REASSURE
Project ID: 731591, Horizon 2020
http://www.reassure.eu

REASSURE

Deliverable D5.2

Data management plan

Editor:	F. Koeune (UCL)
Deliverable nature:	R
Dissemination level: (Confidentiality)	PU
Delivery date:	July 14, 2017
Version:	1.1
Total number of pages:	9
Keywords:	side-channel attacks, open access, leakage traces

Horizon 2020
European Union funding
for Research & Innovation

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 731591.

Executive summary

This document represents the first version of the Data Management Plan (DMP) of the REASSURE project. It is a living document that will be updated throughout the project. The document focuses on identifying the type of data that will be openly shared, namely leakage traces acquired and processed during a side-channel analysis of an embedded security device, as well as the intended target audience and the data format.

List of authors

Company	Author
UCL	François Koeune

 ©REASSURE consortium 2017

Revision history

Revision number	Date	Comment
1.0	June 29, 2017	First version of document
1.1	July 14, 2017	Update based on partners' reactions

Contents

 ©REASSURE consortium 2017

1 Introduction

1.1 REASSURE objectives

Implementing cryptography on embedded devices is an ongoing challenge: every year new implementation flaws are discovered and new attack paths are being used by real life adversaries. Whilst cryptography can guarantee many security properties, it crucially depends on the ability to keep the used keys secret even in face of determined adversaries.

Over the last two decades a new type of adversary has emerged, able to obtain, from the cryptographic implementation, side channel leakage such as recording of response times, power or EM signals, etc. To account for such adversaries, sophisticated security certification and evaluation methods (Common Criteria, EMVCo, FIPS...) have been established to give users assurance that security claims have withstood independent evaluation and testing. Recently the reliability of these evaluations has come into the spotlight: the Taiwanese citizen card proved to be insecure, and Snowden's revelations about NSA's tampering with FIPS standards eroded public confidence. REASSURE will:

1. improve the efficiency and quality of all aspects of certification using a novel, structured detect-mapexploit approach that will also improve the comparability of independently conducted evaluations,
2. cater for emerging areas such as the IoT by automating leakage assessment practices in order to allow resistance assessment without immediate access to a testing lab,
3. deliver tools to stakeholders, such as reference data sets and an open-source leakage simulator based on instruction-level profiles for a processor relevant for the IoT,
4. improve existing standards by actively pushing the novel results to standardization bodies.

REASSURE's consortium is ideal to tackle such ambitious tasks. It features two major circuits manufacturers (NXP, MORPHO), a highly respected side channel testing lab (Riscure), an engaged governmental representative (ANSSI), and two of the most prominent research institutions in this field (UCL, University of Bristol).

1.2 Data sharing policy

Very early in the construction of the REASSURE project, it was decided that not all leakage traces would be made available publicly (the consortium explicitly opted out of the open access pilot plan). Indeed, parts of the power traces manipulated by industrial partners correspond to evaluations of their own, or their customer's, product, and these data sets are critical from a security point of view, and also from a customer's confidence point of view. Similarly, some elements, such as the exact identification of the product being evaluated, might not always be fully described in order not to expose company-specific products.

Yet, as discussed in Sec.2.2, it is the conviction of the consortium that largely sharing experimental data is paramount to improve the comparability and quality of evaluations. Consequently, whilst REASSURE will not commit to sharing all experimental data, we will in practice provide as many practically useful data sets as possible.

2 Data summary

2.1 Purpose of data collection/generation

Data generated in the framework of REASSURE mainly consists in data sets related to side-channel analysis, i.e. leakage traces (e.g. power or electromagnetic traces) acquired or simulated when performing side-channel analysis of a given device. This data will then be processed in order to try and expose the cryptographic keys manipulated by the device or, on the contrary, to assess the efficiency of side-channel countermeasures by showing that these keys cannot be recovered.

2.2 Purpose of data sharing

The efficiency of a side-channel attack depends on a large number of factors, including the quality of the physical data (which in turn depends on the measuring equipment, amount of noise, skills and knowledge of the person performing the measurement), the knowledge of the implementation and device's behaviour, the

quality of the exploitation strategy... As a consequence, comparing different attack techniques is very difficult: when a paper describing a new attack technique appears in the literature, it is not always easy to decide whether it yields better results because of its intrinsic quality, or due to more favorable initial data. For the same reason, comparing the efficiency different countermeasures is quite hard to achieve.

By sharing leakage traces at various stages of their processing, REASSURE aims to:

- Allow reproducing experiments, either to validate claims or as a tool facilitating learning.
- Allow a fair comparison of specific substeps of attacks, by making it possible to compare the efficiency of different methods when exploiting *the same data*.
- Provide a common reference basis: in the future, it is our hope to see new results documented by applying them on some openly-accessible reference leakage traces.
- Remove the burden of having to actually run the acquiring of physical leakage traces. This should prove useful for researchers working on side channel evaluation who have a mathematical/algorithmical background (with strong expertise in signal processing, information extraction or statistical evaluation), but little to no knowledge on physical measurements or access to laboratories allowing them to practically generate and collect traces.

2.3 Target audience

REASSURE-related data are expected to be of use for:

- academic researchers;
- device manufacturers;
- implementers of embedded cryptographic algorithms;
- evaluators assessing the side-channel resistance bodies;
- working groups and standardization bodies, such as the JHAS, ANSSI or BSI, to help them decide on the relevance of newly published attacks.

2.4 Types and formats of data

Data include:

1. Raw data acquired when performing physical measurements of a given chip using specific laboratory equipment.
2. Simulated data, generated by running a simulator emulating a chip's behaviour based on a more or less sophisticated model.
3. Post-processed data, generated based on raw or simulated data by applying some post-treatment (e.g. point-of-interest identification, template modeling...). Post-treatment is an iterative process, so various levels of post-processed data may exist.

Raw and simulated data typically consist in a set of one-dimensional (e.g. power consumption) or multi-dimensional (e.g. electromagnetic emanations) information collected per unity of time. Post-processed data may take a similar form, or be the output of a statistical treatment – and, more generally, of any mathematical function – applied to the initial data. Typical examples are reduced data sets after points-of-interest identification or dimensionality reduction, or a probability density function when a template attack is performed.

2.5 Expected size of data

Depending on the sampling frequencies and measurement precision, the size of leakage data can greatly vary, from a couple of megabytes to several gigabytes.

 ©REASSURE consortium 2017

3 Data sharing principles

Data generated by REASSURE must be shared according to the FAIR principle, meaning that it should be findable, accessible, interoperable and re-usable. We describe below the general principles that will be followed to enforce this. These principles will be further refined as the project evolves.

3.1 Sharing policy

As our objective is to allow as widespread use of the data as possible, open licensing schemes similar to the Creative Commons, are being considered. Consortium members, especially academic partners, consider aligning themselves with the University of Bristol "Open Access Software License", which seems well in line with REASSURE's intents. The compliance of this with partner's institutional regulations is being investigated.

3.2 Data localization

The exact storage process to be used by REASSURE is still under evaluation. Most likely will it be a mixture of partner's institutional storage facilities and, especially for very large data pieces, mass storage solutions, e.g. cloud-based. Data replication will be considered to ensure both easy access and disaster recovery.

3.3 Data referencing

Data will be made findable mostly through REASSURE's dissemination activities. Links towards data repositories will be provided on the project's website, as well as in all related scientific publication, position papers and deliverables. In particular, scientific publications will clearly identify the data used to obtain results and, whenever not prohibited by confidentiality reasons, will provide direct access to these data.

3.4 (Meta-)Data format

Each data set will be accompanied by a clear description of

- The device under evaluation.
- The equipment (hardware and/or software) used for data acquisition or generation.
- The treatment applied to the data.
- The identity of the involved team(s).
- The format of the data file.
- Any additional information deemed relevant.

Considering the large number of possible configurations and variability of potentially relevant parameters in the measurement setup, no standard naming convention or metadata format will be used. Instead, the configuration will consist in a text describing the device, equipment and measurement setup in a clear and unambiguous way.

As described in the introduction, some elements, such as the exact identification of the product being evaluated, might not always be fully described in order not expose company-specific products. More generic information, such as the bus size and operating frequency, will then be provided.

3.5 Data lifetime

Nothing in the nature of the data – such as privacy-preserving reasons – makes it necessary to destroy them after a given period of time, hence no action will be taken to delete them nor limit their replication.

The consortium's goal is to maintain data availability for at least three years after project end.

Appendices

A University of Bristol – Open Access Software Licence

Copyright (c) 2016, The University of Bristol, a chartered corporation having Royal Charter number RC000648 and a charity (number X1121) and its place of administration being at Senate House, Tyndall Avenue, Bristol, BS8 1TH, United Kingdom.
All rights reserved
Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution. THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.
Any use of the software for scientific publications or commercial purposes should be reported to the University of Bristol (OSI-notifications@bristol.ac.uk and quote reference REASSURE, H2020 project 731591). This is for impact and usage monitoring purposes only.
Enquiries about further applications and development opportunities are welcome. Please contact elisabeth.oswald@bristol.ac.uk

[end of document]