Project Title

Implementation of SAT-attack to break Logic Encryption techniques

Ronak Gupta (21307R010)
Shubham Makar (21307R012)
Tanmaya Shrivastv (21307R016)
MTech. ICS

Department of Electrical Engineering

Indian Institute of Technology Bombay

Abstract

Logic Locking is used as a countermeasure against IP piracy and overproduction where key gates and inputs are added to the existing gate level netlist so that correct functionality is only obtained when the correct key is applied. A promising protection to safeguard hardware Intellectual Properties (IPs) against IP piracy and IP overproduction is Logic Locking. However, numerous attacks have been proposed to compromise the security of logic-locked combinational circuits by retrieving their secret key. An attacker can also exploit the test interfaces of the chip to extract the secret key of the logic-locked sequential circuits. To combat such threats, scan access protection techniques have been developed. Functional logic locking techniques, when used in conjunction with scan access protection methods, provide promising solutions known as scan-based logic locking schemes. SAT attack was very effective against the encryption schemes then. It consists of both high corruptibility and low corruptibility blocks. However, this report demonstrates an attack scenario where a malicious circuitry inserted by an untrusted foundry can effectively bypass the scan locking defense strategies. This in turn, allows the SAT attack to defeat functional logic locking easily.

Threats in VLSI Design Flow

Persistent demand of improving the performance of IC has led to an increase in its complexity. Designers are faced with the challenge of not only improving the performance but also maintaining the same cost or decreasing it. In order to facilitate that, the design and manufacturing of a chip is spread across multiple entities. The designers are adopting the reuse approach to reduce time and cost. They purchase IP blocks to streamline the process and reduce time to market. Also, the manufacturing of chips is done by a separate entity because of the huge cost associated with it. It has expenses like the overall cost of fabrication, wafer sorting, dicing, packaging, package test, and access to state-of-the-art technologies. Hence, offshore foundries are responsible for the manufacturing part of the chip. This has propelled the use of IC supply chain's horizontal model. In the horizontal model of the IC supply chain, different steps like design, fabrication, testing, packaging, and integration of ICs are handled by different players. A distributed supply chain is created as a result of this. Due to outsourcing and the participation of several parties at different stages of the IC supply chain, the cost and time-to-market of the chip have been significantly reduced. There are a lot of advantages associated with the distributed IC design. However, in a globalized IC supply chain, a number of potentially unreliable parties may have access to the physical IC or soft IP. Original designers and IP owners/vendors will have now much less control over the supply chain. It results in various hardware security threats, including but not limited to IP piracy, IC overproduction, counterfeiting, insertion of hardware trojans in an IC, and reverse engineering of the chip. The original IP vendors incur losses when an untrusted entities overproduces chips and sells at a lower prices because they (original IP vendors) miss out on that portion of the revenue. The overproduced ICs can have trojan present inside them because they undergo minimal testing . This questions the reliability of the chip. The trojan could reveal confidential information to the attacker or lock the chip at critical juncture. This has serious consequences in case of safety critical applications. Counterfeit ICs are clones of the original ICs that were fabricated illegally to resemble them almost exactly. They include out-of-spec, remarked, and recycled ones, typically salvaged from discarded electronic appliances. Reverse engineering of an IC is the process of obtaining the design/technology information of an IC using imaging techniques.

Logic Locking

To protect against these threats, various techniques have been introduced. The term logic locking was coined by EPIC [3]. It involved the introduction of key gates in the given

gate-level netlist. The logic-locking secret, also known as the key, is the only way to restore the locked circuit's proper functionality. Only approved or reliable parties, such as IP owners or original design houses, have access to this key. Post manufacturing, the correct key value is loaded to restore the original functionality. The key inputs could be loaded from and stored in a tamper-proof on-chip memory. In this way, key gates hide the functionality of IC from untrusted entities. Even if the attacker can extract the gate-level netlist from reverse engineering, it will be locked because of the key gates. 'n' key inputs can have 2^n possibilities. Original functionality is only recovered if the correct key is loaded.

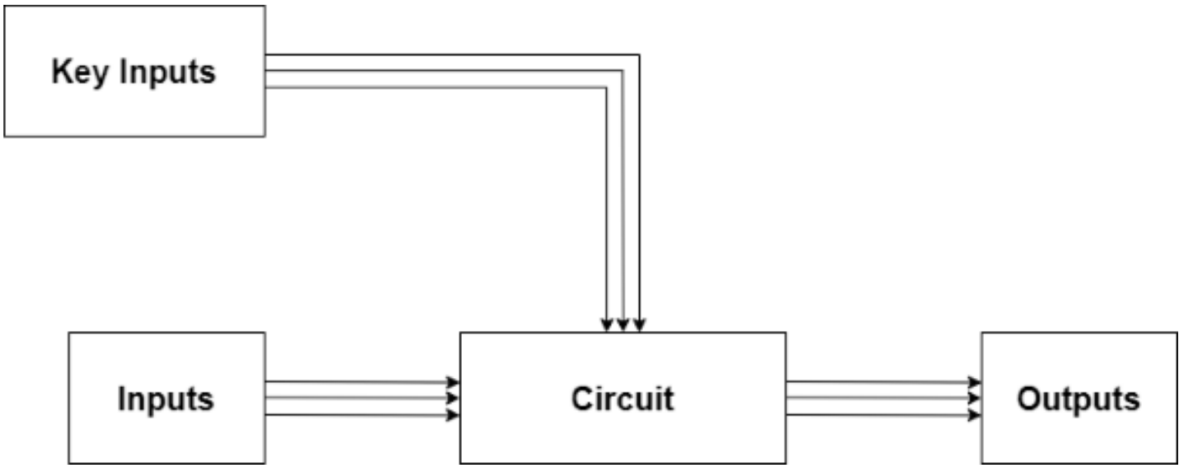

Fig. Overview of Logic Encryption

The SAT attack

Prior to 2015, most of the logic locking schemes focussed on increasing the output corruption of logic circuits. However in 2015, a potent attack was introduced, called SAT attack, which made use of boolean satisfiability to break most of the logic locking schemes. It requires two things to attack; one is locked gate level net-list and other being activated chip (oracle). In this attack, a miter circuit is constructed of the locked netlist such that there are two different keys but same primary inputs. The clauses are constructed for the above miter and a satisfying clause for primary input is found out such that outputs are different for two keys. This is referred to as distinguishing input pattern (DIP). DIP is then applied to oracle to find the correct output. The correct output for the given DIP is added as a constraint to SAT. It ensures that for the particular DIP,

solver eliminates all the keys having outputs different from correct output. Thus, in one iteration multiple keys can be eliminated. This process is repeated until there are no DIPs. After that, the key which satisfies all the DIP constraints emerges as the correct one.

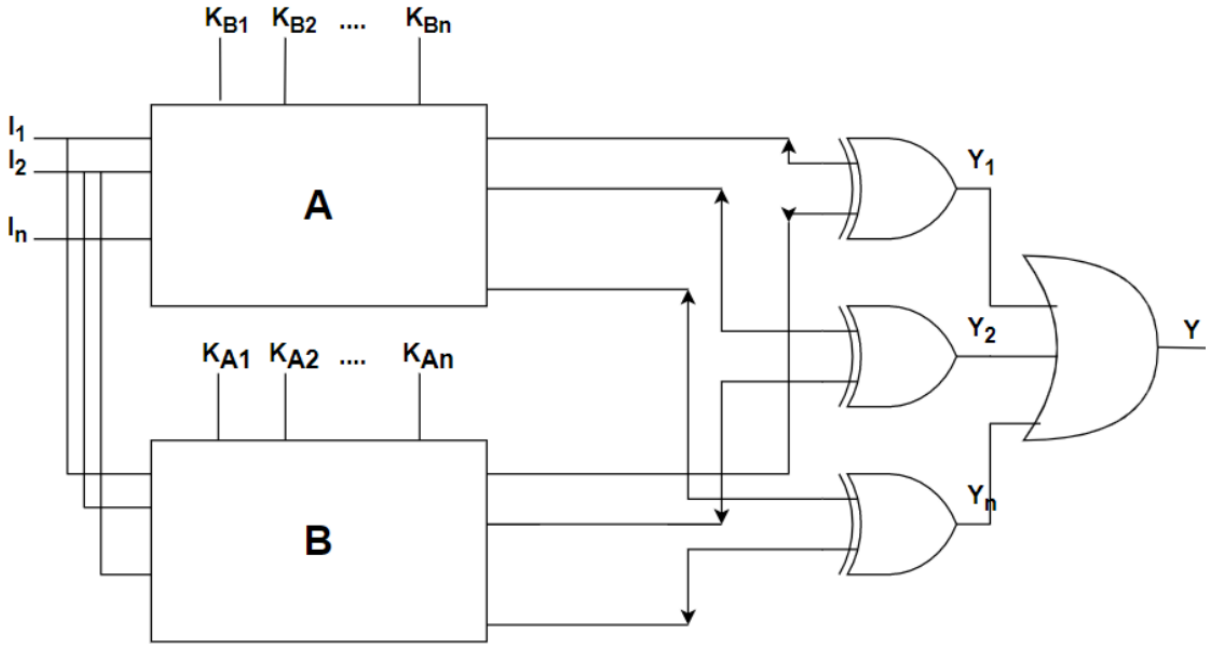

Figure: Miter Circuit

Function: *decrypt*.
Inputs: C and *eval*.
Output: $\vec{K_C}$.

1: $i := 1$
2: $F_1 = C(\vec{X}, \vec{K_1}, \vec{Y_1}) \wedge C(\vec{X}, \vec{K_2}, \vec{Y_2})$
3: **while** $sat[F_i \wedge (\vec{Y_1} \neq \vec{Y_2})]$ **do**
4: $\quad \vec{X_i^d} := sat\_assignment_{\vec{X}}[F_i \wedge (\vec{Y_1} \neq \vec{Y_2})]$
5: $\quad \vec{Y_i^d} := eval(\vec{X_i^d})$
6: $\quad F_{i+1} := F_i \wedge C(\vec{X_i^d}, \vec{K_1}, \vec{Y_i^d}) \wedge C(\vec{X_i^d}, \vec{K_2}, \vec{Y_i^d})$
7: $\quad i := i + 1$
8: **end while**
9: $\vec{K_C} := sat\_assignment_{\vec{K_1}}(F_i)$

Figure: Algorithm

Results

ISCAS-85 Circuit	Circuit Functionality
c432	27-channel interrupt controller
c499	32-bit SEC circuit
c1908	16-bit SEC/DED circuit
c5315	9-bit ALU
c6288	16x16 multiplier
c7552	32-bit adder/comparator

Figure ISCAS-85 Circuits

Benchmark Name	Key Length	# SAT Iterations	Attack Time (seconds)
c17	9	4	0.015625
c432	32	21	0.203125
	64	26	0.296875
	128	60	1.65625
c499	32	6	0.296875
	64	14	0.421875
	128	30	2.578125
c1908	32	10	0.578125
	64	18	1.03125
	128	31	2.5625
c5315	32	10	1.296875
	64	17	1.8125
	128	32	3.0625

Figure ISCAS-85 Circuits Results

Benchmark Name	Key Length	# SAT Iterations	Attack Time (seconds)
b14	32	16	26.53125
	64	32	33.875
	128	57	43.859375
b15	32	16	23.390625
	64	22	23.296875
	128	34	29.4375
b17	32	32	203.234375
	64	36	283.84375
	128	52	155.921875

Figure ISCAS-99/ITC-99 Circuits Results

Benchmark Name	Key Length	# SAT Iterations	Attack Time (seconds)
b20	32	20	66.625
	64	37	75.640625
	128	54	93.59375
b21	32	22	65.359375
	64	34	83.859375
	128	62	116.84375
b22	32	22	199.859375
	64	40	124.03125
	128	59	154.40625

Figure ISCAS-99/ITC-99 Circuits Results