Efficient Control-Flow Subgraph Matching for Detecting Hardware Trojans in RTL Models
Luca Piccolboni1,2, Alessandro Menon2, Graziano Pravadel1
1 Columbia University, New York, NY, USA
2 University of Verona, Verona, Italy

Hardware Trojans: An Incoming Threat

- Hardware Trojans are defined as malicious and intentional alterations of an integrated circuit that result in undesired behaviors
- trigger logic: activates the malicious behavior under specific conditions
- payload logic: implements the malicious behavior (affects functionality)

Our Approach: Control-Flow Subgraph Matching

- Design Under Verification (DUV)
  - Verilog/VHDL code
- Hardware Trojan (HT) Library
  - Verilog/VHDL code + configuration files
- HT Report
  - matches with a confidence value

Extraction Algorithm

- Hardware Trojan (HT) Library containing parametrizable HTs
- Extraction Algorithm to obtain a CFG from the DUV and the HTs
- Detection Algorithm to identify and locate the HTs in the DUV

In which situations is this useful?
- verify in-house designs at RTL
- verify third-party RTL modules
- verify the results of CAD tools

Detection Algorithm

- The algorithm extracts also the confidence value if matches with a Trojan CFGs in the DUV

Determing the Confidence

- Each trigger can be parameterized with a configuration file that specifies how to extend the CFG to represent other camouflaged instances of the trigger (by using the extension directives)
- The structural characteristics of each trigger are used during the matching (by using the confidence directives)
- The payloads can be used as another metric to calculate the confidence

Experimental Results

- All the benchmarks are injected with one HT, except Crypto-T000 that has zero HTs and Crypto-T100 that has two HTs.