COGITO

Deliverable D1.2.2

Specification of the Use Cases, final version

Editor Jean-Louis Lanet
Authors by alphabetical order:
Thierno Barry, CEA
Damien Couroussé, CEA
Philippe Jaillon, ENSMSE
Jean-Louis Lanet, INRIA
Hélène Le Bouder, INRIA
Bruno Robisson, ENSMSE

Version Revision: 263
Date lun., 19 oct. 2015 13:55:19 +0200
CEA ref. DACLE – 15-0698

Contents

List of Figures

1 Executive Summary

This report focuses on the different use cases of the COGITO project. We have chosen different algorithms that have different characteristics in terms of usage frequency, security level and side channel signature. The first one is the cryptographic algorithm AES; its side channel signature is of prime importance while its usage is moderate but could be complex to implement. The second one, the PIN code verification algorithm, is rarely used; its side channel signature is very important and it has a very low complexity. The last one, the memory prefetch in a Java Virtual Machine, is very frequently used, presents a very low complexity but is rarely protected against side channel and security aspect is often considered as less important.

The case studies within the COGITO project contribute to the gathering of evidence for the benefits of using code polymorphism. The results are aimed at raising awareness of the benefits of using the `deGoal` technology in the smart card domain and more generally in the domain of secure embedded devices. The objective is to evaluate the advantages and drawbacks of using code polymorphism in different phases of the smart card development. For that purpose, beside the description of the algorithms we describe the evaluation process and the related metrics we want to collect in order to asses if this technology is affordable for this application domain.

It must be pointed out, that the metrics presented herecan evolve in the future, reflecting our progresses within the project. For this reason, this document will probably be updated with the final results.

2 Cryptographic Algorithm: AES

2.1 Description

AES has become a kind of benchmark reference when it comes to evaluate the effectiveness of a generic protection scheme.

The AES cipher executes a number of round transformations on the input plain-text, where the output of each round is the input of the next one. The number of round r is determined by the initial key length: a 128-bit key uses 10 rounds, a 192-bit uses 12 and 256-bit key uses 14. Overall the implementation of AES is achieved with only three types of operations: `xor` operation, table lookups and 1-byte shifts. Each step operates on 128-bit blocks of data viewed as a 4×4 matrices of bytes.

```
input : A plaintext $p$ of length 128 bits
input : A key $k$ of length 128 bits
output: A ciphertext $c$ of length 128 bits
$\langle k_0, k_1, \ldots, k_r \rangle \leftarrow$ KeySchedule($k$);
$c \leftarrow$ AddRoundKey ($p, k_0$);
for $i \leftarrow 1$ to $r$ do
    $c \leftarrow$ SubBytes ($c$);
    $c \leftarrow$ ShiftRows ($c$);
    $c \leftarrow$ MixColumns ($c$);
    $c \leftarrow$ AddRoundKey ($c$);
end
$c \leftarrow$ SubBytes ($c$);
$c \leftarrow$ ShiftRows ($c$);
$c \leftarrow$ AddRoundKey ($c$);
```

Algorithm 1: The AES encryption algorithm

Algorithm 1 describes AES encryption. Each round but the latter is composed of four processing stages:

- `AddRoundKey`: this stage perform a bytes-addition in $\mathbb{F}_2$ between the plain-text and the round key meaning a byte-`xor` operation.
- `SubBytes`: each byte in the plain-text is substituted by another one from the look-up table (S-Box).
- `ShiftRows`: each row from $i = 0$ to $i = 3$ of the 4x4 *state* matrix is cyclically i-bytes left shifted
- `MixColumns`: the four bytes of each column of the *state* are combined using an invertible linear transformation.

Each round is composed of the same steps, except for the first that starts with an extra addition of a round key and the last where the `MixColumns` operation is skipped.

2.2 Use of code polymorphism

Each processing stage in AES can be provided with a polymorphic implementation. From the point of view of DPA attacks, the weakest points in AES are the execution of the first `SubBytes` and the last `AddRoundKey` [MOP07]. In this use case, we will consider the overhead cost of using code polymorphism for each of the sub-stages of AES with regards to security aspects. In particular, considering the flexibility of the COGITO protection scheme, it is possible to selectively apply code polymorphism on the code portions that are more likely to be under attacks. Otherwise said: on the one hand, a full polymorphic implementation is expected to provide the highest degree of security but with a higher overhead because of the execution time required for code generation. On the other hand, carefully selecting the portions of AES where code polymorphism is applied is likely to reduce the protection cost at the risk of providing lower security.

Several usage scenarii will be considered, considering a full polymorphic implementation of AES, or selective application of polymorphism. We will provide for each scenario figures for the security performance and for the protection overhead.

3 Secure Algorithm: PIN Code verification

3.1 Motivation

The PIN code verification procedure is one of the most attacked. Its purpose is to verify if the proposed PIN value is equal to the stored one. Many attacks have been elaborated to retrieve the code including attacks on the terminal itself [CC11] using terminal sensors. But the most current attack are timing attack against the verification algorithm. While the algorithm compares the proposed value byte by byte one should infer analysing the behaviour of the target on which byte the algorithm stops the verification. Hereafter is a naive verification algorithm.

At the beginning (line 8) the program checks if the user did not more than 3 trials before. Then it compares line 13 byte by byte if the stored value `pin` is equal or not to the proposed value `buffer`. If it does not match, the trial counter is decremented and the program memorizes this state and returns a negative answer.

```
 1 #define maxTries 3
 2 int triesLeft = maxTries;
 3
 4 boolean verify (short[] buffer, short ofs, short len)
 5 {
 6   // No comparison if PIN is blocked
 7   authenticated = false;
 8   if (triesLeft < 0)
 9     return false ;
10   // Main comparison
11   for(short i=0; i < len; i++)
12   {
13     if (buffer[ofs+i] != pin[i])
14     {
15       triesLeft— ;
16       return false ;
17     }
18   }
19   // Comparison is successful
20   triesLeft = maxTries ;
21   authenticated = true ;
22   return true ;
23 }
```

Figure 1: An implementation example of the secure algorithm in C

Of course, the time needed by the algorithm to evaluate each digit increases with time and thus offers a side channel to the attacker who needs just to evaluate the response time of the algorithm to infer if the proposed digit was correct or not.

Such a basic attack on a naive implementation of the PIN verification algorithm is very easy to set up. For this reason smart card manufacturer have developed algorithms that are resistant to timing attacks but also fault attacks. A fault attack needs to be synchronized, so often the first step of this attack is to reverse the algorithm in order to have the know-how on the precise instant to fire the laser beam.

We propose with this use case to evaluate the usage of `deGoal` to generate code that will avoid this synchronization process. This algorithm is rarely used only once per session except while the card is under attack. The time needed to generate the code is not a constraint.

3.2 Description

The algorithm that we use for this project is an improved version of the naive C version proposed in the previous section. In particular, we have an integrity protection of the different fields (mainly redundancy with bit inversion), which are checked before each usage. Of course PIN trial incrementation before use is the basic coding rule. We implement also the secure conditional to detect transient fault. The secure condition test twice the same value; if the second test is different from the precedent without having modified any variable it implies that the environment (a fault attack) did it. We implement also step counters to avoid control flow transfer due to a fault. If the program counter is incremented by the environment (fault attack) it can jump anywhere and potentially bypass some security tests. For this reason we mark important step in the algorithm and we verify before returning that all steps have been executed. We also use constant time evaluation procedure to eliminate the timing attack and secure constants to prevent an attack targeting boolean value. A false value is obtained with all the bit of the cell to zero, and a true value is obtain if the contain of the cell is different of zero. Of course an external event can modify a false value easily with a true value. We transfer the sensitive fields into the RAM at the beginning of the algorithm which allows us to work only on the `C` stack. We ensure consistency of the different fields used in the algorithm to avoid DoS attacks.

Such an algorithm is highly resistant to several attacks. Normally we have to introduce random computations to desynchronize the traces. Code polymorphism will be used for that purpose but also to avoid reverse engineering of the algorithm taking care that the other protections are not affected by code polymorphism. We present hereafter the secure algorithm.

```
boolean verify (short[] try) {
int triesLeft;
  triesLeft = SecuredLoadTries();
  if (triesLeft < 0) {
    return false ;
  }
  else {
    triesLeft-- ;
    SecuredStoreTries(triesLeft) ;
    if( SecuredLoadTries() != triesLeft ) {
      return false ;
```

```
12          }
13          else {
14            if( ConstantTimeComparaison(try , PIN) == true ) {
15              SecuredStoreTries(maxTries) ;
16              if( SecuredLoadTries() != maxTries ) {
17                return false ;
18              }
19              return true ;
20            }
21            else {
22              return false ;
23            }
24          }
25        }
26      }
```

At line 3, the function `SecureLoadTries()` verifies the integrity of the `TriesLeft` field, while the `SecuredStoreTries()` function computes the integrity and performs the secure storage. The comparison of the two arrays must be done in constant time using the function `ConstantTimeComparaison()`.

Each time an abnormal behavior is detected a countermeasure must be taken. This part is out of the study but can be blocking the current application by modifying its life cycle or blocking completely the card (card is mute).

3.3 Use of code polymorphism

In the secure algorithm, the evaluation section, *i.e.* the main comparison, should not be polymorphic until we have the guarantee that the constant time evaluation is not altered. All the other parts of the algorithm can be impacted by polymorphism.

The evaluation process of the code polymorphism will take into account the ability to recognize patterns by side channel but also the level of randomization brought by polymorphism. For that purpose, we will evaluate the EM traces with or without polymorphism in order to assess the benefits. The memory footprint and the code generation process will be also evaluated.

The ability to reverse the code is a metric of the efficiency of the deGoal approach. If acquiring a couple of trace the attacker is able to reverse, to normalize the code *i.e.* eliminate the useless instruction, the attacker will be able to bypass the countermeasure. We will evaluate the ability to reverse the assembly code using our correlation algorithms.

4 General Purpose Algorithm: the JVM

4.1 Motivation

Recently, side channel analysis has become of interest to be used for reverse engineering purposes (e.g. [VWG06], [OSS$^+$]). Reverse engineering of software is primarily known as the process of discovering the source code from the software binaries or executables. It often involves detailed analysis of the program and uses many methods such as analysis through observing information exchange, disassembling or decompiling. There are many tools available on-line that provide all of these functions and that even combine them to acquire the source code. Reverse engineering of Java Card applets is much more difficult because the attacker does not have access to the binary files.

Power or EM analysis can be used to acquire parts of bytecode in order to be reverse engineered. Once the collection of execution traces have been recovered by one of the attack methods, it can be analysed. Then, various techniques may be constructed to affect its function on the card or reveal sensitive information. Of course there is no guarantee that the collection of power traces covers the whole software to analyse. The attacker has to exercise all the data input in order to generate different traces to obtain a high coverage of the traces.

The traces can then be analysed with correlation or pattern recognition. To mitigate such an attack, a solution is to change the pattern of each bytecode in order reduce the probability of recognition. The pattern can be changed in the time axis including random delay or in the form of the pattern thanks to code polymorphism.

4.2 Description

To successfully reverse engineer an unknown Java Card applet from a smart card, first, a dictionary of patterns must be set up using a reference card. The process of reverse engineering begins with the identification of bytecode instructions in the collected power traces. Since the reference card is programmable, it allows the attacker to run testing applets that repeat one known instruction or that repeat a small sequence of known instructions multiple times and then reveal a repeating pattern in the power trace. By comparing the individual parts of the power trace that represent one instruction to each other, a unique template that defines an instruction by its power trace can be constructed. A template is usually constructed as an average power trace of multiple measurements of the same instruction. Moreover using correlation analysis one can recognise the common part of each instruction.

In fact a virtual processor acts as a real one with the same sequence of prefetch, decode and execute cycle. In the JCVM interpretation loop, as described below, these sequences are clearly identifiable: the first part is the preamble, is to say the 'prefetch - decode cycle of a virtual processor, then the second part represents the execute cycle of bytecode, followed by a postamble that depends on the type of bytecode being executed.

```
while (true) {
  bc_item = NULL; /* Preambule */
  handler = bytecode_table[*vm_pc]; /*Prefetch + decode*/
  vm_pc++;
```

```
    bc_action = handler();          /* Execute */
    if (bc_action < 0) {
       if (!handle_excep())
         return false;
  }
    switch (bc_action)     {/* postambule */
      case 0:      continue;
      case ACTION_RETURN:
                    i = handle_return(init_frame);
                    if (i == RUN_RETURN)
                        return true;
                    go = (bool)(i != RETURN_FAIL);
                    break;
      case ACTION_INVOKE:
                    exec_method = (method_t *)bc_item;
                    break;
      case ACTION_NATIVE:
                    go = handle_native();
                    break;
      case ACTION_NEW:
                    go = handle_new();
                    break;
      case ACTION_THROW:
                    go = handle_throw((ref_t *)bc_item);
                    break;
     }
}
```

The `handler` is a function pointer defined as: `typedef int16 (*bc_handler)(void);`. The system table `bytecode_table` associates a bytecode value to a function pointers of type `bc_handler`:

```
const bc_handler bytecode_table[256] =
{
      BC_nop,                          /* 0                    */
      BC_iconst_0,                     /* 1    BC_aconst_null */
      BC_iconst_m1,                    /* 2                    */
      BC_iconst_0,                     /* 3                    */
      ...
      }
```

Each of these functions describes the behaviour of the instructions.

```
int16 BC_iconst_0(void)
{
      return _iconst(0);
}
```

When executing the preamble, the processor handles two parameters that are attached to the bytecode being read: (1) the index `vm_pc` (virtual machine program counter) in the table of bytecodes,

(2) the address of the function handler to be executed. These two parameters are exploitable by an attacker to recover information about the program under execution. `vm_pc` is the pointer to the array of bytecodes representing the method, so it is a pointer to an array of bytes. By using side channel analysis, the attacker can obtain numerous information: the variation in the preamble power curves indicates the value of `handler` or `bc_action`. The first one gives directly the instruction, the second gives the address of the instruction. Then, the execution of the bytecode provides a trace easily identifiable with pattern recognition analysis; the execution time is also highly characteristic of the bytecode value. So an attacker has in each bytecode execution four correlated information about the instruction to be recognised.

4.3 Use of code polymorphism

We will limit the use of polymorphism on the prefetch cycle only. The difficulty with this case study is the high frequency of the code execution. Indeed, considering that an attacker needs a very low number of traces to recover information about the program under execution, it would be desirable to change the form of a polymorphic prefetch at *each* execution. However, considering the overhead of code generation with regards to code execution of the polymorphic code generated, this solution is hardly acceptable in terms of performance. This issue can be a limit of our approach that we want to investigate. We have backup solutions if the ratio compilation time *vs.* execution time is too high.

5 Evaluation process: collecting metrics

The partners EMSE and INRIA have different expertise in this area. EMSE is more familiar with side channel and cryptography while INRIA is is more familiar in software development and reverse engineering using side channel. None of these research laboratories have a prior knowledge on the technology `deGoal` used to achieve code polymorphism.

5.1 Aims

By stating the goals we define the expected effects of using code polymorphism. The aims of the case studies are different. The first one is related to the development of cryptographic functions, the second to a checker algorithm and the last one to a non security oriented function.

The aims of the cryptographic case study is to verify that the EM or power signatures can be mitigated using code polymorphism, in particular with the `deGoal` technology.

For that purpose we need to analyse the power signature of the traces with and without code polymorphism, without any other countermeasure. These kind of algorithms are very often the target of side channel attacks. Numerous techniques are used to eliminate the randomness provided by implementations and polymorphic code generator could be attacked using statistical analysis. Thus, we will be able to evaluate the effectiveness of this technology. The experiments must be done on a general purpose board using the chosen processor to eradicate all side effects. This case study will provide a qualitative assessment but also a security evaluation of the technology.

The aims of the secure case study is to evaluate the cost for engineer to use the `deGoal` technology.

For this case study, we want to provide quantitative evaluation for the development process. In fact, for the adoption of any technology we need to convince management that such a technology is affordable. For this reason, we have to collect metrics concerning the development for evaluating the cost and efficiency of `deGoal` technology. This algorithm rarely used (once per session) is often the target of attacks and is coded very carefully to avoid any error, side channel leakage or fault attack.

The aims of the non secure algorithm case study is to evaluate the temporal overhead of a dynamic solution like `deGoal` *versus* a static diversification solution.

The main idea within this case study is to investigate the limits of this technology by choosing an algorithm which is often used, around each 5-10 millisecond, sensible to side channel attacks for which some static solutions are available. We will focus on the performance of the `deGoal` technology and the generated overhead. We define two main hypotheses: this technology improves the security of smart card products and its is affordable for this application domain.

5.2 Metrics

Distinguish threshold Differential Power Analysis (or Correlation Power Analysis) will be conducted on unprotected and protected implementations. The DPA analysis computes the sample estimation r of Pearson's correlation coefficient ρ between each trace and the model, for each possible

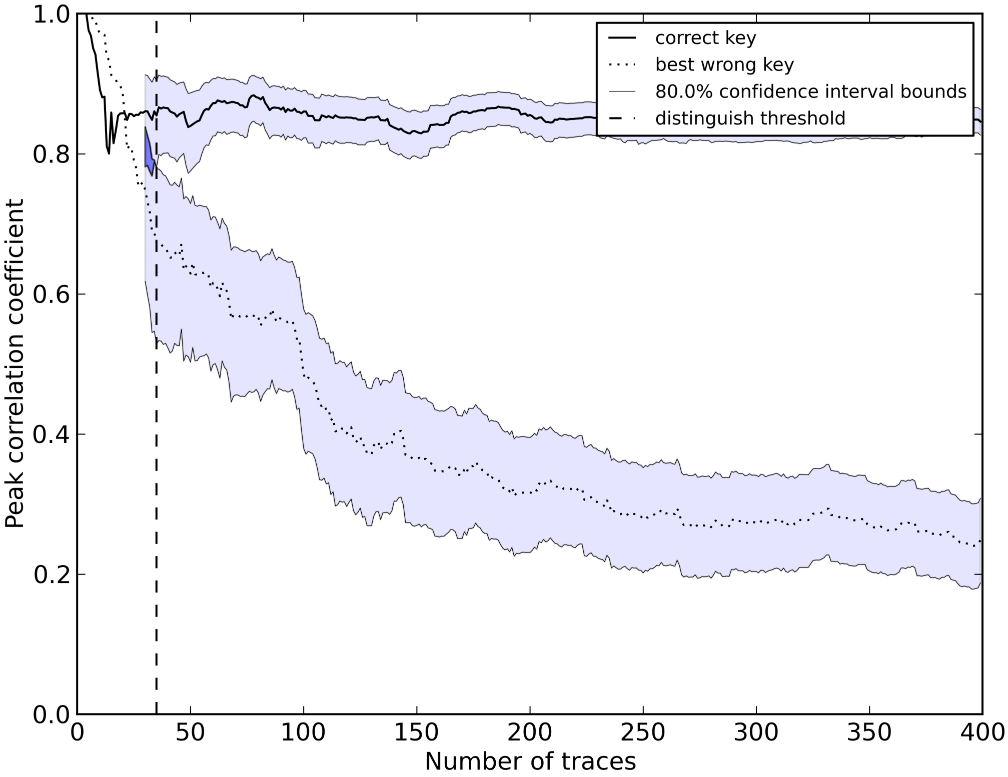

Figure 2: Correlation values and confidence interval for the correct and the best wrong keys on an unprotected implementation of AES 8-bits. The attack is performed on the key byte 0

hypothetical value of the involved key part. The hypothetical key value that provides the maximum correlation values among all but the correct key value is called the *best wrong key*.

We use *confidence intervals* in order to get a quantitative evaluation of the effectiveness of the evaluated protection against DPA. Our aim is to determine an interval around the sample estimation r that contains the correlation coefficient ρ with a probability p [MOP07]. Probability values of 99 % or higher are typically used when computing a confidence interval. In a DPA scenario, we however consider that an attacker does not need such a strong guarantee when trying to distinguish the best key from the correlation traces. Once one or several correlation values distinguish from the others, the attacker will possibly try the related keys on the device under attack. Hence, similarly to [ABP12], we use a probability of 80 % to compute the bounds of confidence intervals.

For an attack to succeed, the confidence interval of the correct key and the confidence intervals of all the other hypothetical keys must separate. We call *distinguish threshold* the number of traces required so that the confidence intervals separate. We measure it as follows: the best correlation values for the correct and the best wrong key are measured from the estimate sample correlations, which respectively correspond to measurement times t_c and t_{bw}. The sample correlation coefficients and their associated confidence intervals are then computed for the correct and best wrong keys at t_c and t_{bw} as a function of the number of sample measurements used to compute the sample correlation. The distinguish threshold corresponds to the minimal value for which the higher bound of the confidence interval for the best wrong key is greater than the lower bound of the confidence interval for the correct key. Figure 2 presents the results of our DPA attack on an unprotected implementation of a 8-bit AES, where the observation traces are measured from an Electro-Magnetic probe. According to our evaluation criterion, the correct key distinguishes from all the other hypothetical key values as soon as 35 traces with a confidence probability above 80 %.

5.2.1 Metrics for system performance

Our setup for code polymorphism using `deGoal` allows to activate several independent sources of polymorphism [COG14]: random register allocation (RA), use of semantic equivalences (SE), insertion of dummy instructions (DI) and instructions shuffling (IS). Let us denote S the state of polymorphic protections activated.

$$\text{state of polymorphism: } S = \{\emptyset, RA, RA+SE, RA+SE+IS, RA+SE+IS+DI\} \tag{1}$$

Code generation interval (ω) We assume that, the more frequently a new polymorphic instance is generated, the higher is the level of security is. Hence we denote ω, the *code generation interval*, which is the ratio between the number of times a new polymorphic instance is generated and the number of executions of all the polymorphic instances (equation 2).

$$\text{code generation interval: } \omega = \frac{\text{nb. executions}}{\text{nb. code generations}} \tag{2}$$

The code generation interval is related to the frequency at which a new polymorphic instance is generated. In terms of security, the best case is achieved when ω equals to 1, and the worst case is achieved when ω is close to 0.

Execution time overhead (k) k denotes the execution time overhead (equation 3). It measures the overhead incurred by the use of code polymorphism as compared to a reference implementation, in terms of execution time. t_{ref}, t_{gen} and t_{exe} respectively denote the execution time of the reference implementation without polymorphism, the execution time of the polymorphic code generator, and the execution time of the polymorphic instance. Considering that we have different parameters that control the level of polymorphism achieved, decribed by S, all the variables k, t_{ref}, t_{gen} and t_{exe} actually depend on S.

$$\text{execution time overhead: } \alpha(S) = \frac{t_{gen}(S) + \omega \times t_{exe}(S)}{\omega \times t_{ref}(S)} \tag{3}$$

In other words, α measures the impact of code polymorphism on execution time for *one* execution of a polymorphic instance as compared to one execution of a reference implementation.

5.3 Metrics of obfuscation by deGoal

The main objective of this case study but also the JVM one is to avoid the reverse by obfuscating the binary code thanks to the embedded compilette. The obfuscation is not directly linked with the software complexity in the sens of Software Engineering (MacCabe, Halstead,...). We need to introduce the concept of resilience which measures how a transformation holds under attack while using an automatic de-obfuscator. The resilience of a transformation is the result of two measures:

- Programmer effort, the amount of time required by the attacker to build its automatic de-obfuscator which should be able to reduce the complexity induced by the compilette,
- De-obfuscator effort, the execution time and space needed the attack program to reduce the complexity of the kernel.

We can say that a transformation is potential if it can confuse a human error but it is considered as resilient if it confuse an automatic tool. For example, some transformation add useless instructions (move r0, r0) so that it does not change its observable behavior (of course the side channel behavior is changed), it just increase the effort for a human to read the code. Theses transformation can be undone with different degree of difficulty using CFG normalization, data flow analysis, *etc*. The resilience in one hand defines the quality of the transformation. We have also to pay attention to the effort needed by the compilette to perform a given transformation (execution cost). For example, an inter-procedural transformation has a high resilience but also a high execution cost in term space and/or memory.

A good obfuscator must have a high resilience with a low execution cost. Of course a trade-off is needed to find the adequate performance of both parameters. In[WW02], the author presents some of well-known transformations. He classifies them into several categories:

- The easiest way of obfuscation is addition of new instructions, covering the view of the real control path. These instructions should in some safe way entangled with obfuscated program, leaving impression of a real program. Any operations can be inserted as far as they do not change the current machine context. Such a transformation has a hight potentiallity but a low resilience. The effect of such transformation can be analyzed with any data flow analyzer.
- Computation transformation, the main approach is to hide the control flow, by extending loop condition which does not change program behaviour, adding dead code controlled with a predicate which value is always the same, remove function call or library call,*etc*.
- Aggregation transformation break up the computation that logically belongs together or merge computation that do not. In-lining is one of the most efficient one but also outlining that turns a sequence of instructions into a subroutine. One of the most efficient transformation against reverse is probably interleaving code where a code fragment performs different activities controlled by an additional parameter.
- Ordering transformation which randomizes the order in which computation are carried out. This transformation has a low potentiality but a very important resilience.
- Creation of pointers to data aliases, the elimination of such constructions is a task computationally very complex and thus has a high resilience.

In this use case, we will evaluate the resilience of the deGoal transformations and evaluate them taking into account the programmer effort and the execution cost.

The analysis of the data depends on the number of collected data. In our case, the number of case study is not enough to use traditional analysis techniques. We will provide the results for the four projects without any statistical treatment.

6 References

[ABP12] Giovanni Agosta, Alessandro Barenghi, and Gerardo Pelosi. A code morphing methodology to automate power analysis countermeasures. In *DAC*, pages 77–82. ACM, 2012.

[CC11] Liang Cai and Hao Chen. Touchlogger: Inferring keystrokes on touch screen from smartphone motion. In *HotSec*, 2011.

[COG14] COGITO. Specification of the tool, first release. public deliverable D1.1.1, 2014.

[MOP07] Stefan Mangard, Elisabeth Oswald, and Thomas Popp. *Power analysis attacks: Revealing the secrets of smart cards*, volume 31. Springer, 2007.

[OSS$^+$] David Oswald, Daehyun Strobel, Falk Schellenberg, Timo Kasper, and Christof Paar. When reverse-engineering meets side-channel analysis – digital lockpicking in practice.

[VWG06] Dennis Vermoen, Marc Witteman, and Georgi N. Gaydadjiev. Reverse engineering of java card applets using power analysis. Technical report, WISTP'2007, LNCS 4462, 2006.

[WW02] Gregory Wroblewski and Gregory Wroblewski. General method of program code obfuscation, 2002.