TRUDEVICE STSM Report

Ke Jiang
Embedded Systems Laboratory (ESLAB)
Linköping University
Sweden

November 13, 2013

1 STSM Information

1.1 Visitor

Visitor: Ke Jiang
Affiliation: Embedded Systems Laboratory (ESLAB), Linköping University, Sweden

1.2 Host

Host: Prof. Lejla Batina
Affiliation: Digital Security Group (DS Group), Radboud University Nijmegen, The Netherlands

1.3 Period

September 22, 2013 - October 25, 2013

2 STSM Purpose

Embedded systems (ESs) have been widely used in all aspects of our daily lives, and are increasingly being applied in new applications, e.g., automotive control and health monitoring. Many of these applications are in safety-critical domains, in which the functional correctness and information security are vital. However, new communication interfaces that are integrated in the ES for pursing higher reliability and better performance expose the system to more severe security threats. Thereby, it is indispensable to protect the critical embedded systems against malicious security attacks.

Both of the involved institutions in the STSM have great interests and solid experiences in the design and manufacture of secure embedded devices. However, the research focuses and expertise of the two groups are quite different.

I (ESLAB, Linköping University) previously had an emphasis on system level design of secure embedded systems under various constraints, e.g., timing, resource, and power constraints; while, Prof. Lejla Batina (DS Group, Radboud University Nijmegen) has focused more on cryptography and side-channel attacks (SCA). In my previous works, I assumed that the attacker has no physical access to system details, e.g., power consumption. If a very strong attacker, who has the possibility of mounting SCA, is present, the system may be compromised. So design optimization of real-time systems against SCA is also an important problem to study.

Prof. Batina has solid experience and knowledge in SCA and corresponding countermeasures (CM). Thus, it became much easier for me to start working on the idea if I could learn from and discuss with Prof. Batina directly. In order to bridge the big knowledge gap between the two groups, and maximally exploit the excellence of both sides, I decided to visit Prof. Batina in Nijmegen as the very first step for collaborations. The concrete objective of my visit was to learn the principles behind SCA and find holistic approaches of modelling and designing robust secure devices against SCA.

3 Joint Works

During my visit in Nijmegen, I have found two interesting problems to start with through reading literature and discussing with Prof. Batina. The first problem contains two sub-problems: (1) to propose a new policy for scheduling real-time systems (RTS) and performing CM against SCA at the same time; (2) to study the robustness of different existing scheduling policies. The second problem is to find the best CM designs under different conditions. In this section, I will elaborate more on the joint projects that we would like to work on together at this moment. Let us first look at the common system model shared in both problems.

3.1 System Model

We target a mono-processor RTS. The system interacts (by wire or wireless) with other peers or service centres via a communication module. The central embedded processor runs a set of periodic tasks. If there is a dependency between two tasks $\tau_i \rightarrow \tau_j$, then τ_j can only start execution when τ_i finishes. A task τ_i is associated with a set of attributes ($\mathcal{P}_i$, $\mathcal{E}_i$, $\mathcal{L}_i$). $\mathcal{P}_i$ is the release period of τ_i and also its relative deadline. $\mathcal{E}_i$ is the Worst-Case Execution Time (WCET) of the task. $\mathcal{L}_i$ is the set of messages via which task τ_i interacts with the outside world. Each message $m_{ij} \in \mathcal{L}_i$ is associated with a length l_{ij} (in number of blocks) and a weight w_{ij} representing its relative importance (criticality).

In these works, we focus on protecting the confidentiality of communication, which is arguably the most important security concept for embedded devices. In order to provide communication confidentiality, we carry out AES encryp-

tion/decryption. To be more illustrative, we abstract the AES operations on communication messages as independent tasks, and refer to them as AES tasks. An AES task τ_{ij}^{AES} shares the same period and deadline with τ_i. And its execution time depends on the length of the message, i.e.,

$$\mathcal{E}_{ij}^{AES} = l_{ij} * \mathcal{E}^{AES}, \tag{1}$$

where $\mathcal{E}^{AES}$ is the WCET of the AES operation for one block of text.

3.2 Problem 1: Scheduling Policy as Countermeasure in Real-Time Systems

Many embedded systems, especially those in critical applications, carry out real-time computing. Such systems are referred to as real-time systems (RTS), in which the underlying scheduling policy is of central importance for ensuring functional correctness. In order to protect the cryptographic operations (we focus on AES) in RTS against SCA, two concepts can be exploited as mentioned in [1] that are

- reducing the signal-noise-ratio (SNR), $SNR = \frac{Var(Q)}{Var(N)}$, of the power consumption
- randomly disarranging the starting time of AES tasks.

Assuming AES is implemented in software, then they are not the only executions on the embedded processor which also handles other executions, e.g., data processing and device control tasks. The non-AES tasks can be modelled as noises to the AES tasks using the idea of SNR analysis. The scheduling policy has quantifiable impacts on the disarrangement of the AES operations. Thus, scheduling policies can be used to provide certain CM effects, and result in little or no architecture/compiler modifications or time overhead, thus, much more efficient in practice. Now we present the robustness quantification methods used in this work.

3.2.1 Robustness Analyses

The communication messages are protected by AES encryption as aforementioned. However, the AES operations are vulnerable to differential-power analysis attacks. So, we need to employ CM. Old CMs either require architecture modification, or are not suitable for general embedded platforms, e.g., embedded processors and controllers handling different executions.

We are interested in finding a CM that has the least or no impact on existing platforms and architectures. So, we try to obfuscate the attacker from obtaining a straight-forward correlation between the hypothetical power consumption and the obtained power traces by burying the AES tasks into the non-security task executions and shuffling the release time of all AES tasks.

In order to quantify how good the obfuscation is, we have to do a theoretical analysis. The correlation between his hypothetical power consumption

H and the obtained power traces P is measured using the Pearson correlation coefficient, i.e.

$$\rho(H,P) = \frac{Cov(H,P)}{\sqrt{Var(H)Var(P)}} = \frac{E(HP) - E(H)E(P)}{\sqrt{Var(H)Var(P)}}. \quad (2)$$

The correlation between H and P based on samples are captured as

$$r(< h_1, ..., h_S >, < p_1, ..., p_s >) = \frac{\sum_{s=1}^{S}(x_s - \bar{x})(y_s - \bar{y})}{\sqrt{\sum_{s=1}^{S}(x_s - \bar{x})^2 \sum_{s=1}^{S}(y_s - \bar{y})^2}}. \quad (3)$$

The number of samples needed to mount a successful DPA attack mainly depends on the highest occurred correlation $\rho_{max} = \rho_{k_c,t_c}$ calculated as follows,

$$\rho_{max} = \frac{\rho(H,Q)}{\sqrt{1 + \frac{1}{SNR}}} * \hat{p} * \sqrt{\frac{Var(P)}{Var(\hat{P})}}. \quad (4)$$

$\rho(H,Q)$ is the Pearson correlation between the attacker's hypothetical power consumption for the correct subkey and the power consumption of the device caused by the attacked intermediate result. SNR is the signal-noise-ratio. Let us denote the time that the highest correlation ρ_{max} between the correct hypothetical power consumption and the power consumption of the device happens as $\hat{t_c}$ and the probability that a power consumption at this time is caused by processing of an attacked intermediate result as $\hat{p}$, respectively. $Var(P)$ and $Var(\hat{P})$ are the variance of power consumption caused by an attacked intermediate result and power consumption of the device at $\hat{t_c}$, respectively.

Since we have no control on how much knowledge the attacker has, we must make an conservative assumption of $\rho(H,Q) = 1$. At the first design step, $Var(P)$ is assumed to be equal with $Var(\hat{P})$. But this can be updated to correspond to practice when more information is available, e.g., the designer can make real measurements on the processor with the intended tasks to be executed. Then, Eq. 4 can be simplified to

$$\rho_{max} = \frac{\hat{p}}{\sqrt{1 + \frac{1}{SNR}}}. \quad (5)$$

After obtaining ρ_{max}, we can calculate the lower bound of the number of samples for successfully mounting a DPA attack as follows,

$$S = 3 + 8\left(\frac{Z_a}{ln(\frac{1+\rho_{max}}{1-\rho_{max}})}\right)^2. \quad (6)$$

S is the determinant value for measuring how good the shuffling decision is.

As can be seen in Eq. 5 and 6, we can reduce SNR and $\hat{p}$ to reduce ρ_{max}, thereby, increase the difficulty of mounting DPA attacks. These two methods

can be achieved by by burying the AES tasks into the non-security task executions and shuffling the release time of all AES tasks. Because AES tasks are executing on the same processor with other tasks, the power consumption of the processor include also the power when executing the normal tasks. So it is much more difficult to find a high correlation on the hypothetical and real power traces. The power consumption of normal tasks can be modelled as noises to the attacker, which can adopt the idea of SNR calculation.

In addition, we can randomize the time point of doing AES, such that t_c is different in every release (period) of an AES encryption task. In this way, there is no repetitive pattern in the power trace for AES. Therefore, ρ_{max} is significantly reduced. Let us denote the moment of time when the highest correlation in DPA happens with the maximum of the probability distribution is located as $\hat{t}$. The maximal probability and the power consumption of the device at $\hat{t}$ is denoted as $\hat{p}$ and $\hat{P}$, respectively. Thus, the probability that $\hat{P}$ is caused by processing an attacked intermediate result is $\hat{p}$. Similarly, $\hat{P}$ is caused by processing other tasks is $(1 - \hat{p})$.

3.2.2 Two Sub-Problems

We can utilize the method discussed in the previous section to measure how good a scheduling policy is. It can be used in two different directions, i.e., to use it as the design optimization objective of a new scheduling policy or as the quantification for existing scheduling policies as stated in the following two sub-problems.

A Novel Scheduling Policy In this problem, we want to find a quasi-dynamic scheduling policy that maximizes the strength against DPA attacks, similar to [2, 3]. Our proposed solution takes the set of tasks, as well as their parameters, and messages to be protected as inputs. Then it breaks all tasks into fractions. That is, the normal tasks can be interrupted at any time instances, while AES tasks are broken between rounds. After that, it shuffles the timing order of the task fractions under various constraints, e.g., dependency and deadline. The output is a quasi-dynamic schedule that provides the strongest resistance against DPA attacks.

Study of Existing Scheduling Policies There exists a collection of dynamic scheduling policies, which delivers certain resistance against DPA attacks (though may not be very robust from the CM point of view). Therefore, it is an interesting topic to study their performances with respect to the DPA attack resistance using the quantification methods discussed previously. Some examples of scheduling policies to look at are, for example, Earliest Deadline First (EDF) and Rate-Monotonic (RM).

3.3 Problem 2: Scalable Countermeasures for Constrained Environments

In this problem, we would like to find the best combinations of implementing existing CMs such that the RTS delivers the highest resistance against DPA attacks under potential constraints. We will consider a set of CM (as well as combinations of different CM) available in literature, e.g., masking [4] and random delay insertion [2], and analyse their strengths and induced overheads, e.g., time, energy, and complexity. A table holding the trade-offs between strength and overhead among different CM is expected after the small survey study. Then we would like to find the best design decision of implementing a combination of CM for all AES operations that delivers the highest robustness against DPA attacks, and meets all constraints, e.g., cost and timing.

We assume that the system is scheduled by the earliest-deadline-first (EDF) policy. Then the timing constraint can be transferred to processor utilization test. First, the load caused by the execution of τ_i, together with the AES on its communication messages, is

$$U_{\tau_i} = \frac{\mathcal{E}_i}{\mathcal{P}_i} + \sum_{m_{ij} \in \mathcal{L}_i} \frac{\mathcal{E}_{ij}^{AES} * \mathcal{O}_{ij}^{CM}}{\mathcal{P}_i}. \tag{7}$$

$\mathcal{O}_{ij}^{CM}$ is the time overhead ratio of the chosen CM technique for AES on message m_{ij}. Now we can test the system schedulability: a set of tasks is schedulable by EDF if and only if the total utilization of all tasks is not more than 100%. The utilization $U_{\tau_i}(f)$ of task executions was defined in Eq. 7. Thereby, the schedulability of the system can be examined by the following condition

$$U_M = \sum_{\tau_i \in M} U_{\tau_i}(f) \leq 1. \tag{8}$$

Other constraints can also be formulated correspondingly. The global design objective is then to maximize the global robustness of the system against DPA attacks, i.e.,

$$\text{maximize } R = \prod \mathcal{S}_{ij}^{CM}, \tag{9}$$

where $\mathcal{S}_{ij}^{CM}$ is the strength of the chosen CMs on protecting message m_{ij}.

4 Future Collaboration

Besides the problems presented in Section 3, I also had interesting discussions with colleagues in Nijmegen on several other problems. Here are several examples for future collaborations.

4.1 Real Implementation of Proposed Techniques

We would like to carry out real implementations on the techniques proposed for solving the aforementioned problems, and then analyse the robustness and performance of the techniques.

4.2 Adaptive Intrusion Detection System for Real-Time Systems

During my visit in Nijmegen, I talked with Prof. Peter Schwabe and Pol Van Aubel about implementing adaptive intrusion detection in RTS. Although detecting system intrusion in RTS is crucial, previous intrusion detection systems lead to significant system overhead, and are not tunable under different conditions. This hinders the feasibility of having intrusion detection in RTS. So in this work, we would like to propose an adaptive method for detecting intrusions that can be adjusted depending on the currently available resources, and then delivers the best system affordable intrusion detection accuracy. How to quantify the accuracy of intrusion detection systems is still an open problem, but we would like to make some early approaches of our own.

4.3 Efficient Heuristic for Approaching Best System Designs

I also had discussions with Stjepan Picek, a PhD student of Prof. Batina, whose previous research was focused on genetic algorithm (GA). GA is a metaheuristic that solves complex optimization problems efficiently. Thereby, we would like to collaborate on my future works, in which he could validate his GA framework on the combinatorial optimization problems, and I could have efficient solution techniques.

5 Planned Publication

At this moment, we have planned two potential publications resulted from the STSM.

- A Novel Scheduling Policy for Protecting Real-Time Systems Against Differential-Power Analysis Attacks, *Ke Jiang, Lejla Batina, Petru Eles, Zebo Peng, ...*, TBD
- System-Level Design Optimization of DPA Resistant Embedded Systems, *Ke Jiang, Lejla Batina, Petru Eles, Zebo Peng, ...*, TBD

These two papers correspond to the works mentioned in Section 3. The works mentioned in Section 4 are still in early discussions, and have not yet reached concrete publication plans.

6 Other Notes and Reflections

The Digital Security Group where Prof. Batina works has a very interesting seminar setup on Fridays, which is referred to as the "lunch colloquium". People take their own lunch to a room, and listen to a presentation while enjoying their food. This idea in fact works very efficiently, since most of the group members

will be available during lunch time. Thus it reduces the possibility of collisions, and maximizes the number of available members. It gives the group members a perfect chance of gathering together and learning new stuff. I got to attend three lunch colloquiums, and the brief ideas of the talks are as follows,

- Talk 1: A KPMG employee gave a talk on the attacks, e.g., fishing and frauding, existed in modern online banking systems. There is even a forum that people sell and buy different attacking tools. It can be extremely easy, but powerful, malwares that affects 3 percent of all Windows computers. Mainly targeted systems are Windows and Android.
- Talk 2: A DS member talked about the design of secure circuits at low hardware levels.
- Talk 3: A DS member talked about secure execution environments, e.g., Intel IPT (identity Protection Technology) and ARM TrustZone. Both tries to isolate the security related executions. IPT even has a dedicated RISC processor running JAVA VitrualMachine for security.

I really had a nice experience doing this STSM in Radboud University Nijmegen. The colleagues there were a superb company. All of them were very considerate and nice who invited me in all kinds of occasions, for example, forming lunch groups and eating out for dinner in different restaurants. It was also beneficial for me to talk with people from a different research community who have totally different mindsets and knowledge. This difference triggered a lot of interesting discussions, and some of them will potentially turn into joint publications.

Acknowledgement

I would like to thank **TRUDEVICE** (COST action IC1204) for funding my visit to Prof. Lejla Batina in the Netherlands. It was an excellent opportunity and memorable experience for me to learn from and interact with another research society. This short visit triggered a lot of interesting discussions, some of which may turn into potential joint publications in the topic of designing secure embedded devices.

References

[1] S. Mangard, "Hardware countermeasures against dpa–a statistical analysis of their effectiveness," in *Topics in Cryptology–CT-RSA 2004*. Springer, 2004, pp. 222–235.

[2] Y. Lu, M. P. O'Neill, and J. V. McCanny, "Fpga implementation and analysis of random delay insertion countermeasure against dpa," in *International Conference on ICECE Technology*. IEEE, 2008, pp. 201–208.

[3] J. A. Ambrose, R. G. Ragel, and S. Parameswaran, "Rijid: random code injection to mask power analysis based side channel attacks," in *Proceedings of the 44th annual Design Automation Conference.* ACM, 2007, pp. 489–492.

[4] J. Blömer, J. Guajardo, and V. Krummel, "Provably secure masking of aes," in *Selected Areas in Cryptography.* Springer, 2005, pp. 69–83.