DEFENSE AGAINST CACHE BASED MICRO-ARCHITECTURAL SIDE CHANNEL ATTACKS

MS Thesis
Sahil Bhat
April 29, 2019. 10:00 AM
ENGR 3202

Advisor: Dr. Houman Homayoun

Abstract:
The hardware security domain in recent years has seen many state-of-the-art cache based Side channel attacks (SCAs) which have posed and continue to pose threats to the integrity of our data. These attacks function by exploiting the side-channels which invariably leak important data during various operations of its (application) execution. These attacks have been successful to steal the private keys from RSA encryption by observing the sequence of operations. Shutting down the side channels is not a feasible approach due to various restrictions it would pose to system performance, hence it is necessary to reduce the entropy of the side channels to reduce the leakage and thus, thwart such attacks. Moreover, to overcome the performance overheads incurred by the traditional software-based malware detection techniques, Hardware-assisted Malware Detection (HMD) using machine learning (ML) classi ers has emerged as a panacea to detect malicious applications and se-cure the systems. To classify benign and malicious applications, HMD primarily relies on the generated low-level microarchitectural events captured through Hardware Performance Counters (HPCs).
We hereby propose a method to minimize the side channel leakage thus thwarting the attack. A wrapper code adds perturbations to the data leaked by the victim application thereby reducing entropy which makes the data on the attacker's side resemble leaked secret data but with perturbations added which makes it arduous to retrieve the original secret data. The wrapper code 'Entropy Shield' can be implemented to protect any encryption algorithm with only a few tweaks. Also, this work creates an adversarial attack on the HMD systems to tamper the security by introducing the perturbations in the HPC traces with the aid of an adversarial sample generator application. To craft the attack, we rst deploy an adversarial sample predictor to predict the adversarial HPC pattern for a given application to be misclassi ed by the deployed ML classi er in the HMD. Further, as the attacker has no direct access to manipulate the HPCs generated during runtime, based on the output of the adversarial sample predictor, we devise an adversarial sample generator wrapped around a normal application to produce HPC patterns similar to the adversarial predictor HPC trace. As the crafted adversarial sample generator application does not have any malicious operations, it is not detectable with traditional signature-based malware detection solutions. With the proposed attack, malware detection accuracy has been reduced to 18.04% from 82.76%.