Research work:
Modeling Control-Flow Attackers of a Secure Cryptoprocessor

Supervisors: Sylvain Boulmé and Marie-Laure Potet
mailto:Sylvain.Boulme@univ-grenoble-alpes.fr

Our pervasive computing world (with Internet of things, Cyber-physical systems, etc) makes us very vulnerable to computer attacks. In particular, control-flow attacks makes the chip run another code than the one intended by its designer (for instance, see Return-oriented programming as detailed in Buchanan et al. [2008]). Control-flow attacks may exploit software weaknesses (e.g. a lack of protection against buffer overflows), or they may also “create” the weakness by hardware fault injections Bukasa et al. [2018].

Several techniques, under the umbrella of Control-Flow Integrity (CFI) defined by Abadi et al. [2005a], have been proposed to detect and prevent such attacks, by monitoring program branching behaviors. In particular, we consider a processor, called “IntrinSec”, and designed by Oliver Savry and his team at CEA-Leti. IntrinSec provides CFI from instruction-level encryption, following ideas of a previous experiment described in Hiscock et al. [2019].

In order to ensure CFI, IntrinSec requires that the compiler generate some special instructions before each branching instructions. At Verimag, Sylvain Boulmé and Paolo Torrini have implemented these modifications on the CompCert compiler\textsuperscript{1}. CompCert is actually a certified compiler, meaning that the functional correctness of the compiler is mathematically proved, and that this proof is automatically checked within a proof assistant (here Coq). While, the functional correctness of our compiler is still a proof in progress, they have already defined a functional model of IntrinSec and of its special instructions to protect the control-flow.

This research work will tackle the issue of evaluating the protection of IntrinSec against control-flow attacks. This requires to define models of control-flow attackers, such as de Clercq and Verbauwhede [2017], Potet et al. [2014] or Jacomme et al. [2017]. Hence, the goal of this work is to define/adapt such models in order to evaluate CFI of IntrinSec. For example, following ideas of Abadi et al. [2005b], can we extend our CompCert model of IntrinSec in order to include models of attackers? Can we prove that these attacks make the processor aborts? Indeed, ideally, we expect to be able to mathematically prove that some kinds of attacks are not possible (or difficult). Such mathematical proofs are now very frequent on security protocols : for instance, see Scerri [2015].

\textsuperscript{1}http://compcert.inria.fr/
References


