Towards Cognitive Obfuscation Analyzing Human Factors to Impede Hardware Reverse Engineering

Steffen Becker, Carina Wiesen
Horst-Görtz-Institut für IT-Sicherheit, Ruhr-Universität Bochum
(Prof. Dr. Nikol Rummel, Prof. Dr.-Ing. Christof Paar)

In a world in which interconnected digital systems permeate almost all facets of our lives, cyber security attacks form devastating threats with catastrophic consequences. Hardware components are the root of trust in virtually any computing system and are valuable targets of cyberattacks. In order to detect fabrication faults, copyright infringements, counterfeit products, or malicious manipulations hardware reverse engineering is usually the tool-of-choice. While hardware reverse engineering is a highly complex and universal tool for legitimate purposes, it can also be employed with illegitimate intentions, undermining the integrity of ICs via piracy, subsequent weakening of security functions, or insertion of hardware Trojans. In particular, Intellectual Property (IP) piracy has become a major concern for the semiconductor industry which causes losses in the range of several billion dollars. Governments and armed forces classify hardware reverse engineering as a major security threat, since malicious manipulations of hardware components in mission-critical systems can even have lethal consequences. Due to the serious threats posed by attacks based on hardware reverse engineering, strong countermeasures are indispensable. The security of most existing obfuscation techniques is assessed exclusively based on technical measures. However, the process of hardware reverse engineering cannot be fully automated yet and the lack of holistic tools forces human analysts to combine several semi-automated steps. Accordingly, cognitive processes and strategies applied by humans in the context of hardware reverse engineering must be considered for the development of sound countermeasures.

Our research aims to develop cognitive obfuscation methods considering both technical and human factors during the hardware reverse engineering processes. Based on the psychological theory of reverse engineering by Lee & Johnson-Laird (2013) we define human processes in hardware reverse engineering as a special kind of problem solving. Our research is focused on understanding how human analysts reverse parts of unknown hardware designs in realistic scenarios. Therefore, we perform several psychological studies and analyze the behavior of engineers at different levels of expertise. By observing human engineers who are analyzing gate-level netlists of unknown chip designs we aim to identify problems in the process of hardware reverse engineering which are hard to solve. Based on these problems and failures we aim to derive novel quantification metrics enabling obfuscation metrics, that are hard to solve for both machines and humans. In our talk, we will present the psychological background, study design and first results in more detail.