On a New Way to Read Data from Memory

David Samyde[1], Sergei Skorobogatov[2], Ross Anderson[2] and Jean-Jacques Quisquater[1]

1: Université catholique de Louvain, UCL Crypto Group
Place du Levant, 3, B-1348 Louvain-la-Neuve, Belgium
{jjq,samyde}@dice.ucl.ac.be, http://www.dice.ucl.ac.be/crypto

2: Computer Laboratory, JJ Thompson Avenue, Cambridge CB3 0FD, England
{Ross.Anderson,Sergei.Skorobogatov}@cl.cam.ac.uk

Abstract

This paper explains a new family of techniques to extract data from semiconductor memory, without using the read-out circuitry provided for the purpose. What these techniques have in common is the use of semi-invasive probing methods to induce measurable changes in the analogue characteristics of the memory cells of interest. The basic idea is that when a memory cell, or read-out amplifier, is scanned appropriately with a laser, the resulting increase in leakage current depends on its state; the same happens when we induce an eddy current in a cell. These perturbations can be carried out at a level that does not modify the stored value, but still enables it to be read out. Our techniques build on a number of recent advances in semi-invasive attack techniques [1], low temperature data remanence [2, 3], electromagnetic analysis [4] and eddy current induction [5]. They can be used against a wide range of memory structures, from registers through RAM to FLASH. We have demonstrated their practicality by reading out DES keys stored in RAM without using the normal read-out circuits. This suggests that vendors of products such as smartcards and secure microcontrollers should review their memory encryption, access control and other storage security issues with care.

Keywords. Smartcards, tamper resistance, data remanence, electromagnetic security, semi-invasive attacks, optical probing, eddy current attack.

1 Introduction

The goal of this work was to explore new ways of recovering data directly from the memory of smartcards and other security processors without using the read operations provided by their vendors for that purpose, thereby circumventing any access controls and reading out secret data directly.

The traditional way of reading out data from smartcard memories involved an invasive attack using mechanical probing, usually of the processor's bus [6, 7]. Such attacks involve physically depackaging the chip and reading out its internal state by making direct electrical connections to internal components using microprobes. This is becoming more difficult for a number of reasons, ranging from shrinking feature sizes to the use of hardware access control circuits for on-chip memory.

Recently, our two teams have been developing semi-invasive attacks, in which the chip is still depackaged, but where no direct electrical contact is made and the chip passivation remains intact. Examples of such attacks include optical probing [1], in which a laser is used to induce a transient fault in one or more gates in such a way as to cause information leakage; and eddy current attacks in which a similar effect is achieved by bringing a small coil close to the surface of the chip and inducing a large transient magnetic field [5].

The natural progression from this attack

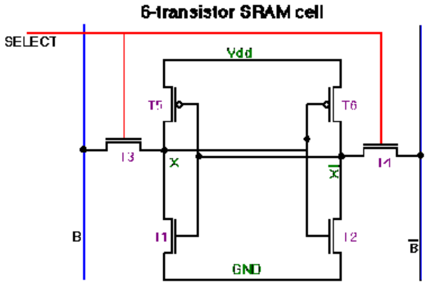

Figure 1: The architecture of an SRAM cell

technology was to investigate whether semi-invasive techniques can be used to read out the state of a memory cell in a non-destructive way. The answer, as we shall see, is yes. We will describe the techniques in the context of CMOS RAM, but they have much wider applicability.

2 Optical Read-out of CMOS RAM

The structure of a typical SRAM cell is shown in figure 1. Two inverters are built from pairs of p- and n-channel transistors. The output of the first inverter is connected to the input of the second, and vice versa. Two n-channel transistors are used to read data from it and write data into it. A read-write amplifier based on a differential structure gives access to the cell (figure 2).

To analyze the structure of SRAM memory we used a red laser focused on the chip surface using a microscope. As photons from the red laser (650nm wavelength) have energy larger than the silicon band gap, they will ionize active areas inside the chip. If the photons reach the area near p-n junctions, a photocurrent will be produced due to the photovoltaic effect. When the photons hit the p- or n-channel area, this will decrease the resistance of the channel by injecting free carriers. In each CMOS inverter, there are six p-n junctions; there are also two resistors corresponding to n- and p-channels.

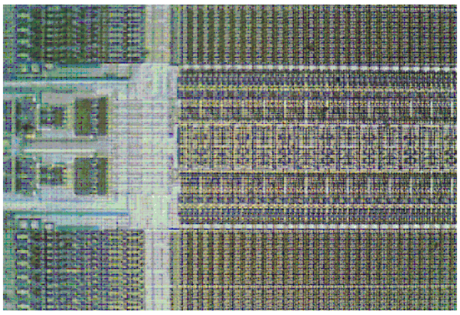

Figure 2: Internal structure of a RAM (Amplifiers and cells)

The fact that enables us to read a memory cell's state is that the decrease in resistance is noticeable for closed channels, and almost negligible for open channels. Thus, by aiming the laser beam at an appropriate transistor or transistors, we can distinguish between the two possible memory states. (A similar technique was used in [1] to switch the state of memory bits; nondestrictive read-out involves using a lower-power laser beam.)

In our first experiment, we built a map of the active areas in a microcontroller by measuring the photocurrent induced by laser scanning the chip surface. The chip was mounted on an X-Y motorized stage with 0.1μm resolution. The result of the scan is shown in figure 3. The active areas can be seen as they produce higher current, but most of the chip is covered with metal layers which the laser cannot penetrate, so these areas do not produce any current. We used this picture as a reference to the results obtained from a powered chip.

Our next experiment was done with an operating chip. It was programmed to allow us to upload any value into its RAM and then stop the chip operation. The result of the scanning with memory cells loaded with random data is shown in figure 4. It can be seen that memory cells have different states: where the cell holds a '1' the top is brighter, and where it is a '0' the bottom is. Thus the sixteen bits held in the locations scanned are

1	1	0	0
1	1	1	0
1	1	1	1
1	1	1	1

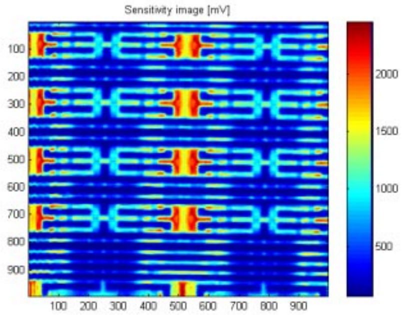

Figure 3: Laser scan of unpowered memory

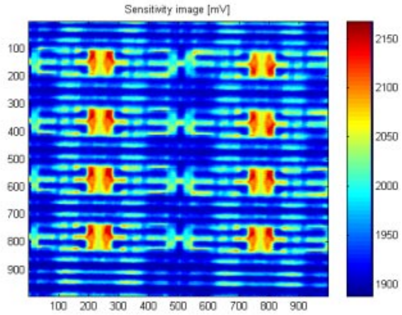

Figure 4: Laser scan of powered-up memory with state

Our experiments are somewhat similar to results published by Sandia Labs [8], but with a number of differences. They were done without using extremely expensive laser scanning microscopes; we scanned the chip from its top side; and instead of sending constant current through the chip, we used a constant voltage supply and measured current as in a standard power analysis [9].

3 Electromagnetic attack

It is also possible to use electromagnetic induction to scan a semiconductor. In [5] we explained a low-cost attack in which we used a camera flash, a needle and some wire to insert faults into a cryptographic processor. We built a miniature inductor by wrapping several hundred turns of fine wire around the tip of a microprobe needle. A current injected into this coil will create a magnetic field, and the needle will concentrate the field lines. We obtained the current from a camera, by connecting the coil where the flashbulb should have been. The test probe was then placed a few microns over the surface of the target processor. The magnetic field creates an eddy current in the chip, and we sensed this in order to build a map of the chip (Figure 5).

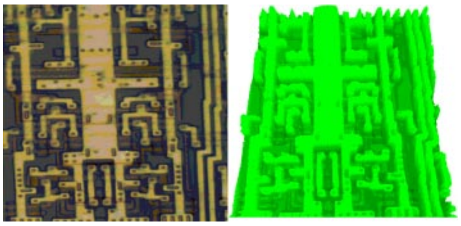

Figure 5: A map built using eddy current and a picture of the same area

We experimented to see whether this fault induction technique could also be used for nondestructive readout. With the same sensor we used to scan the chip, we created a small perturbation on a memory cell. Our idea was simple: to move for a very short time the polarization point of the transistors. As long as the polarization point does not return to its initial state as the same speed in both case, it is possible to know if the transistor is locked at "0" or at "1". We therefore tried to do a timing attack. In practice we found that the timing difference was not enough to distinguish memory states; however, the intensity of the current necessary to recover the initial value of the polarization point was noticeably different between the zero state and the one state.

We managed to recover several bytes from static RAM and FLASH. The two architectures are very different when it is time to look at one cell. But as long as the transistors do not react in the same way when their polarization point is not the same, it seems to be possible to measure the difference.

With our crude equipment, it turned out to be fairly difficult to create enough current on the chip without disturbing the content of any memory cells. In particular, read-write amplifiers are rather sensitive, and even a little perturbation of one of these components will drive the output of a whole row or column to a fixed value.

We have therefore focussed our practical research on refining the laser read-out method. However, with some combination of better equipment, improved lab technique and more sophisticated signal processing, we believe it may be practical to use electromagnetic techniques for memory read-out. It is certainly possible for us on a small scale, and needs to be considered for high-assurance products.

Thus, although it is helpful to give a smartcard chip an opaque passivation layer, it is not sufficient. A continuous metal layer would be preferable – though even this would not block attacks based on the use of infrared lasers through the rear of the chip, or the use of X-rays. For that, more active countermeasures are indicated, as we will discuss below.

4 Freezing and probing

The direct memory read-out teachiques described above are effective but slow. They are adequate for reading out data from chips that can be stopped in the target state; however, smartcard chips typically have defences against under-clocking such as reset circuitry or even some use of dynamic logic [7].

In [3] we explained how to freeze a static RAM in order to maintain the integrity of the data once the power has been switched off. We used the same technique, but replaced the Peltier plate by a cooling spray or liquid nitrogen. Frozen RAM maintains its content for significantly longer – from minutes to hours. We used this method to maintain data in SRAM in order to read its content off line. In particular, we froze a static RAM and recovered a 56-bit DES key.

We tested our attack on several static random access memories from different silicon manufacturers, and a few flash memories. We always managed to extract data by one method or another.

5 Countermeasures

A modern high-security smartcard will have its CPU implemented using random place-and-route, so that there are no visible registers; the transistors that make up the registers are scattered across the silicon. (Of course, for performance reasons they cannot be scattered too widely.) It will also have some kind of memory encryption, so that data written to and read from the bulk memory structures are at least lightly enciphered (doing more than a few rounds of a block cipher may impose a noticeable performance penalty; see [12]). However, in current designs, not all memory can be enciphered; the boot code and the master key have to be kept somewhere. Also, where bulk memory read-out becomes economical, ad-hoc ciphering techniques are likely to become vulnerable.

More attention should be paid to techniques such as the use of logic with built-in alarm propagation [10]. At a very least, it seems prudent to include low-temperature alarm sensors in smartcards, as well as sensors for ionising radiation of various kinds from infrared through X-rays.

As feature sizes shrink, the opportunity should be seized to beef up memory encryption to the maximum extent consistent with allowable memory latency. The use of self-timed circuits can also help, as it makes it harder for an attacker to know when to freeze a circuit for analysis. Techniques for alarmed off-chip storage of cryptographic keys, as in [11], also bear further study. In the G3Card project, we have developed prototype smartcard microcontrollers based on self-timed redundant logic with built-in alarm propagation, which can deal with many of the concerns raised by the attack techniques described in this paper [12].

6 Conclusion

If valuable data are present in the clear in memory for just one clock cycle in a location that an attacker can deduce, and the state can be frozen (whether physically, using low temperature, or by some other means such as stopping the clock), then it is likely to be possible for an attacker to read this data out using optical or electromagnetic probing techniques. The investment in skills and equipment required to carry out such attacks is significantly lower than that needed for full invasive attacks. Hardware countermeasures will be necessary for any processors required to resist capable hardware attacks.

7 Acknowledgments

We want to thank Cedric Volant for his help. We also want to thank Paul Kocher for a very interesting discussion about selection function for differential power analysis at Crypto 2002. Finally, this work was carried out under the aegis of the EU G3Card project, whose aim is to design next generation smartcard processors that are resistant to semi-invasive attacks [12].

References

[1] SP Skorobogatov, RJ Anderson, "Optical Fault Induction Attacks", Cryptographic Hardware and Embedded Systems 2002, Springer Lecture Notes in Computer Science, to appear; `http://www.cl.cam.ac.uk/ftp/users/rja14/faultpap3.pdf`

[2] P Gutmann, "Data Remanence in Semiconductor Devices", 10th USENIX Security Symposium, 2001; `http://www.usenix.org/events/sec01/full_papers/gutmann/gutmann.pdf`

[3] SP Skorobogatov, "Low temperature data remanence in static RAM", Technical report UCAM-CL-TR-536, University of Cambridge Computer Laboratory, June 2002; `http://www.cl.cam.ac.uk/TechReports/UCAM-CL-TR-536.pdf`

[4] D Agrawal, B Archambeault, JR Rao and P Rohatgi, "The EM Side-channel(s)", Cryptographic Hardware and Embedded Systems 2002, Springer Lecture Notes in Computer Science, to appear

[5] JJ Quisquater and D Samyde, "Eddy Current for Magnetic Analysis with Active Sensor", Proceedings of ESmart 2002, pp 185–194, Eurosmart.

[6] RJ Anderson, MG Kuhn, "Tamper Resistance – a Cautionary Note", in *Proceedings of the Second Usenix Workshop on Electronic Commerce* (Nov 96) pp 1–11; `http://www.cl.cam.ac.uk/users/rja14/tamper.html`

[7] O Kömmerling, MG Kuhn, "Design Principles for Tamper-Resistant Security Processors", USENIX Workshop on Smartcard Technology, 1999; `http://www.cl.cam.ac.uk/Research/Security/tamper/`

[8] C Ajluni, "Two New Imaging Techniques Promise To Improve IC Defect Identification", in Electronic Design v 43 no 14 (10 July 1995) pp 37-38

[9] P Kocher, J Jaffe and B Jun, *Differential Power Analysis*, In M. Wiener, editor, Advances in Cryptology - CRYPTO'99, vol. 1666 of Lecture Notes in Computer Science, pp. 388-397,Springer-Verlag, 1999. Also available at http://www.cryptography.com/dpa

[10] S Moore, R Anderson, P Cunningham, R Mullins, G Taylor, "Improving Smart Card Security Using Self-Timed Circuits", in Proceedings Async 2002 pp 211–218; `http://www.cl.cam.ac.uk/ftp/users/rja14/async2002paperV2.pdf`

[11] O Kömmerling, F Kömmerling, "Anti Tamper Encapsulation for an Integrated Circuit", patent application PCT WO 01/50530 A1 (2001)

[12] S Moore, R Anderson, R Mullins, G Taylor, and J Fournier, "Balanced Self-Checking Asynchronous Logic for Smartcard Applications", submitted to *Journal of Microprocessors and Microsystems*; available at `http://www.cl.cam.ac.uk/~swm11/papers/smartcards/micromicro.pdf`