SPRING 2016 SEMINAR SERIES

You Get Where You're Looking For: The Impact of Information Sources on Code Security

Doowon Kim
Department of Computer Science, University of Maryland, College Park

April 14, 2016 (Thursday) @ 3:30 – 4:45 PM, CSB - Room 210

Doowon Kim is currently a Ph.D. student at UMD CS, advised by Prof. Michelle Mazurek. His research interests are in computer security, privacy, and usable security.

Abstract: Vulnerabilities in Android code - including but not limited to insecure data storage, unprotected inter-component communication, broken TLS implementations, and violations of least privilege – have enabled real-world privacy leaks and motivated research cataloguing their prevalence and impact. Researchers have speculated that application promotes security problems, as it increasingly allows inexperienced laymen to develop complex and sensitive apps. Anecdotally, Internet resources such as Stack Overflow are blamed for promoting insecure solutions that are naively copy-pasted by inexperienced developers. In this paper, we for the first time systematically analyzed how the use of information resources impacts code security. We first surveyed 295 app developers who have published in the Google Play market concerning how they use resources to solve security-related problems. Based on the survey results, we conducted a lab study with 54 Android developers (students and professionals), in which participants wrote security- and privacy relevant code under time constraints. The participants were assigned to one of four conditions: free choice of resources, Stack Overflow only, official Android documentation only, or books only. Those participants who were allowed to use only Stack Overflow produced significantly less secure code than those using, the official Android documentation or books, while participants using the official Android documentation produced significantly less functional code than those using Stack Overflow. To assess the quality of Stack Overflow as a resource, we surveyed the 139 threads our participants accessed during the study, finding that only 25% of them were helpful in solving the assigned tasks and only 17% of them contained secure code snippets. In order to obtain ground truth concerning the prevalence of the secure and insecure code our participants wrote in the lab study, we statically analyzed a random sample of 200,000 apps from Google Play, finding that 93.6% of the apps used at least one of the API calls our participants used during our study. We also found that many of the security errors made by our participants also appear in the wild, possibly also originating in the use of Stack Overflow to solve programming problems. Taken together, our results confirm that API documentation is secure but hard to use, while informal documentation such as Stack Overflow is more accessible but often leads to insecurity. Given time constraints and economic pressures, we can expect that Android developers will continue to choose those resources that are easiest to use; therefore, our results firmly establish the need for secure-but-usable documentation.

Contact Dr. Soo-Yeon Ji (sji@bowiestate.edu) if you have any question.