012, we collected 1062 news articles about computer security that were published in 2011. These articles came from 16 regional, national, and international newspapers. We manually coded them for 23 topics.

Education documents: In 2012 and 2013, we collected 518 computer security education documents from universities, companies, and government institutions and coded them for format and education tactic. We then ran topic modeling on the text of these documents using MALLET to determine 12 distinct topics [5].

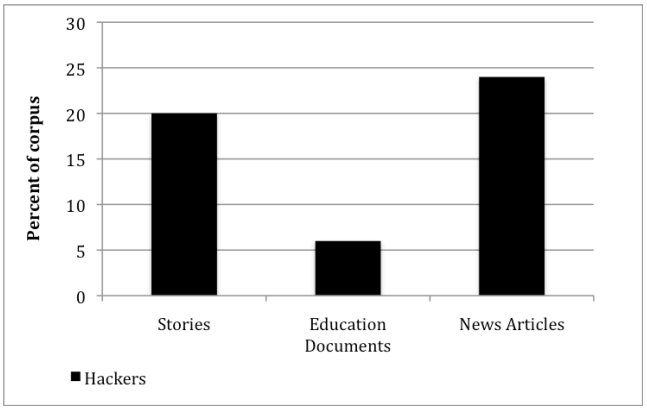
3. RESULTS

Since the number of entries in each study varies highly, (stories=301, documents=518, and articles=1062) the findings in this section are by the percentage of documents within each corpus coded as each topic, not by absolute count.

In order to better view trends across all three of these datasets, we grouped each set of topics into nine meta-topics: Cyber Harassment, Hackers, Identity Theft, Organizational Security, Personal Security, Phishing + Spam, Viruses + Malware, and Other. We focus on the findings in three of these meta-topics: hackers, viruses + malware, and phishing + spam.

3.1 Hackers

Figure 1. Percent coverage of Hackers meta-topic



As can be seen in Figure 1, hackers received substantial attention in the stories and news articles (20% and 24% coverage, respectively) but did not receive much attention from the

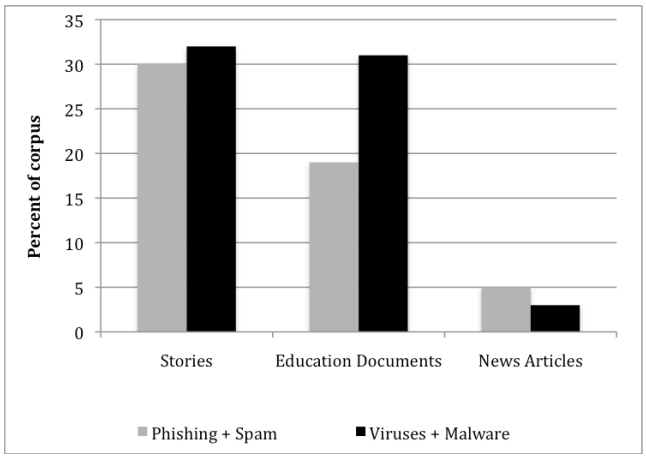
education documents (6%). Our story survey asked respondents to indicate whether or not they had changed their behavior in reaction to the computer security story they had heard, and those who had heard stories about hackers were least likely to change their behavior [3]. The difference in prevalence of the hacker meta-topic over these datasets may play into the markedly small number of people who reported changing their behavior after hearing a story about hackers (32%). The average behavior-change rate for the people who related stories in other topics was 52%.

Many of the hacker education documents were explanatory and did not provide any security advice, preventative or reactionary. The code that determined this, "Just the Facts", was the second most common code in the hacker documents. This suggests that the information people were given about hackers and hacking from education documents only defined the problem for them and did not offer any actions they could take to prevent an attack, or how they should react to an attack that had already occurred. Participants in the story survey were concerned about hacking, but the education materials did not provide them with specific advice.

The discrepancy in popularity of hacker-related materials among the datasets could be explained by the disproportionate amount of attention given to hacking events by the stories and news articles. Part of this sample was from the time Sony/Nintendo was hacked into in 2011, and a subset of the stories and the news articles also addressed this occurrence. Events about hacking in the news tend to happen to either high-profile organizations, or organizations where a security breach would impact a large number of people. Since security breaches at high-profile companies impact so many people, this might explain the disproportionate prevalence of the Hackers meta-topic in the stories and news articles.

3.2 Viruses + malware and phishing + spam

Figure 2. Percent coverage of Viruses + Malware and Phishing + Spam meta-topics



As can be seen in Figure 2, there are also notable differences across datasets regarding the Viruses + Malware and Phishing + Spam meta-topics. Both of these meta-topics received substantial attention in the stories and education documents, but not in the news articles. Viruses + Malware covered 32% of stories and 31% of education documents, but only 3% of news articles. Similarly, Phishing +Spam covered 24% of stories and 45% of documents, but only 5% of articles.

We hypothesize that these two meta-topics have become mundane in the eyes of the mainstream media, and that newspapers subsequently do not publish many articles on them. Another possibility is that, since the majority of these security breaches happen to average people, rather than high-profile organizations, the news may believe that they are not worth media attention.

4. DISCUSSION

Our finding that hackers are not often addressed in education documents is important because the education documents are not addressing a major point of end-user concern, whether this is because of the disproportionate amount of attention given to hacking in the media, the widespread effects of hacking in the video game community, or the lack of focus of education materials on hacking and hackers. End-users may be concerned about hacking because they hear about it often in the news and from other people, but few education documents address this meta-topic.

Our viruses and phishing findings are important because there is a mismatch between public concern and news coverage of Viruses + Malware and Phishing + Spam in the news. People are still concerned about these security breaches, and education documents cover these meta-topics to a commensurate degree, but these occurrences have become too mundane for newspapers to cover them to any great extent. This discrepancy results in a lack of media visibility for a meta-topic that concerns end-users.

5. ACKNOWLEDGMENTS

Thank you to everyone at the BITLab for your support and input throughout this project, particularly Lauren McKown for collecting and coding the news articles, Nate Zemanek for collecting the education materials, and Zack Girouard for helping with coding the education documents.

6. REFERENCES

[1] Sasse, M., Brostroff, S. and Weirich, D. Transforming the 'weakest link' – a human/computer interaction approach to usable and effective security.

[2] Cranor, L. 2008 A Framework for Reasoning About the Human in the Loop. In the UPSEC'08 Proceedings of the 1st Conference on Usability, Psychology, and Security.

[3] Rader, E., Wash, R., and Brooks, B. 2012 Stories as Informal Lessons about Security. In the Proceedings of the Symposium on Usable Privacy and Security.

[4] McCallum, A. 2002 MALLET: A Machine Learning for Language Toolkit. http://mallet.cs.umass.edu