

Hardware Trojan Horse Detection

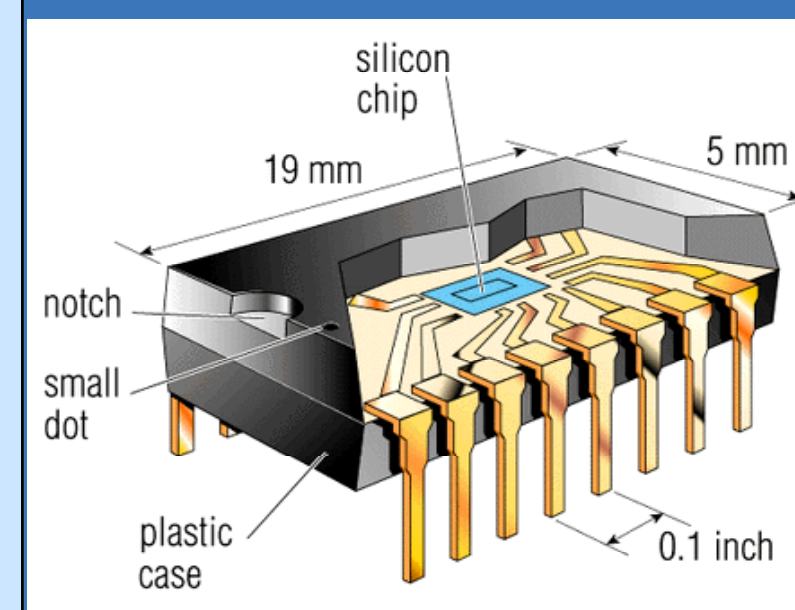
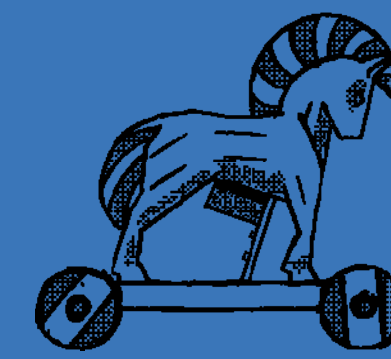


Miodrag Potkonjak (UCLA); Farinaz Koushanfar (Rice U); John Lach (U of Virginia)
www.cs.ucla.edu/~miodrag; www.ece.rice.edu/~fk1; www.ece.virginia.edu/~jcl7d

Trojan horse: unauthorized alternation of design and circuits embedded within the IC and used for malicious purposes

- Untrusted tools, untrusted designers, untrusted foundries, untrusted testing
- Current testing technology does not answer: random vs. intentional errors
- Damages: reveal sensitive information, tampering with the IC, facilitating SW attacks and piracy
- We consider several HTH attacks, including:
 - Gate sizing or addition of extra gates
 - Alteration of interconnects or filling
 - Errors due to IPs and errors introduced by tools

The designed IP is transparent to the fabs, but the fab process, and added circuitry to the Manufactured ICs by the foundry are unknown to the designers and IP providers.



The IC's internals are opaque. Limited controllability and Observability and Unknown actual functionality are the Sources of problems

Approach and Impact

New approach

- Noninvasive detection methods that find the gates and interconnect characteristics on each IC
- To evaluate the algorithm, we study the probability of detection (P_D)
- We also study the probability of false alarm (P_{FA})
- Studying the relationship between P_D and P_{FA} as a function of relevant parameters

Research impact

- Post-silicon characterization of the gates
- Post-silicon characteristics of cross-talk and interconnect
- Security of ICs and protection of security-sensitive devices
- Timing testing of ICs
- Testing for HTH / design against HTH
- Statistical extraction of spatial correlation structure

Technical description: noninvasive gate characterization

Gate-level characterization by non-invasive measurements of timing or power during testing

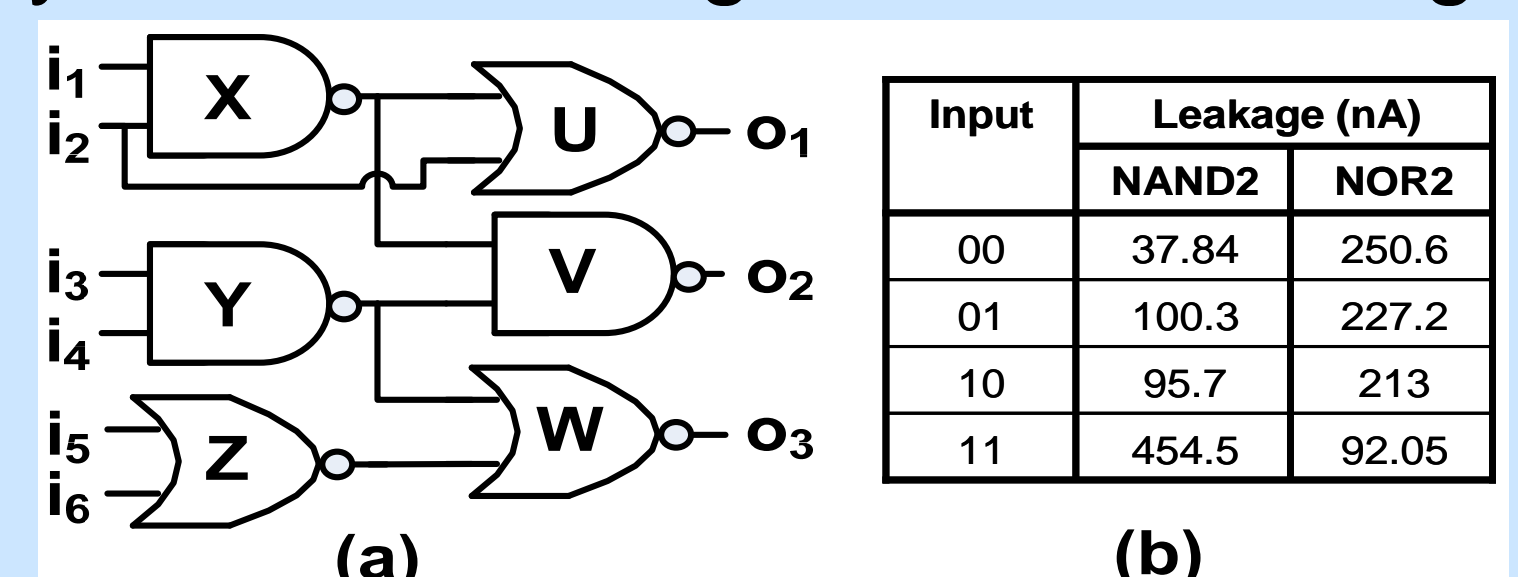
System of linear equations for various input vectors

- *Optimization problem:* minimize sum of errors
- $\sum_n |e_m(t_n)|$, subject to the system of equations
- Unknowns are *scale factors* of the gates (gate sizes)

Challenges in non-invasive post-silicon feature extraction:

- Creating a system of equations with *full-rank*
- Solving the equations in presence of *measurement errors*

Example: find the gates' scale factors by external leakage current reading



$$\begin{aligned}
 I_{leak}(000000) + e_1 &= s_X I_{NAND}(00) + s_Y I_{NAND}(00) \\
 &\quad + s_Z I_{NOR}(00) + s_U I_{NOR}(01) \\
 &\quad + s_V I_{NAND}(11) + s_W I_{NOR}(11); \\
 I_{leak}(010101) + e_2 &= s_X I_{NAND}(01) + s_Y I_{NAND}(01) \\
 &\quad + s_Z I_{NOR}(01) + s_U I_{NOR}(01) \\
 &\quad + s_V I_{NAND}(00) + s_W I_{NOR}(01); \\
 I_{leak}(100101) + e_3 &= s_X I_{NAND}(10) + s_Y I_{NAND}(01) \\
 &\quad + s_Z I_{NOR}(01) + s_U I_{NOR}(00) \\
 &\quad + s_V I_{NAND}(00) + s_W I_{NOR}(01); \\
 I_{leak}(\dots) + e_i &= \dots; \quad i = 4, \dots, M
 \end{aligned}$$