Activity Report 2017

Team TAMIS

Threat Analysis and Mitigation for Information Security

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Rennes - Bretagne-Atlantique

THEME
Security and Confidentiality

Table of contents

Team TAMIS

Creation of the Team: 2016 January 01, updated into Project-Team: 2018 January 01

Keywords:

Computer Science and Digital Science:

A4. - Security and privacy
A4.1. - Threat analysis
A4.2. - Correcting codes
A4.4. - Security of equipment and software
A4.5. - Formal methods for security
A4.8. - Privacy-enhancing technologies
A4.9. - Security supervision

Other Research Topics and Application Domains:

B6. - IT and telecom
B6.4. - Internet of things
B6.5. - Information systems
B6.6. - Embedded systems
B8.1. - Smart building/home
B8.2. - Connected city
B8.4. - Security and personal assistance
B9.8. - Privacy
B9.9. - Risk management
B9.10. - Ethics

1. Personnel

Research Scientists

Axel Legay [Team leader, Inria, Researcher, HDR]
Christian Grothoff [Inria, Advanced Research Position, until Aug 2017, HDR]
Annelie Heuser [CNRS, Researcher]
Jean-Louis Lanet [Inria, Senior Researcher, HDR]
Olivier Zendra [Inria, Researcher]
Fabrizio Biondi [Centrale-Supelec, Researcher, "Chaire Malware"]
Kim Larsen [Inria, International Chair, Advanced Research Position]

Post-Doctoral Fellows

Najah Ben Said [Inria]
Ludovic Claudepierre [Inria, from Oct 2017]
Ronan Lashermes [Inria, until Oct 2017]
Hélène Le Bouder [Inria, until Mar 2017]
Tania Richmond [Inria, from Sep 2017]

PhD Students

Delphine Beaulaton [UBS Vannes, from September 2017]
Sebanjila Bukasa [Inria]
Mounir Chadli [Algerian Ministry of Defense, until Dec 2017]

Olivier Decourbe [Inria]
Florian Dold [Inria]
Mihai Enescu [Inria, until Sep 2017]
Alexandre Gonzalvez [IMT Atlantique]
Nisrine Jafri [Inria]
Martin Moreau [SECURE-IC SAS, from May 2017]
Ruta Moussaileb [IMT Atlantique, from Oct 2017]
Tristan Ninet [Thales]
Lamine Noureddine [Inria, from Nov 2017]
Leopold Ouairy [Inria, from Oct 2017]
Aurélien Palisse [Inria]
Emmanuel Tacheau [CISCO, from Sep 2017]
Alexander Zhdanov [Inria, from Oct 2017]

Technical staff

Jeffrey Paul Burdges [Inria]
Sébastien Campion [Inria]
Thomas Given-Wilson [Inria]
Bruno Lebon [Inria, from Oct 2017]
Laurent Morin [Inria]
Jean Quilbeuf [Inria]
Sean Sedwards [Inria, until Mar 2017]
Marcello Stanisci [Inria, until Jul 2017]
Gabor Toth [Inria, until Apr 2017]
Louis Marie Traonouez [Inria]

Interns

Nicolas Bellec [Inria, from May 2017 until Jul 2017]
Benjamin Bouguet [Inria, from Feb 2017 until Jun 2017]
Vincent Redoute [Centrale-Supélec, from Jun 2017 until Aug 2017]
Vivek Verma [Inria, from Jun 2017 until Aug 2017]

Administrative Assistant

Cecile Bouton [Inria]

External Collaborators

Francois-Renaud Escriva [DGA]
Sebastien Josse [DGA]
Colas Le Guernic [DGA]

2. Overall Objectives

2.1. Context

Security devices are subject to drastic security requirements and certification processes. They must be protected against potentially complex exploits that result from the combination of software and hardware attacks. As a result, a major effort is needed to develop new research techniques and approaches to characterize security issues, as well as to discover multi-layered security vulnerabilities in complex systems.

In recent years, we have witnessed two main lines of research to achieve this objective.

The first approach, often called *offensive security*, relies on engineering techniques and consists in attacking the system with our knowledge on its design and our past expertise. This is a creative approach that supports (1) checking whether a system is subject to existing vulnerabilities, i.e. classes of vulnerabilities that we already discovered on other systems, and (2) discovering new types of vulnerabilities that were not foreseen and that may depend on new technologies and/or programming paradigms. Unfortunately, this approach is limited to systems whose complexity remains manageable at the human level. This means that exploits which combine several vulnerabilities may be hard to identify. The second and more formal approach builds on formal models (also known as *formal methods*) to automatically detect vulnerabilities, or prove their absence. This is applicable to systems whose complexity is beyond human reasoning, but can only detect existing classes of vulnerabilities, i.e., those that have been previously characterized by offensive security.

2.2. Approach and motivation

The claim made by TAMIS is that *assessing security requires combining both engineering and formal techniques*.

As an example, security exploits may require combining classes of well-known vulnerabilities. The detection of such vulnerabilities can be made via formal approaches, but their successful combination requires human creativity. TAMIS's central goal is thus to demonstrably narrow the gap between the vulnerabilities found using formal verification and the issues found using systems engineering. As a second example, we point out that there are classes of attacks that exploit both the software and hardware parts of a system. Although vulnerabilities can be detected via formal methods in the software part, the impact of attacking the hardware still needs to be modeled. This is often done by observing the effect of parameter changes on the system, and capturing a model of them. To address this situation, the TAMIS team bundled resources from scalable formal verification and secure software engineering for *vulnerability analysis*, which we extend to provide methods and tools to (a) *analyze (binary) code including obfuscated malware*, and (b) *build secure systems*.

Very concrete examples better illustrate the differences and complementarity of engineering and formal techniques. First, it is well-known that formal methods can be used to detect buffer overflows. However, the definition of buffer overflows itself was made first in 1972 when the Computer Security Technology Planning study laid out the technique and claimed that over sizing could be exploited to corrupt a system. This exploit was then popularized in 1988 as one of the exploits used by the Morris worm, and only at that point systematic techniques were developed to detect it. Another example is the work we conducted in attacking smart cards. The very firsts experiments were done at the engineering level, and consisted of retrieving the key of the card in a brute force manner. Based on this knowledge, we generated user test-cases that characterize what should not happen. Later, those were used in a fully automatized model-based testing approach [66].

3. Research Program

3.1. Axis 1: Vulnerability analysis

This axis proposes different techniques to discover vulnerabilities in systems. The outcomes of this axis are (a) new techniques to discover system vulnerabilities as well as to analyze them, and (b) to understand the importance of the hardware support.

Most existing approaches used at the engineering level rely on testing and fuzzing. Such techniques consist in simulating the system for various input values, and then checking that the result conforms to a given standard. The problem being the large set of inputs to be potentially tested. Existing solutions propose to extract significant sets by mutating a finite set of inputs. Other solutions, especially concolic testing developed at Microsoft, propose to exploit symbolic executions to extract constraints on new values. We build on those existing work, and extend them with recent techniques based on dissimilarity distances and learning. We also account for the execution environment, and study techniques based on the combination of timing attacks with fuzzing techniques to discover and classify classes of behavior of the system under test.

Techniques such as model checking and static analysis have been used for verifying several types of requirements such as safety and reliability. Recently, several works have attempted to adapt model checking to the detection of security issues. It has clearly been identified that this required to work at the level of binary code. Applying formal techniques to such code requires the development of disassembly techniques to obtain a semantically well-defined model. One of the biggest issues faced with formal analysis is the state space explosion problem. This problem is amplified in our context as representations of data (such as stack content) definitively blow up the state space. We propose to use statistical model checking (SMC) of rare events to efficiently identify problematic behaviors.

We also seek to understand vulnerabilities at the architecture and hardware levels. Particularly, we evaluate vulnerabilities of the interfaces and how an adversary could use them to get access to core assets in the system. One particular mechanism to be investigated is the DMA and the so-called Trustzone. An ad-hoc technique to defend against adversarial DMA-access to memory is to keep key material exclusively in registers. This implies co-analyzing machine code and an accurate hardware model.

3.2. Axis 2: Malware analysis

Axis 1 is concerned with vulnerabilities. Such vulnerabilities can be exploited by an attacker in order to introduce malicious behaviors in a system. Another method to identify vulnerabilities is to analyze malware that exploits them. However, modern malware has a wide variety of analysis avoidance techniques. In particular, attackers obfuscate the code leading to a security exploit. For doing so, recent black hat research suggests hiding constants in program choices via polynomials. Such techniques hinder forensic analysis by making detailed analysis labor intensive and time consuming. The objective of research axis 2 is to obtain a full tool chain for malware analysis starting from (a) the observability of the malware via deobfuscation, and (b) the analysis of the resulting binary file. A complementary objective is to understand how hardware attacks can be exploited by malwares.

We first investigate obfuscation techniques. Several solutions exist to mitigate the packer problem. As an example, we try to reverse the packer and remove the environment evaluation in such a way that it performs the same actions and outputs the resulting binary for further analysis. There is a wide range of techniques to obfuscate malware, which includes flattening and virtualization. We will produce a taxonomy of both techniques and tools. We will first give a particular focus to control flow obfuscation via mixed Boolean algebra, which is highly deployed for malware obfuscation. We recently showed that a subset of them can be broken via SAT-solving and synthesis. Then, we will expand our research to other obfuscation techniques.

Once the malware code has been unpacked/deobfuscated, the resulting binary still needs to be fully understood. Advanced malware often contains multiple stages, multiple exploits and may unpack additional features based on its environment. Ensuring that one understands all interesting execution paths of a malware sample is related to enumerating all of the possible execution paths when checking a system for vulnerabilities. The main difference is that in one case we are interested in finding vulnerabilities and in the other in finding exploitative behavior that may mutate. Still, some of the techniques of Axis 1 can be helpful in analyzing malware. The main challenge for axis 2 is thus to adapt the tools and techniques to deal with binary programs as inputs, as well as the logic used to specify malware behavior, including behavior with potentially rare occurrences. Another challenge is to take mutation into account, which we plan to do by exploiting mining algorithms.

Most recent attacks against hardware are based on fault injection which dynamically modifies the semantics of the code. We demonstrated the possibility to obfuscate code using constraint solver in such a way that the code becomes intentionally hostile while hit by a laser beam. This new form of obfuscation opens a new challenge for secure devices where malicious programs can be designed and uploaded that defeat comprehensive static analysis tools or code reviews, due to their multi-semantic nature. We have shown on several products that such an attack cannot be mitigated with the current defenses embedded in Java cards. In this research, we first aim at extending the work on fault injection, then at developing new techniques to analyze such hostile code. This is done by proposing formal models of fault injection, and then reusing results from our work on obfuscation/deobfuscation.

3.3. Axis 3: Building a secure network stack

To evaluate the techniques developed in Axes 1 and 2, we analyze concrete systems developed not only with industry partners, but also within the team. By using our own systems, we can co-evolve best-practices, while externally developed systems provide realistic challenges especially with respect to analyzing obfuscated malware in the hardware or complex vulnerabilities. In this context, Christian Grothoff (ARP Inria) is currently developing a new Internet, which is supposed to be more secure. This introduces interesting challenges both in terms of vulnerability and malware analysis, and hence should be a great opportunity to mix the competences of all the members of the team.

More precisely, this system intends to challenge the idea that network security is an administrative task, where network administrators shield users with passwords, firewalls, intrusion detection systems and policies. Instead, we want to eliminate administrators that have power over user's data, and as such administrators themselves are liabilities, and because a network design that permits administrative intrusion inherently adds vulnerabilities. Instead, the system should ensure secure communication mechanisms without trusted third parties.

Key challenges we work on include (a) improving scalable secure ad-hoc decentralized routing, including key-value lookup, unicast and multicast communication, (b) protecting meta-data in the overlay using advanced decentralized onion routing, (c) a unified public-key infrastructure and identity management solution that is suitable to replace the Web-of-Trust, X.509, DNSSEC and other legacy methods for naming and identifying services, (d) secure synchronous and asynchronous messaging at scale, providing decentralized alternatives to common online social applications and addressing challenges in protocol evolution and compatibility. Finally, we are currently working on GNU Taler, a new secure privacy-preserving payment system where users never have to authenticate. This system in particular can be used as a concrete test case for the methods developed in the team.

To support this research work, we develop a framework named GNUnet. It provides a clear separation into layers, which facilitates testing and verifying the various components. However, we see that often existing formal verification techniques still do not scale to typical subsystems encountered in practice. Our objective is thus to exploit efficient and scalable formal techniques techniques proposed in Axis 1 together with engineering skills in order to guide the validation (message synchronization, data protection, ...) and reach the best compromise. An additional complication is that we need a validation process that not merely covers the software itself, but also all of its dependencies (such as database, cryptographic libraries and networking libraries). For the Taler-specific hardware, we are envisioning an NFC-powered device, which creates new challenges in terms of securing cryptographic computations in a setting where the adversary has control over the power supply. In such a case, the attacker can drive the environment and modify the behavior of the system as we have shown in Axis 2. Providing the control of the environment is a new vector for attackers.

Christian Grothoff, who leads this axis, got a position in Bern in 2017. This axis is expected to follow him in the future, although Tamis still holds expertise and members to finish ongoing work with the team. Cooperations with Bern are expected in the future.

4. Application Domains

4.1. System analysis

The work performed in Axes 1 and 2 and the methods developed there are applicable to the domain of system analysis, both wrt. program analysis and hardware analysis.

4.2. Cybersecurity

The work done in the 3 axes above aims at improving cybersecurity, be it via vulnerability analyses, malware analyses and the development of safer networking mechanisms.

4.3. Safe Internet

The work done in Axis 3 above very directly contributes to the goal of a safer Internet.

5. Highlights of the Year

5.1. Highlights of the Year

"Chaire Analyse de Menaces" (Threat Analysis)

Participants: Axel Legay, Fabrizio Biondi

Creation of the "Chaire Analyse de Menaces" (Threat Analysis), that has been assigned to Fabrizio Biondi.

Thales Air Operations partnership

Participants: Axel Legay, Louis-Marie Traonouez

Creation of a partnership with Thales Air Operations for machine learning algorithms to detect anomalies in ground-to-air communications.

6. New Software and Platforms

6.1. GNUnet

SCIENTIFIC DESCRIPTION: The GNUnet project seeks to answer the question what a modern Internet architecture should look like for a society that care about security and privacy. We are considering all layers of the existing well-known Internet, but are also providing new and higher-level abstractions (such as voting protocols, Byzantine consensus, etc.) that are today solved in application-specific ways. Research questions include the desired functionality of the overall stack, protocol design for the various layers as well as implementation considerations, i.e. how to implement the design securely.
FUNCTIONAL DESCRIPTION: GNUnet is a framework for secure peer-to-peer networking that does not use any centralized or otherwise trusted services. Our high-level goal is to provide a strong free software foundation for a global network that provides security and in particular respects privacy.

GNUnet started with an idea for anonymous censorship-resistant file-sharing, but has grown to incorporate other applications as well as many generic building blocks for secure networking applications. In particular, GNUnet now includes the GNU Name System, a privacy-preserving, decentralized public key infrastructure.

- Participants: Alvaro Garcia Recuero, Florian Dold, Gabor Toth, Hans Grothoff, Jeffrey Paul Burdges and Sree Hrsha Totakura
- Partner: The GNU Project
- Contact: Hans Grothoff
- URL: https://gnunet.org/

6.2. MHD

GNU libmicrohttpd
KEYWORDS: Embedded - Web 2.0
SCIENTIFIC DESCRIPTION: We are providing a standards compliant and complete implementation of the HTTP server protocol that allows developers to easily write correct HTTP servers. Key challenges include code size minimization (for IoT devices), performance (zero copy, scalability to 100k concurrent connections), portability and security. MHD is already widely used in production by both academic and industrial users. Ongoing research challenges include formal verification.

FUNCTIONAL DESCRIPTION: GNU libmicrohttpd is a small C library that is supposed to make it easy to run an HTTP server as part of another application.

- Participants: Evgeny Grin, Hans Grothoff and Sree Hrsha Totakura
- Partner: The GNU Project
- Contact: Hans Grothoff
- URL: http://www.gnu.org/software/libmicrohttpd/

6.3. PLASMA Lab

KEYWORDS: Energy - Statistics - Security - Runtime Analysis - Model Checker - Statistical - Model Checking - Aeronautics - Distributed systems
SCIENTIFIC DESCRIPTION: Statistical model checking (SMC) is a fast emerging technology for industrial scale verification and optimisation problems. SMC only requires an executable semantics and is not constrained by decidability. Therefore we can easily apply it to different modelling languages and logics. We have implemented in PLASMA Lab several advanced SMC algorithms that combine formal methods with statistical tests, which include techniques for rare events estimation and non-deterministic models.
FUNCTIONAL DESCRIPTION: PLASMA Lab is a compact, efficient and flexible platform for statistical model checking of stochastic models. PLASMA Lab includes simulators for PRISM models (Reactives Modules Language-RML) and Biological models. It also provides plugins that interface external simulators in order to support Matlab/Simulink, SytemC and LLVM . PLASMA Lab can be extended with new plugins to support other external simulators, and PLASMA Lab API can be used to embed the tool in other softwares. PLASMA Lab provide fast SMC algorithms, including advanced techniques for rare events simulation and nondeterministic models. These algorithms are designed in a distributed architecture to run large number of simulations on several computers, either on a local area network or grid. PLASMA Lab is implemented in Java with efficient data structures and low memory consumption
NEWS OF THE YEAR: In 2017 we have extended PLASMA Lab with a new simulator plugin that allows to verify LLVM code.

- Participants: Axel Legay, Jean Quilbeuf, Benoît Boyer, Kevin Corre, Louis-Marie Traonouez, Matthieu Simonin and Sean Sedwards
- Contact: Axel Legay
- URL: https://project.inria.fr/plasma-lab/

6.4. Taler

GNU Taler
KEYWORD: Privacy
SCIENTIFIC DESCRIPTION: Taler is a Chaum-style digital payment system that enables anonymous payments while ensuring that entities that receive payments are auditable. In Taler, customers can never defraud anyone, merchants can only fail to deliver the merchandise to the customer, and payment service providers can be fully audited. All parties receive cryptographic evidence for all transactions, still, each party only receives the minimum information required to execute transactions. Enforcement of honest behavior is timely, and is at least as strict as with legacy credit card payment systems that do not provide for privacy.

The key technical contribution underpinning Taler is a new refresh protocol which allows fractional payments and refunds while maintaining untraceability of the customer and unlinkability of transactions. The refresh protocol combines an efficient cut-and-choose mechanism with a link step to ensure that refreshing is not abused for transactional payments.

We argue that Taler provides a secure digital currency for modern liberal societies as it is a flexible, libre and efficient protocol and adequately balances the state's need for monetary control with the citizen's needs for private economic activity.

FUNCTIONAL DESCRIPTION: Taler is a new electronic payment system. It includes an electronic wallet for customers, a payment backend for merchants and the main payment service provider logic called the exchange. Taler offers Chaum-style anonymous payments for citizens, and income-transparency for taxability.

- Participants: Florian Dold, Gabor Toth, Hans Grothoff, Jeffrey Paul Burdges and Marcello Stanisci
- Partner: The GNU Project
- Contact: Hans Grothoff
- URL: http://taler.net/

6.5. HyLeak

Hybrid Analysis Tool for Information Leakage
KEYWORD: Information leakage
FUNCTIONAL DESCRIPTION: HyLeak is an evolution of the QUAIL tool, also developed by the TAMIS team. HyLeak divides the input program into (terminal) components and decides for each of them whether to analyze it using precise or statistical analysis, by applying heuristics that evaluate the analysis cost of each component. Then, HyLeak composes the analysis results of all components into an approximate joint probability distribution of the secret and observable variables in the program. Finally, the tool estimates the Shannon leakage and its confidence interval.

- Partner: AIST Tsukuba
- Contact: Fabrizio Biondi

6.6. SimFI

Tool for Simulation Fault injection
KEYWORDS: Fault injection - Fault-tolerance
FUNCTIONAL DESCRIPTION: Fault injections are used to test the robust and security of systems. We have developed SimFI, a tool that can be used to simulate fault injection attacks against binary files. SimFI is lightweight utility designed to be integrated into larger environments as part of robustness testing and fault injection vulnerability detection.

- Contact: Nisrine Jafri
- URL: https://github.com/nisrine/Fault-Injection-Tool

6.7. DaD

Data-aware Defense
KEYWORD: Ransomware
FUNCTIONAL DESCRIPTION: DaD is a ransomware countermeasure based on a file system minifilter driver. It is a proof of concept and in its present condition cannot be used as a replacement of the existing antivirus solutions. DaD detects randomness of the data by monitoring the write operations on the file system. We monitor all the userland threads, and also the whole file system (i.e., not restricted to Documents). It blocks the threads that exceed a specific threshold. The malicious thread is not killed, we only block its next I/O operations.

- Contact: Aurélien Palisse

6.8. MASSE

Modular Automated Syntactic Signature Extraction
KEYWORDS: Malware - Syntactic analysis

FUNCTIONAL DESCRIPTION: The Modular Automated Syntactic Signature Extraction (MASSE) architecture is a new integrated open source client-server architecture for syntactic malware detection and analysis based on the YARA, developed with Teclib'. MASSE includes highly effective automated syntactic malware detection rule generation for the clients based on a server-side modular malware detection system. Multiple techniques are used to make MASSE effective at detecting malware while keeping it from disrupting users and hindering reverse-engineering of its malware analysis by malware creators. MASSE integrates YARA in a distributed system able to detect malware on endpoint systems using YARA, analyze malware with multiple analysis techniques, automatically generate syntactic malware detection rules, and deploy the new rules to the endpoints. The MASSE architecture is freely available to companies and institutions as a complete, modular, self-maintained antivirus solution. Using MASSE, a security department can immediately update the rule database of the whole company, stopping an infection on its tracks and preventing future ones.

- Contact: Axel Legay

6.9. Behavioral Malware Analysis

KEYWORDS: Artificial intelligence - Malware - Automatic Learning - Concolic Execution
FUNCTIONAL DESCRIPTION: Our approach is based on artificial intelligence. We extract graphs from programs, that represent their behaviors. Such graphs are called system call dependency graphs (SCDGs). Our software learns to distinguish malware from cleanware on a large set of malwares and cleanwares. Whenever we want to analyze a new program, we extract its graphs and use the result of the training to decide whether the new program to analyze is a malware.

- Partner: Cisco
- Contact: Axel Legay
- URL: https://team.inria.fr/tamis/

6.10. VITRAIL - Visualisation Tool

Real-Time, Advanced, Immersive Visualization of Software / Visualizer
KEYWORD: Visualization of software
SCIENTIFIC DESCRIPTION: It is difficult for developers to explore and understand the source code of large programs, for example in objet-oriented languages programs featuring thousands of classes. Visualization methods based on daily life metaphors have thus been proposed. The VITRAIL Visualization tool (or VITRAIL Vizualizer) makes it possible to display, visualize and explore Java programs in a metaphorical way, using the city metaphor. An execution trace of the Java (byte)code provided by VITRAIL JBInstrace tool, is provided as input to VITRAIL Visualizer which displays a city-like metaphorical world showing the static structure of the code as well as some dynamic elements (calls).
FUNCTIONAL DESCRIPTION: This program makes it possible to displays, visualizes and explores Java programs in a metaphorical way (using the city metaphore). Useful for complex application developers/architects.

RELEASE FUNCTIONAL DESCRIPTION: Early release

- Participants: Damien Bodenes, Olivier Demengeon and Olivier Zendra
- Contact: Olivier Zendra
- URL: http://vitrail.loria.fr

6.11. VITRAIL 6 JBInsTrace

Real-Time, Advanced, Immersive Visualization of Software / Java Bytecode Instrumenter and Tracer
KEYWORDS: Execution trace - Profiling - Instrumentation - Bytecode - Java - Basic block

SCIENTIFIC DESCRIPTION: VITRAIL JBInsTrace is a program to instrument Java bytecode to trace its execution. The trace contains both static and dynamic information (calls). It is produced by intercepting the JVM class loader and replacing it by ours. Thus Java bytecode file are not modified, since instrumentation is performed on the fly, in memory. This makes it possible to instrument the whole program code, including libraries. Java source code is not needed. The trace which is then fed into our program VITRAIL Visualizer is an XML-like file.

FUNCTIONAL DESCRIPTION: VITRAIL JBInsTrace is a program to instrument Java bytecode files to trace their execution. The trace is then fed into our VITRAIL Visualizer tool.

- Participants: Olivier Zendra and Pierre Caserta
- Contact: Olivier Zendra
- URL: http://vitrail.loria.fr

6.12. Platforms

6.12.1. Malware'o'Matic

This LHS platform is dedicated to the collect, the categorization and the analyze of malware. We are currently interested in a specific kind of malware the ransomware. The platform grabs periodically samples of public data bases, executes the ransomware without virtualization on a victim PC and evaluate the implemented detection mechanisms. Once a ransomware has been executed the image of the OS is automatically restored and a new sample is evaluated. The platform is fully automatic and target Windows platforms (seven, W10) in both 32 bits and 64 bits versions. More recent developments can be seen in the LHS Activity Report.

6.12.2. Faustine

This LHS platform is dedicated to the EM fault injection experiments. It is composed of a motion table (XY), a pulse generator, an amplifier and a control PC. It injects EM pulses in a controlled way on a targeted device using an EM probe. It controls with a high precision the timing and the edges of the pulse. A recent development consists in adding a FPGA board to control the trigger in a more convenient and precise way. Then, the pulse can be triggered while a specific information is sent to the board under attack. More recent developments can be seen in the LHS Activity Report.

7. New Results

7.1. Results for Axis 1: Vulnerability analysis

7.1.1. Statistical Model Checking of LLVM Code

Participants: Axel Legay, Louis-Marie Traonouez.

We have extended PLASMA Lab statistical model-checker with a new plugin that allows to simulate LLVM bitcode. The plugin is based on an external simulator LODIN. This simulator implements a probabilistic semantics for a LLVM program. At its core the semantics consist of the LLVM program given as a labelled transition system. The labels are function calls to an environment that implements functions outside the LLVM core language. The environment is also responsible for assigning probabilities to individual transitions

By interfacing the LODIN simulator with PLASMA Lab we can apply all the statistical model-checking algorithms provided by PLASMA Lab, including rare events verification algorithms like importance splitting. We have applied LODIN and PLASMA Lab to several case studies, including the analysis of some security vulnerability, like the PTrace privilege escalation attack that could be performed on earlier versions of the Linux Kernel. This work has been submitted to a conference this year [61], and is currently under review.

[61] We present our work in providing Statistical Model Checking for programs in LLVM bitcode. As part of this work we develop a semantics for programs that separates the program itself from its environment. The program interact with the environment through function calls. The environment is furthermore allowed to perform actions that alter the state of the C-program-useful for mimicking an interrupt system. On top of this semantics we build a probabilistic semantics and present an algorithm for simulating traces under that semantics. This paper also includes the development of the new tool component Lodin that provides a statistical model checking infrastructure for LLVM programs. The tool currently implement standard Monte Carlo algorithms and a simulator component to manually inspect the behaviour of programs. The simulator also proves useful in one of our other main contributions; namely producing the first tool capable of doing importance splitting on LLVM code. Importance splitting is implemented by integrating Lodin with the existing statistical model checking tool Plasma-Lab.

7.1.2. Verification of IKEv2 protocol

Participants: Axel Legay, Tristan Ninet, Louis-Marie Traonouez, Olivier Zendra.

The IKEv2 (Internet Key Exchange version 2) protocol is the authenticated key-exchange protocol used to set up secure communications in an IPsec (Internet Protocol security) architecture. It guarantees security properties like mutual-authentication and secrecy of the exchanged key. To obtain an IKEv2 implementation as secure as possible, we use model checking to verify the properties on the protocol specification, and smart fuzzing to test the implementation, and try to detect implementation flaws like buffer overflows or memory leaks.

Two weaknesses had previously been found in the specification, but were harmless. We showed that the first weakness does not actually exist. We demonstrated that the second weakness is not harmless, and we designed a Denial-of-Service attack that exploits it, the deviation attack. As a counter-measure, we propose a modification of IKEv2, and use model checking to prove that the modified version is secure.

This work is being prepared for responsive disclosure and publication.

7.1.3. High-Level Frameworks for Scheduling Systems

Participants: Mounir Chadli, Axel Legay, Louis-Marie Traonouez.

Formal model-based techniques are more and more used for the specification and verification of scheduling systems. These techniques allow to consider complex scheduling policies beyond the scope of classical analytical techniques. For instance, hierarchical scheduling systems (HSS) integrates a number of components into a single system running on one execution platform. Hierarchical scheduling systems have been gaining more attention by automotive and aircraft manufacturers because they are practical in minimizing the cost and energy of operating applications. Model-based techniques can also be used to solve new problems like energy optimization or runtime monitoring. However, one limitation of formal model-based approaches is that they require high technical knowledge about the formalims and tools used to design models and write properties.

In a previous work [62], we have presented a model-based framework for the verification of HSS. It is based on a stochastic extension of timed automata and statistical model checking with the tool UPPAAL. We have also developed a graphical high-level language to represent complex hierarchical scheduling systems. To bridge the gap between the formalisms, we exploit Cinco, a generator for domain specific modeling tools to generate an interface between this language and the one of UPPAAL. Cinco allows to specify the features of a graphical interface in a compact meta-model language. This is a flexible approach that could be extended to any formal model of scheduling problem.

We have extended the previous work in journal paper [55] published this year, where we provide another high-level framework for the verification of energy-aware scheduling systems. We also present two new analysis techniques. One that performs runtime monitoring in order to detect alarming change in the scheduling system, and one that performs energy optimization.

> [55] Over the years, schedulability of Cyber-Physical Systems (CPS) has mainly been performed by analytical methods. These techniques are known to be effective but limited to a few classes of scheduling policies. In a series of recent work, we have shown that schedulability analysis of CPS could be performed with a model-based approach and extensions of verification tools such as UPPAAL. One of our main contributions has been to show that such models are flexible enough to embed various types of scheduling policies, which goes beyond those in the scope of analytical tools.
>
> However, the specification of scheduling problems with model-based approaches requires a substantial modeling effort, and a deep understanding of the techniques employed in order to understand their results. In this paper we propose simplicity-driven high-level specification and verification frameworks for various scheduling problems. These frameworks consist of graphical and user-friendly languages for describing scheduling problems. The high-level specifications are then automatically translated to formal models, and results are transformed back into the comprehensible model view. To construct these frameworks we exploit a meta-modeling approach based on the tool generator Cinco.
>
> Additionally we propose in this paper two new techniques for scheduling analysis. The first performs runtime monitoring using the CUSUM algorithm to detect alarming change in the system. The second performs optimization using efficient statistical techniques. We illustrate our frameworks and techniques on two case studies.

7.1.4. Side-channel Analysis of Cryptographic Substitution Boxes

Participants: Axel Legay, Annelie Heuser.

With the advent of the Internet of Things, we are surrounded with smart objects (aka things) that have the ability to communicate with each other and with centralized resources. The two most common and widely noticed artefacts are RFID and Wireless Sensor Networks which are used in supply-chain management, logistics, home automation, surveillance, traffic control, medical monitoring, and many more. Most of these applications have the need for cryptographic secure components which inspired research on cryptographic algorithms for constrained devices. Accordingly, lightweight cryptography has been an active research area over the last 10 years. A number of innovative ciphers have been proposed in order to optimize various performance criteria and have been subject to many comparisons. Lately, the resistance against side-channel attacks has been considered as an additional decision factor.

Side-channel attacks analyze physical leakage that is unintentionally emitted during cryptographic operations in a device (e.g., power consumption, electromagnetic emanation). This side-channel leakage is statistically dependent on intermediate processed values involving the secret key, which makes it possible to retrieve the secret from the measured data.

Side-channel analysis (SCA) for lightweight ciphers is of particular interest not only because of the apparent lack of research so far, but also because of the interesting properties of substitution boxes (S-boxes). Since the nonlinearity property for S-boxes usually used in lightweight ciphers (i.e., 4×4) can be maximally equal to 4, the difference between the input and the output of an S-box is much smaller than for instance for AES. Therefore, one could conclude that from that aspect, SCA for lightweight ciphers must be more difficult. However, the number of possible classes (e.g., Hamming weight (HW) or key classes) is significantly lower, which may indicate that SCA must be easier than for standard ciphers. Besides the difference in the number of classes and consequently probabilities of correct classification, there is also a huge time and space complexity advantage (for the attacker) when dealing with lightweight ciphers.

In [65], [64] we give a detailed study of lightweight ciphers in terms of side-channel resistance, in particular for software implementations. As a point of exploitation we concentrate on the non-linear operation (S-box) during the first round. Our comparison includes SPN ciphers with 4-bit S-boxes such as KLEIN, PRESENT, PRIDE, RECTANGLE, Mysterion as well as ciphers with 8-bit S-boxes: AES, Zorro, Robin. Furthermore, using simulated data for various signal-to-noise ratios (SNR) we present empirical results for Correlation Power Analysis (CPA) and discuss the difference between attacking 4-bit and 8-bit S-boxes.

An extension of this work is given in [10]. We investigate whether side-channel analysis is easier for lightweight ciphers than e.g. for AES. We cover both profiled and non-profiled techniques where we are interested in recovering secret (round)keys or intermediate states. In the case of non-profiled attacks, we evaluate a number of S-boxes appearing in lightweight ciphers using the confusion coefficient and empirical simulations.

First, we investigate in the scenario where the attacker targets the first round and thus exploits the S-box computation. We observe that the 8-bit S-boxes from AES, Zorro, and Robin perform similarly, whereas for 4-bit S-boxes we have a clear ranking, with the S-box of Piccolo being the weakest to attack and the S-box of KLEIN and Midori (1) the hardest. Interestingly, when considering the last round and thus the inverse S-box operation the ranking changes such that Mysterion is the weakest and PRESENT/LED is the most side-channel resistant cipher from the ones investigated. Moreover, we could observe that attacking the last round is equal or less efficient for all considered ciphers. Finally, we use the information gained from both rounds together, where this approach is of interest when the cipher does not use round keys from a key scheduling algorithm but rather uses the same (or a straightforward computable) key in each round. LED fulfils this requirement. For a reasonable low SNR, to reach a success rate of 0.9 an attack on both rounds only requires 100 traces, whereas an attack using the first round requires 200 traces and on the last 400 traces. This example highlights the important role the confusion coefficient (relationship between predicted intermediate states under a leakage model from different key hypotheses), and that not only the SNR (even if low) is a key factor influencing the success rate. Additionally, our result show that we cannot conclude that the 4-bit S-boxes are generally significantly less resistant than the investigated 8-bit S-boxes. In particular, when considering inverse S-boxes we showed that 4-bit S-boxes may be more resistant.

For profiled attacks, we analyze several machine learning techniques to recover 4-bit and 8-bit intermediate states. Our results show that attacking 4-bit is somewhat easier than attacking 8-bit, with the difference mainly stemming from the varying number of classes in one or the other scenario. Still, that difference is not so apparent as one could imagine. Since we work with only a single feature and yet obtain a good accuracy in a number of test scenarios, we are confident (as our experiments also confirm) that adding more features will render classification algorithms even more powerful, which will result in an even higher accuracy. Finally, we did not consider any countermeasures for the considered lightweight algorithms, since the capacity for adding countermeasures is highly dependent on the environment (which we assume to be much more constrained than in the case of AES). However, our results show that a smart selection of S-boxes results in an inherent resilience (especially for 4-bit S-boxes). Moreover, we show that in case of highly restricted devices, in which countermeasures on the whole cipher are not practically feasible, a designer may choose to only protect the weakest round (first round) in the cipher to increase the side-channel resistant until a certain limit.

Our work in [23] concentrates on how to improve SCA resilience of ciphers without imposing any extra cost. This is possible by considering the inherent resilience of ciphers. We particularly concentrate on block ciphers which utilize S-boxes and therefore study the resilience of S-boxes against side-channel attacks. When discussing how to improve side-channel resilience of a cipher, an obvious direction is to use various masking or hiding countermeasures. However, such schemes come with a cost, e.g. an increase in the area and/or reduction of the speed. When considering lightweight cryptography and various constrained environments, the situation becomes even more difficult due to numerous implementation restrictions. However, some options are possible like using S-boxes that are easier to mask or (more on a fundamental level), using S-boxes that possess higher inherent side-channel resilience. In [23] we investigate what properties should an S-box possess in order to be more resilient against side-channel attacks. Moreover, we find certain connections between those properties and cryptographic properties like nonlinearity and differential uniformity. Finally, to strengthen our theoretical findings, we give an extensive experimental validation of our results.

[64] Side-channel Analysis of Lightweight Ciphers: Current Status and Future Directions

[65] Side-channel Analysis of Lightweight Ciphers: Does Lightweight Equal Easy?

[10] Lightweight Ciphers and their Side-channel Resilience.

[23] Trade-Offs for S-Boxes: Cryptographic Properties and Side-Channel Resilience

[24] Do we need a holistic approach for the design of secure IoT systems? hal-01628683

7.1.5. New Advances on Side-channel Distinguishers

Participants: Axel Legay, Annelie Heuser.

[16] *Template Attack vs Bayes Classifier*

Side-channel attacks represent one of the most powerful category of attacks on cryptographic devices with profiled attacks in a prominent place as the most powerful among them. Indeed, for instance, template attack is a well-known real-world attack that is also the most powerful attack from the information theoretic perspective. On the other hand, machine learning techniques have proven their quality in a numerous applications where one is definitely side-channel analysis. As one could expect, most of the research concerning supervised machine learn- ing and side-channel analysis concentrated on more powerful machine learning techniques. Although valid from the practical perspective, such attacks often remain lacking from the more theoretical side. In this paper, we investigate several Bayes classifiers, which present simple supervised techniques that have significant similarities with the template attack. More specifically, our analysis aims to investigate what is the influence of the feature (in)dependence in datasets with different amount of noise and to offer further insight into the efficiency of machine learning for side-channel analysis.

[46] *Side-channel analysis and machine learning: A practical perspective* The field of side-channel analysis has made significant progress over time. Analyses are now used in practice in design companies as well as in test laboratories, and the security of products against side-channel attacks has significantly improved. However, there are still some remaining issues to be solved for analyses to be more effective. Side-channel analysis ac- tually consists of two steps, commonly referred to as identification and exploitation. The identification consists of understanding the leakage in order to set up a relevant attack. On the other hand, the exploitation consists of using the identified leakages to extract the secret key. In scenarios where the model is poorly known, it can be approximated in a profiling phase. There, machine learning techniques are gaining value. In this paper, we conduct extensive analysis of several machine learning techniques, showing the importance of proper parameter tuning and training. In contrast to what is perceived as common knowledge in unrestricted scenarios, we show that some machine learning techniques can significantly outperform template attack when properly used. We therefore stress that the traditional worst case security assessment of cryptographic implementations that includes mainly template attacks might not be accurate enough. Besides that, we present a new measure called the Data Confusion Factor that can be used to assess how well machine learning techniques will perform on a certain dataset.

[30] *Codes for Side-Channel Attacks and Protections*

This article revisits side-channel analysis from the standpoint of coding theory. On the one hand, the attacker is shown to apply an optimal decoding algorithm in order to recover the secret key from the analysis of the side-channel. On the other hand, the side-channel protections are presented as a coding problem where the information is mixed with randomness to weaken as much as possible the sensitive information leaked into the side-channel. Therefore, the field of side-channel analysis is viewed as a struggle between a coder and a decoder. In this paper, we focus on the main results obtained through this analysis. In terms of attacks, we discuss optimal strategy in various practical contexts, such as type of noise, dimensionality of the leakage and of the model, etc. Regarding countermeasures, we give a formal analysis of some masking schemes.

[38] *Climbing Down the Hierarchy: Hierarchical Classification for Machine Learning Side-Channel Attacks*

Machine learning techniques represent a powerful paradigm in side-channel analysis, but they come with a price. Selecting the appropriate algorithm as well as the parameters can sometimes be a difficult task. Nevertheless, the results obtained usually justify such an effort. However, a large part of those results use simplification of the data relation and in fact do not consider allthe available information. In this paper, we analyze the hierarchical relation between the data and propose a novel hierarchical classification approach for side-channel analysis. With this technique, we are able to introduce two new attacks for machine learning side-channel analysis: Hierarchical attack and Structured attack. Our results show that both attacks can outperform machine learning techniques using the traditional approach as well as the template attack regarding accuracy. To support our claims, we give extensive experimental results and discuss the necessary conditions to conduct such attacks.

[14] *Stochastic Collision Attack*

On the one hand, collision attacks have been introduced in the context of side-channel analysis for attackers who exploit repeated code with the same data without having any knowledge of the leakage model. On the other hand, stochastic attacks have been introduced to recover leakage models of internally processed intermediate secret variables. Both techniques have shown advantages and intrinsic limitations. Most collision attacks, for instance, fail in exploiting all the leakages (e.g., only a subset of matching samples are analyzed), whereas stochastic attacks cannot involve linear regression with the full basis (while the latter basis is the most informative one). In this paper, we present an innovative attacking approach, which combines the flavors of stochastic and collision attacks. Importantly, our attack is derived from the optimal distinguisher, which maximizes the success rate when the model is known. Notably, we develop an original closed-form expression, which shows many benefits by using the full algebraic description of the leakage model. Using simulated data, we show in the unprotected case that, for low noise, the stochastic collision attack is superior to the state of the art, whereas asymptotically and thus, for higher noise, it becomes equivalent to the correlation-enhanced collision attack. Our so-called stochastic collision attack is extended to the scenario where the implementation is protected by masking. In this case, our new stochastic collision attack is more efficient in all scenarios and, remarkably, tends to the optimal distinguisher. We confirm the practicability of the stochastic collision attack thanks to experiments against a public data set (DPA contest v4). Furthermore, we derive the stochastic collision attack in case of zero-offset leakage that occurs in protected hardware implementations and use simulated data for comparison. Eventually, we underline the capability of the new distinguisher to improve its efficiency when the attack multiplicity increases.

[15] *Optimal side-channel attacks for multivariate leakages and multiple models*

Side-channel attacks allow to extract secret keys from embedded systems like smartcards or smartphones. In practice, the side-channel signal is measured as a trace consisting of several samples. Also, several sensitive bits are manipulated in parallel, each leaking differently. Therefore, the informed attacker needs to devise side-channel distinguishers that can handle both multivariate leakages and multiple models. In the state of the art, these two issues have two independent solutions: on the one hand, dimensionality reduction can cope with multivariate leakage; on the other hand, online stochastic approach can cope with multiple models. In this paper, we combine both solutions to derive closed-form expressions of the resulting optimal distinguisher in terms of matrix operations, in all situations where the model can be either profiled offline or regressed online. Optimality here means that the success rate is maximized for a given number of traces. We recover known results for uni- and bivariate models (including correlation power analysis) and investigate novel distinguishers for multiple models with more than two parameters. In addition, following ideas from the AsiaCrypt?2013 paper ?Behind the Scene of Side-Channel Attacks,? we provide fast computation algorithms in which the traces are accumulated prior to computing the distinguisher values.

[39] *Stochastic Side-Channel Leakage Analysis via Orthonormal Decomposition*

Side-channel attacks of maximal efficiency require an accurate knowledge of the leakage function. Template attacks have been introduced by Chari et al. at CHES 2002 to estimate the leakage function using available training data. Schindler et al. noticed at CHES 2005 that the complexity of profiling could be alleviated if the evaluator has some prior knowledge on the leakage function. The initial idea of Schindler is that an engineer can model the leakage from the structure of the circuit. However, for some thin CMOS technologies or some advanced countermeasures, the engineer intuition might not be sufficient. Therefore, inferring the leakage function based on profiling is still important. In the state-of-the-art, though, the profiling stage is conducted based on a linear regression in a non-orthonormal basis. This does not allow for an easy interpretation because the components are not independent. In this paper, we present a method to characterize the leakage based on a Walsh-Hadamard orthonormal basis with staggered degrees, which allows for direct interpretations in terms of bits interactions. A straightforward application is the characterization of a class of devices in order to understand their leakage structure. Such information is precious for designers and also for evaluators, who can devise attack bases relevantly.

[17] *On the optimality and practicability of mutual information analysis in some scenarios*

The best possible side-channel attack maximizes the success rate and would correspond to a maximum likelihood (ML) distinguisher if the leakage probabilities were totally known or accurately estimated in a profiling phase. When profiling is unavailable, however, it is not clear whether Mutual Information Analysis (MIA), Correlation Power Analysis (CPA), or Linear Regression Analysis (LRA) would be the most successful in a given scenario. In this paper, we show that MIA coincides with the maximum likelihood expression when leakage probabilities are replaced by online estimated probabilities. Moreover, we show that the calculation of MIA is lighter that the computation of the maximum likelihood. We then exhibit two case-studies where MIA outperforms CPA. One case is when the leakage model is known but the noise is not Gaussian. The second case is when the leakage model is partially unknown and the noise is Gaussian. In the latter scenario MIA is more efficient than LRA of any order.

[59] *On the Relevance of Feature Selection for Profiled Side-channel Attacks*

In the process of profiled side-channel analysis there is a number of steps one needs to make. One important step that is often conducted without a proper attention is selection of the points of interest (features) within the side-channel measurement trace. Most of the related work start with an assumption that the features are selected and various attacks are then considered and compared to find the best approach. In this paper, we concentrate on the feature selection step and show that if a proper selection is done, most of the attack techniques offer satisfactory results. We investigate how more advanced feature selection techniques stemming from the machine learning domain can be used to improve the side-channel attack efficiency. Our results show that the so-called Hybrid feature selection methods result in the best classification accuracy over a wide range of test scenarios and number of features selected.

[60] *Profiled SCA with a New Twist: Semi-supervised Learning*

Profiled side-channel attacks represent the most powerful category of side-channel attacks. In this context, the attacker gains ac- cess of a profiling device to build a precise model which is used to attack another device in the attacking phase. Mostly, it is assumed that the attacker has unlimited capabilities in the profiling phase, whereas the attacking phase is very restricted. We step away from this assumption and consider an attacker who is restricted in the profiling phase, while the attacking phase is less limited as in the traditional view. Clearly, in general, the attacker is not hindered to exchange any available knowledge between the profiling and attacking phase. Accordingly, we propose the concept of semi-supervised learning to side-channel analysis, in which the attacker uses the small amount of labeled measurements from the profiling phase as well as the unlabeled measurements from the attacking phase to build a more reliable model. Our results show that semi-supervised learning is beneficial in many scenarios and of particular interest when using template attack and its pooled version as side-channel attack techniques. Besides stating our results in varying scenarios, we discuss more general conclusions on semi-supervised learning for SCA that should help to transfer our observations to other settings in SCA.

7.1.6. Side-channel analysis on post-quantum cryptography

Participants: Axel Legay, Annelie Heuser, Tania Richmond, Martin Moreau.

In recent years, there has been a substantial amount of research on quantum computers ? machines that exploit quantum mechanical phenomena to solve mathematical problems that are difficult or intractable for conventional computers. If large-scale quantum computers are ever built, they will be able to break many of the public-key cryptosystems currently in use. This would seriously compromise the confidentiality and integrity of digital communications on the Internet and elsewhere. The goal of post-quantum cryptography (also called quantum-resistant cryptography) is to develop cryptographic systems that are secure against both quantum and classical computers, and can interoperate with existing communications protocols and networks. At present, there are several post-quantum cryptosystems that have been proposed: lattice-based, code-based, multivariate cryptosystems, hash-based signatures, and others. However, for most of these proposals, further research is needed in order to gain more confidence in their security and to improve their performance. Our interest lies in particular on the side-channel analysis and resistance of these post-quantum schemes. We first focus on code-based cryptography and then extend our analysis to find common vulnerabilities between different families of post-quantum crypto systems.

7.1.7. Binary Code Analysis: Formal Methods for Fault Injection Vulnerability Detection

Participants: Axel Legay, Thomas Given-Wilson, Annelie Heuser, Nisrine Jafri, Jean-Louis Lanet.

Formal methods such as model checking provide a powerful tool for checking the behaviour of a system. By checking the properties that define correct system behaviour, a system can be determined to be correct (or not).

Increasingly fault injection is being used as both a method to attack a system by a malicious attacker, and to evaluate the dependability of the system. By finding fault injection vulnerabilities in a system, the resistance to attacks or faults can be detected and subsequently addressed.

A process is presented that allows for the automated simulation of fault injections. This process proceeds by taking the executable binary for the system to be tested, and validating the properties that represent correct system behaviour using model checking. A fault is then injected into the executable binary to produce a mutant binary, and the mutant binary is model checked also. A different result to the validation of the executable binary in the checking of the mutant binary indicates a fault injection vulnerability.

This process has been automated with existing tools, allowing for easy checking of many different fault injection attacks and detection of fault injection vulnerabilities. This allows for the detection of fault injection vulnerabilities to be fully automated, and broad coverage of the system to be formally shown.

The work is implemented in the SimFi tool.

[56] (J; submitted) Fault injection has increasingly been used both to attack software applications, and to test system robustness. Detecting fault injection vulnerabilities has been approached with a variety of different but limited methods. This paper proposes an extension of a recently published general model checking based process to detect fault injection vulnerabilities in binaries. This new extension makes the general process scalable to real-world implementations which is demonstrated by detecting vulnerabilities in different cryptographic implementations.

7.1.8. Security at the hardware and software boundaries

Participants: Axel Legay, Jean-Louis Lanet, Ronan Lashermes, Kevin Bukasa, Hélène Le Bouder.

7.1.8.1. Side-channel attacks (SCA)

SCA exploit the reification of a computation through its physical dimensions (current consumption, EM emission, etc.). Focusing on Electromagnetic Analyses (EMA), such analyses have mostly been considered on low-end devices: smartcards and micro-controllers. In the wake of recent works, we analyze the effects of a modern micro architecture [31] on the efficiency of EMA (here Correlation Power Analysis and template attacks). We show that despite the difficulty to synchronize the measurements, the speed of the targeted core and the activity of other cores on the same chip can still be accommodated. Finally, we confirm that enabling the secure mode of TrustZone (a hardware-assisted software countermeasure) has no effect whatsoever on the EMA efficiency. Therefore, critical applications in TrustZone are not more secure than in the normal world with respect to EMA, in accordance with the fact that it is not a countermeasure against physical attacks. We hint that such techniques may be more common in the future to overcome the true difficulty with high-end devices: dealing with time precision (problem even worse with an OS or a virtual machine). Here again TrustZone or the activity of other cores have no incidence. But with these attacks, managing the big amount of data generated by our measures may prove to be the limiting factor, requiring better computing resources.

We investigate the way the compiler works and new attack paths have been discovered. In particular we demonstrated experimentally on an ARM7m the possibility to execute arbitrary code, generate buffer overflow even in presence of compiler assisted canary and ROP attacks. This raises a new challenge: any code fragment of an embedded program is sensitive to a fault attack. Thus an attacker increases the success rate of its attack while targeting a non sensitive part of the program for the injection. Then it becomes easy to extract security materials from the device. Then, the verification of the absence of a potential vulnerability must be checked on the whole program and not only on the cryptographic primitives. Thus the prevention analysis that was possible thanks to formal methods becomes unreachable with these new attack paths [40].

7.1.8.2. SCA based fuzzer

One of the main challenges during the development of system is to give a proof of evidence that its functionalities are correctly implemented and that no vulnerability remains. This objective is mostly achieved via testing techniques, which include software testing to check whether a system meets its functionalities, or security testing to express what should not happen. For the latter case, fuzzing is considered as first class citizen. It consists in exercising the system with (randomly) generated and eventually modified inputs in order to test its resistance. While fuzzing is definitively the fastest and the easiest way for testing applications, it suffers from severe limitations. Indeed, the precision of the model used for input generation: a random and/or simple model cannot reach all states and significant values. Moreover, a higher model precision can result in a combinatorial explosion of test cases.

We suggest a new approach [11] whose main ingredient is to combine timing attacks with fuzzing techniques. This new approach, allows not only reducing the test space explosion, but also to simplify the fuzzing process configuration. This new testing scenario is based on observing several executions of the system and by freezing some of its parameters in order to establish a partial order on their timing evaluation. The root of our technique is to exploit timing information to classify the input data into sub-domains according to the behavior observed for specific values of the parameters. Our approach is able to discover hidden unspecified commands that may trigger computations in the tested software. Due to the specific nature of the application (the domain of the parameters is the byte) and its programming model we can also retrieve the control flow graph of the application. The limits of the approach have been identified, and it has been tested on two applications. Validation via a coverage tool has been established.

7.1.9. System Vulnerability Analysis

Participants: Jean-Louis Lanet, Abdelhal Mesbah, Razika Lounas, Chaharezd Yayaoui.

We present in this section our effort to detect and correct some misbehaviors encountered with some firmware. We start with an attack on a secure device, such that we are able to reverse a code while the ISA is unknown and the code itself is not available. Then, we propose a formal specification of the update process of a firmware which provides the guarantee that the updated program respects the semantics of the language. In a last aspect, we try to predict the ability of a program to be attacked thanks to a Machine Learning algorithm. We demonstrated in section 7.1.8 that a state exploration is useless until the whole program is examined, we demonstrated here that approximative solutions can deal with real live programs with an affordable response time.

7.1.9.1. Reverse engineering

We believe that an adversary can gain access to different assets of the system using a black box approach.This implies of course the absence of the source code, but also sometime the absence of the binary code (romized within the soc or micro-controller, no update mechanism, no jtag, no memory extraction, no read function, and so on). In that case, the first step consists in extracting the binary code from the system. The attacker is just allowed to load data. He has then to infer enough information on the system internals and then he should be able to gain access to the native layers. In [43], we demonstrate the advantage of a graphical representation of the data in the memory can help the reverse process thanks to the abstraction provided. Our graphical tool links all the objects with a relationship based on the presence of a pointer.

In a Java based secure element, a Java application is considered as data executed by the executed program (the virtual machine) by the native processor. We introduce a first weakness in the program that allows to read an instance as an array which violate the Java type system. This weakness allows us to dump a short part of the memory which contains the meta data on a set of arrays. Thanks to this information, we generate a mimicry attack by forging pointer illegally [41]. In turns, it open the possibility to read large part of the memory as element of a forged array. Then we succeed in characterizing the memory management algorithm [12]. At the end, we transform the initial problem of finding a vulnerability in the code of a device in a black box approach to a white box problem after de-assembling the binary code.

In another work [44], we studied the byte code verification process towards an unchecked code. We found that this verification is not complete and can be bypassed. The verifier checks the semantics of the Java Card byte code. This process is split in two parts. First, the verifier loads the methods' byte code and checks the package content. For the method segment, it checks that the control flow remain inside the methods, the jump destinations are correct and so on. Secondly, for each entry point and only for these, it controls the semantics and the type correctness of the code. This step is not performed for unreachable code, while the specification states that no unreachable code should remain in the file. However, during our analysis we discovered that the verifier does some verification on the semantics of the unreachable code. Then, thanks to a fault attack (the return byte code is noped) we diverted the control flow into this unchecked area were we stored our ill-typed code leading to the execution of an aggressive shell code which in turn dumped the native layers of the card giving access to the secret key material in plain text.

7.1.9.2. Safe system update mechanism

Dynamic Software Updating (DSU) consists in updating running programs on the fly without any downtime. This feature is interesting in critical applications that must run continuously. Because updates may lead to security breaches, the question of their correctness is raised. Formal methods are a rigorous means to ensure the correctness required by applications using DSU. We propose [13] a formal verification of correctness of DSU in a Java-based embedded system. Our approach is based on three steps. First, a formal interpretation of the semantics of update operations to ensure type safety of the update. Secondly, we rely on a functional representation of byte code, the predicate transformation calculus, and a functional model of the update mechanism to ensure the behavioral correctness of the updated programs. It is based on the use of Hoare predicate transformation to derive a specification of an updated byte code. In the last step, we use the functional representation to model the safe update point detection mechanism. This mechanism guarantees that none of the updated method active methods are active. This property is called activeness safety. We propose a functional specification that allows to derive proof obligations that guarantee the safety of the mechanism.

7.1.9.3. Prediction of system divergence

Fault attack represents one of the serious threats against embedded system security. The result of the fault injection could lead to a mutation of the code in such a way that it becomes hostile or execute a unwanted sequence of code as we demonstrated in 7.1.8. Any successful attack may reveal a secret information stored in the card or grant an undesired authorization. We propose a methodology [5] to recognize, during the development step, the sensitive patterns to the fault attack. It is based on the concepts from text categorization and machine learning. In fact, in this method we represented the patterns using opcodes n-grams as features and we evaluated different machine learning classifiers.

In the first experiment, we evaluated all the combination of n-gram size (for n=2, n=3 and n=4), number of features using GR method to select 100, 200, ..., 500 and 1000 relevant n-grams, n-gram weighting (Term Frequency (TF), Term Frequency Inverse Document Frequency (TFIDF) and binary representations), and five classification algorithms (Naive Bayes network (NB), Decision Tree (DT), Support Vector Machine (SVM), and the boosted version of these two lasts (BDT and BSVM)) to determine the best setting. We used accuracy measure to evaluate performance of the classifiers. In addition to accuracy, we used F1, TP rate and FP rate measures to evaluate how the algorithms classified the dangerous patterns. In the first experiment, we noted that 2-gram outperformed others. Nearly 2-gram, TFIDF, 1000 features with boosted algorithm outperformed the other settings. The F1 results have shown that the classifiers are more accurate at classifying examples of the class of non dangerous pattern compared to other classes. We suggest that this might be due to the imbalance of our data set. In the second experiment, we investigated the imbalance problem. We applied SMOTE and NCR resampling techniques to overcome this class imbalance problem. We found that the outperforming setting in the resampled data set was St_{270} also with BSVM classifier. Resampled data set improves accuracy of the smallest class and keeps the accuracy of other classes.

The experimental results indicated that the resampling techniques improved the accuracy of the classifiers. In addition, our proposed method reduces the execution time of sensitive patterns classification in comparison to the mutant generator tool micro seconds instead of hours.

7.2. Results for Axis 2: Malware analysis

The detection of malicious programs is a fundamental step to be able to guarantee system security. Programs that exhibit malicious behavior, or *malware*, are commonly used in all sort of cyberattacks. They can be used to gain remote access on a system, spy on its users, exfiltrate and modify data, execute denial of services attacks, etc.

Significant efforts are being undertaken by software and data companies and researchers to protect systems, locate infections, and reverse damage inflicted by malware. Our contribution to malware analysis include the following fields:

7.2.1. Malware Detection

Participants: Axel Legay, Fabrizio Biondi, Olivier Decourbe, Mike Enescu, Thomas Given-Wilson, Annelie Heuser, Jean-Louis Lanet, Jean Quilbeuf, Alexander Zhdanov, Olivier Zendra.

Given a file or data stream, the malware detection problem consists of understanding if the file or data stream contain traces of malicious behavior. For binary executable files in particular, this requires extracting a signature of the file, so it can be compared against signatures of known clean and malicious files to determine whether the file is malicious. Binary file signatures can be divided in *syntactic* and *semantic*.

Syntactic signatures are based on properties of the file itself, like its length, hash, number and entropy of the executable and data sections, and so on. While syntactic signatures are computationally cheap to extract from binaries, it is also easy for malware creators to deploy *obfuscation* techniques that change the file's syntactic properties, hence widely mutating the signature and preventing its use for malware detection.

Semantic signatures instead are based on the binary's behavior and interactions with the system, hence are more effective at characterizing malicious files. However, they are more expensive to extract, requiring behavioral analysis and reverse-engineering of the binary. Since behavior is much harder to change than syntactic properties, against these signatures obfuscation is used to harden the file against reverse-engineering and preventing the analysis of the behavior, instead of changing it directly.

In both cases, *malware deofbuscation* is necessary to extract signatures containing actuable information that can be used to characterize the binaries as clean or malicious. Once the signatures are available, *malware classification* techniques, usually based on machine learning, are used to automatically determine whether binaries are clean or malicious starting from their signatures. Our contributions on these fields are described in the next sections.

7.2.2. Malware Deobfuscation

Participants: Axel Legay, Fabrizio Biondi, Olivier Decourbe, Mike Enescu, Thomas Given-Wilson, Annelie Heuser, Nisrine Jafri, Jean-Louis Lanet, Jean Quilbeuf.

Given a file (usually a portable executable binary or a document supporting script macros), deobfuscation refers to the preparation of the file for the purposes of further analysis. Obfuscation techniques are specifically developed by malware creators to hinder detection reverse engineering of malicious behavior. Some of these techniques include:

Packing Packing refers to the transformation of the malware code in a compressed version to be dynamically decompressed into memory and executed from there at runtime. Packing techniques are particularly effective against static analysis, since it is very difficult to determine statically the content of the unpacked memory to be executed, particularly if packing is used multiple times. The compressed code can also be encrypted, with the key being generated in a different part of the code and used by the unpacking procedure, or even transmitted remotely from a command and control (C&C) server.

Control Flow Flattening This technique aims to hinder the reconstruction of the control flow of the malware. The malware's operation are divided into basic blocks, and a dispatcher function is created that calls the blocks in the correct order to execute the malicious behavior. Each block after its execution returns control to the dispatcher, so the control flow is flattened to two levels: the dispatcher above and all the basic blocks below.

To prevent reverse engineering of the dispatcher, it is often implemented with a cryptographic hash function. A more advanced variant of this techniques embed a full virtual machine with a randomly generated instruction set, a virtual program counted, and a virtual stack in the code, and uses the machine's interpreter as the dispatcher.

Virtualization is a very effective technique to prevent reverse engineering. To contrast it, we are implementing state-of-the-art devirtualization algorithms in `angr` , allowing it to detect and ignore the virtual machine code and retrieving the obfuscated program logic. Again, we plan to contribute our improvements to the main `angr` branch, thus helping the whole security community fighting virtualized malware.

Opaque Constants and Conditionals Reversing packing and control flow flattening techniques requires understanding of the constants and conditionals in the program, hence many techniques are deployed to obfuscate them and make them unreadable by reverse engineering techniques. Such techniques are used e.g. to obfuscate the decryption keys of packed encrypted code and the conditionals in the control flow.

We have proven the efficiency of dynamic synthesis in retrieving opaque constant and conditionals, compared to the state-of-the-art approach of using SMT (Satisfiability Modulo Theories) solvers, when the input space of the opaque function is small enough. We are developing techniques based on fragmenting and analyzing by brute force the input space of opaque conditionals, and SMT constraints in general, to be integrated in SMT solvers to improve their effectiveness.

7.2.3. *Malware Classification*

Participants: Axel Legay, Fabrizio Biondi, Olivier Decourbe, Mike Enescu, Thomas Given-Wilson, Annelie Heuser, Nisrine Jafri, Jean-Louis Lanet, Jean Quilbeuf.

Once malicious behavior has been located, it is essential to be able to classify the malware in its specific family to know how to disinfect the system and reverse the damage inflicted on it.

While it is rare to find an actually previously unknown malware, morphic techniques are employed by malware creators to ensure that different generations of the same malware behave differently enough than it is hard to recognize them as belonging to the same family. In particular, techniques based on the syntax of the program fails against morphic malware, since syntax can be easily changed.

To this end, semantic signatures are used to classify malware in the appropriate family. Semantic signatures capture the malware's behavior, and are thus resistant to morphic and differentiation techniques that modify the malware's syntactic signatures. We are investigating semantic signatures based on the program's System Call Dependency Graph (SCDG), which have been proven to be effective and compact enough to be used in practice. SCDGs are often extracted using a technique based on pushdown automata that is ineffective against obfuscated code; instead, we are applying concolic analysis via the `angr` engine to improve speed and coverage of the extraction.

Once a semantic signature has been extracted, it has to be compared against large database of known signatures representing the various malware families to classify it. The most efficient way to obtain this is to use a supervised machine learning classifier. In this approach, the classifier is trained with a large sample of signatures malware annotated with the appropriate information about the malware families, so that it can learn to quickly and automatically classify signatures in the appropriate family. Our work on machine learning classification focuses on using SCDGs as signatures. Since SCDGs are graphs, we are investigating and adapting algorithms for the machine learning classification of graphs, usually based on measures of shared subgraphs between different graphs. One of our analysis techniques relies on common subgraph extraction, with the idea that a malicious behavior characteristic of a malware family will yield a set of common subgraphs. Another approach relies on the Weisfeiler-Lehman graph kernel which uses the presence of nodes and their neighborhoods pattern to evaluate similarity between graphs. The presence or not of a given pattern becomes a feature in a subsequent machine learning analysis through random forest or SVM.

In malware detection and classification, it is fundamental to have a false positive rate (i.e. rate of cleanware classified as malware) approaching zero, otherwise the classification system will classify hundred or thousands of cleanware files as malware, making it useless in practice. To decrease the false positive rate, the classifier is also trained with a large and representative database of cleanware, so that it can discriminate between signatures of cleanware and malware with a minimal false positive rate. We use a large database of malware and cleanware to train our classifier, thus guaranteeing a high detection rate with a small false positive rate.

We have put in place a platform for malware analysis, using dedicated hardware provided by Cisco. This platform is now fully operational and receives a daily feed of suspicious binaries for analysis. Furthermore, we developed tools for maintaining our datasets of cleanware and malware binaries, run existing syntactic analysis on them. Our toolchain is able to extract SCDGs from malwares and cleanwares and apply our classification techniques on the SCDGs.

7.2.4. Botnet Trojan Detection

Participants: Axel Legay, Fabrizio Biondi, Vesselin Bontchev, Thomas Given-Wilson, Jean Quilbeuf, Olivier Decourbe, Najah Ben Said.

Botnet trojans are a class of malware that opens a backdoor in a system and waits from further instructions from a C&C server, and possibly replicates itself somehow. A large group of systems infected by such malware is known as a botnet, and can be used by the botnet's controller to distribute spam emails (possibly carrying other malware) and perform distributed denial-of-service (DDoS) attacks. In a DDoS attack, all the systems in the botnet flood a single target with requests amounting to gigabytes or even terabytes of traffic. The target is not able to handle such traffic or to discriminate malicious request from legitimate ones, failing to provide its service.

Detecting and correctly classify botnet trojans in transit is a necessary step to be able to stop their infection. We applied our semantic classification approach on a particular family of malware, the Mirai botnet. With these experiments, we were able to confirm that the classification based on SCDG extraction and common subgraphs mining has a very low false positive rate and a high detection rate. Furthermore, our approach proved to be more accurate than detection based on syntactic signatures, without increasing the number of false positives.

7.2.5. Modular Automated Syntactic Signature Extraction (MASSE)

Participants: Axel Legay, Fabrizio Biondi, Olivier Zendra, Alexander Zhdanov, Bruno Lebon, François Déchelle.

Malware detection techniques based on syntactic signatures (or "rules") are commonly used in antivirus since their low computational cost allows them to be used on scan the files handled by the system without excessively slowing down the system. Semantic analysis techniques are relatively expensive to use, and would slow down a system significantly if used for on-access malware detection. Hence, it is common in antivirus company to use advanced semantic techniques like the SCDG-based ones we develop to detect and analyze known and unknown malware samples, and then to manually write a syntactic rule for the detection of such samples that is uploaded to the client machines.

The MASSE projects aims at providing an open-source, self-contained architecture to deploy this on a given system, company, or infrastructure, without needing to give access to the structure's data to third parties. The architecture is composed of a server executing the computationally-expensive semantic analysis, and of a number of lightweight clients performing inexpensive syntactic analysis on the client's systems. The MASSE server automatically analyzes unknown or suspicious files passing on the clients, detects the malicious ones, synthesizes syntactic signatures for them, and updates the signature databases of the clients, keeping them protected.

The MASSE server exploits modular malware analysis, supporting malware analysis modules using dynamic, static, or hybrid analysis; extracting syntactic, semantic, or hybrid signatures; using signature-based or anomaly-based detection; and any other technique the user desires, thanks to its open source malware analysis APIs. MASSE also implements pseudonymization of the signature databases, preventing an attacker to learn precisely the syntactic signatures in case some of the clients are compromised.

7.2.6. Malware IDS

Participants: Jean-Louis Lanet, Aurélien Palisse, Colas Le Guernic.

7.2.6.1. An efficient IDS for malware detection

Ransomware is a type of malware that prevents legitimate users from accessing their machine or files and demands a payment for restoring the functionalities of the infected computer. There are two classes of ransomware: the *simple lockers*, which block the usage of the computer, and *cryptors*, that encrypt files on the computer. In the case of encryption-based ransomware, the user data can only be restored with the secret key(s) used during the attack if the key is provided by the attacker.

Detecting a malware can use two options:

- The system knows the features of the malware. Features can be structural information: n-gram or graph isomorphism, or behavioral information: APIs call or system calls. Exact pattern matching algorithm or approximative algorithm (Machine learning) can be used. This approach is known as signature based and can only detect known patterns.
- The system knows its correct behavior. Any deviation of this model leads to the detection of hostile programs. This approach can detect any new attack, it does not rely on a model of the bad behavior but on the model of the correct behavior. This approach is also known as IDS (Intrusion Detection System).

In [45], [34] we apply this technique to detect malware at run time (EPS: End Point Solution). Our first solution is based on the dynamic analysis of the data transformation by the program. We propose to monitor file activity. Since it has already been proven a valid approach in terms of detection, our main goal in is to show that a good detection rate can be achieved with little to no impact on system performances. To this end, we limit our monitoring to a minimum. In order to reduce the impact on detection with a low rate of false positive, we use the chi-square goodness-of-fit test instead of Shannon entropy (*i.e.*, sensitive to compressed chunks of data). We also achieve system completeness and fine granularity by monitoring the whole file system for all userland threads. In order to evaluate our prototype implementation, Data Aware Defense (DaD), under realistic conditions, we used the bare-metal analysis platform of the LHS, Malware - O - Matic (MoM), and ran it on a large and heterogeneous (compared to the literature) live ransomware collection. We used *de facto* industry standard benchmarks to get a pertinent and reproducible assessment of the performance penalties. A second model of the correct behavior with better results has been developed (patent pending).

Our countermeasure is efficient and can be deployed on Windows 7/10 machines with a reasonable performance hit, with an average delay of 12 μs per write operation on disk, a few hundred times smaller than previous approaches. Our extensive experiments show that the more sophisticated ransomware already use mimicry attacks. However we successfully detect 99.37 % of the samples with at most 70 MB lost per sample's threads in 90% of cases and less than 7 MB in 70% of cases. Its speed and low negative rate makes it a good candidate as a first line of defense. Once a thread is deemed malicious, instead of blocking disk accesses, other more costly metrics can be used to improve the false positive rate without impacting performance, since it would not be computed for all other threads.

7.2.7. Papers

This section gathers papers that are results common to all sections above pertaining to Axis 2.

[51] (C) The largest DDoS attacks in history have been executed by devices controlled by the Mirai botnet trojan. To prevent Mirai from spreading, this paper presents and evaluates techniques to classify binary samples as Mirai based on their syntactic and semantic properties. Syntactic malware detection is shown to have a good detection rate and no false positives, but to be very easy to circumvent. Semantic malware detection is resistant to simple obfuscation and has better detection rate than syntactic detection, while keeping false positives to zero. This paper demonstrates these results, and concludes by showing how to combine syntactic and semantic analysis techniques for the detection of Mirai.

[19] (C) We present the MASSE architecture, a YARA-based open source client-server malware detection platform. MASSE includes highly effective automated syntactic malware detection rule generation for the clients based on a server-side modular malware detection system. Multiple techniques are used to make MASSE effective at detecting malware while keeping it from disrupting users and hindering reverse-engineering of its malware analysis by malware creators.

[4] (J) Control flow obfuscation techniques can be used to hinder software reverse-engineering. Symbolic analysis can counteract these techniques, but only if they can analyze obfuscated conditional statements. We evaluate the use of dynamic synthesis to complement symbolic analysis in the analysis of obfuscated conditionals. We test this approach on the taint-analysis-resistant Mixed Boolean Arithmetics (MBA) obfuscation method that is commonly used to obfuscate and randomly diversify statements. We experimentally ascertain the practical feasibility of MBA obfuscation. We study using SMT-based approaches with different state-of-the-art SMT solvers to counteract MBA obfuscation, and we show how targeted algebraic simplification can greatly reduce the analysis time. We show that synthesis-based deobfuscation is more effective than current SMT-based deobfuscation algorithms, thus proposing a synthesis-based attacker model to complement existing attacker models.

7.3. Results for Axis 3: Building a secure network stack

7.3.1. Privacy-Preserving Abuse Detection in Future Decentralised Online Social Networks

Participants: Jeffrey Burdges, Alvaro Garcia Recuero, Christian Grothoff.

Future online social networks need to not only protect sensitive data of their users, but also protect them from abusive behavior coming from malicious participants in the network. We investigated the use of supervised learning techniques to detect abusive behavior and describe privacy-preserving protocols to compute the feature set required by abuse classification algorithms in a secure and privacy-preserving way. While our method is not yet fully resilient against a strong adaptive adversary, our evaluation suggests that it will be useful to detect abusive behavior with a minimal impact on privacy.

Our results show how to combine local knowledge with private set intersection and union cardinality protocols (with masking of BLS signature to protect identity of signers/subscribers) to privately derive feature values from users in OSNs. Given an adaptive adversary that would be able to manipulate most features we propose in our supervised learning approach, it is surprising that with just three features resistant to adversarial manipulation, the algorithms still provide useful classifications.

This work was originally presented at DPM 2016 [63] and expanded upon in Álvaro García-Recuero's PhD thesis [1].

7.3.2. Fog of Trust

Participants: Jeffrey Burdges, Christian Grothoff.

The Web of Trust (WoT) used traditionally used by tools for private communication such as PGP is used to to validate individual links between participants. Using the WoT, however, leaks meta data, such that users must opt-in for it – exposing themselves to risks of privacy loss. We proposed a new method, the Fog of Trust (FoT), which uses the privacy-preserving set intersection cardinality protocol originally used in our work on abuse detection in online social networks, to support this critical step of public key verification via collaboration. In the FoT, the social relationships — which are used to verify public keys – remain hidden. This allows keys to be verified via trusted intermediaries that were established beforehand, without the need to verify each individual new contact using Trustwords. Consequently, FoT will can the same functionality as the WoT without its drawbacks to privacy.

7.3.3. Cell tower privacy

Participants: Christian Grothoff, Neal Walfield.

Context-aware applications are programs that are able to improve their performance by adapting to the current conditions, which include the user's behavior, networking conditions, and charging opportunities. In many cases, the user's location is an excellent predictor of the context. Thus, by predicting the user's future location, we can predict the future conditions. In this work, we developed techniques to identify and predict the user's location over the next 24 hours with a minimum median accuracy of 82results include our observation that cell phones sample the towers in their vicinity, which makes cell towers as-is inappropriate for use as landmarks. Motivated by this observation, we developed two techniques for processing the cell tower traces so that landmarks more closely correspond to locations, and cell tower transitions more closely correspond to user movement. We developed a prediction engine, which is based on simple sampling distributions of the form $f(t, c)$, where t is the predicted tower, and c is a set of conditions. The conditions that we considered include the time of the day, the day of the week, the current regime, and the current tower. Our family of algorithms, called TomorrowToday, achieves 89% prediction precision across all prediction trials for predictions 30 minutes in the future. This decreases slowly for predictions further in the future, and levels off for predictions approximately 4 hours in the future, at which point we achieve 82% prediction precision across all prediction trials up to 24 hours in the future. This represents a significant improvement over NextPlace, a well-cited prediction algorithm based on non-linear time series, which achieves appropriately 80% prediction precision (self reported) for predictions 30 minutes in the future, but, unlike our predictors, which try all prediction attempts, NextPlace only attempts 7% of the prediction trials on our data set [67].

7.3.4. Taler protocol improvements

Participants: Jeffrey Burdges, Florian Dold, Christian Grothoff, Marcello Stanisci.

We started modeling the Taler protocol in the framework of Provable Security, precisely defining the formal meaning of income transparency, fairness, anonymity and unforgeablity as security games. The resulting definitions and security proofs allow a more precise statement of the security of Taler in relation to the security assumptions that are being made.

The implementation of the wallet module now supports the full Taler protocol, including the refresh operation for highly efficient and privacy-preserving change.

In addition to improving the stability of the implementation of all Taler components, we added new features to the protocol that (1) allow refunds from merchants without violating privacy and (2) allow merchants to do "customer tipping", which transfers money from merchants directly to customers' wallets as a reward for doing actions on their website.

7.3.5. Mix Networking

Participants: Jeffrey Burdges, Christian Grothoff.

We have begun implementing our ratcheting scheme for providing hybrid post-quantum and forward security to the Sphinx mix network packet format. We also began collaborating with the Panoramix project and LEAP to help resolve numerous practical challenges to deploying a mix network. We shall speak about this ongoing work at the Chaos Computer Club's annual congress 34c3 in December 2017.

7.4. Other research results

7.4.1. Privacy and Security: Information-Theoretical Quantification of Security Properties

Participants: Axel Legay, Fabrizio Biondi, Olivier Zendra, Thomas Given-Wilson, Annelie Heuser, Sean Sedwards, Jean Quilbeuf, Mike Enescu.

Information theory provides a powerful quantitative approach to measuring security and privacy properties of systems. By measuring the *information leakage* of a system security properties can be quantified, validated, or falsified. When security concerns are non-binary, information theoretic measures can quantify exactly how much information is leaked. The knowledge of such information is strategic in the developments of component-based systems.

The quantitative information-theoretical approach to security models the correlation between the secret information of the system and the output that the system produces. Such output can be observed by the attacker, and the attacker tries to infer the value of the secret information by combining this information with their prior knowledge of the system.

Armed with the produced output of the system, the attacker tries to infer information about the secret information that produced the output. The quantitative analysis we consider defines and computes how much information the attacker can expect to infer (typically measured in bits). This expected leakage of bits is the information leakage of the system. This is computed by symbolically exploring the code to be analyzed, and using the symbolic constraints accumulated over the output together with a model counting algorithm to quantify the leakage.

The quantitative approach generalizes the qualitative approach and thus provides superior analysis. In particular, a system respects non-interference if and only if its leakage is equal to zero. In practice very few systems respect non-interference, and for those that don't it is imperative to be able to distinguish between the systems leaking very small amounts of secret information and systems leaking a significant amount of secret information, since only the latter are considered to pose a security vulnerability to the system.

While quantitative leakage computation is a powerful technique to detect security vulnerabilities, computing the leakage of complex programs written in low-level languages is a hard and computationally intensive task. The most common language for low-level implementation of security protocols is C, due to its efficiency, hence much of the effort in developing tools to detect vulnerabilities in source code focus on C. Recently, we have improved the state of the art in leakage quantification from C programs by proposing the usage of approximated model counting instead of precise model counting. We have shown how the approximation can improve the efficiency of leakage quantification by orders of magnitude against a logarithmic decrease in the precision of the result, often producing the same result as precise model counters much faster, and often being able to analyze cases where precise model counters would have failed. We demonstrated this technique by providing the first quantitative leakage analysis of the C code of the Heartbleed bug, showing that our technique can detect the bug in the code.

A different but equally interesting approach is followed by our new HyLEak tool. HyLeak is also able to analyze a system and compute its information leakage, i.e. the amount of information that an observer would gain by about the value of system's secret by observing its output. Contrarily to other techniques, HyLeak can analyze randomized systems, and correctly distinguish between the randomness injected in the system and the uncertainty on the secret value. This allows HyLeak to be used both on systems with explicit randomization and systems that depend on stochastic properties, like cyber-physical systems.

HyLeak uses static code analysis to divide the system to be analyzed in components. For each component, HyLeak evaluates whether it is more convenient to analyze the component using precise or statistical analysis. Each component is analyzed with the most appropriate strategy, and then the results for all components are combined together and information leakage is estimated.

The hybrid approach provides better results than both the precise and the statistical ones in terms of computation time and precision of the result. Also, it bridges the gap between cheap but imprecise statistical techniques and precise but expensive formal techniques, allowing the user to control the required precision of the result according to the computation time they have available. We evaluated HyLeak against QUAIL's precise approach and the statiatical approach implemented in the LeakWatch tool, showing that HyLeak outperforms them both. HyLeak is open source and available at https://project.inria.fr/hyleak/

Applied to shared-key cryptosystems, the information-theoretical approach allows precise reasoning about the information leakage of any secret information in the system including, the key, and the message. Recent work on max-equivocation has generalised perfect secrecy and shown the maximum achievable theoretic bounds for the security of the key and message. Achieving these theoretic maximal bounds has been proven to be achievable by Apollonian Cell Encoders (ACEs). ACEs not only allow the maximum security possible in a shared-key cryptosystem, but also allow for infinite key reuse when the key has less entropy than the message. Further, ACEs are straightforward to construct and have a compact representation making them feasible to use in practice.

Another application is to use information leakage to reason about leakage through shared resources, representing various side-channel attacks. Developmens here allow for the formalising of the leakage model through shared resources, and quantifying how significant the leakage can be. This improves on the state-of-the-art that uses only qualified leakage, and so can be precise about how much is leakage through a shared resource. Such quantification of leakage allows for scheduling of the shared resource to exploit this information to minimise leakage. Such minimisation of leakage allows for scheduling and utilisation of resources that would fail a simple quanlified test, providing solutions when prior state-of-the-art would claim impossibility. Further, a reasoned trade-off can be made between acceptable leakage and utility of the shared resource, allowing solutions that are acceptable even if not perfect.

[53] (C; submitted) Preserving privacy of private communication against an attacker is a fundamental concern of computer science security. Unconditional encryption considers the case where an attacker has unlimited computational power, hence no complexity result can be relied upon for encryption. Optimality criteria are defined for the best possible encryption over a general collection of entropy measures. This paper introduces Apollonian cell encoders, a class of shared-key cryptosystems that are proven to be universally optimal. In addition to the highest possible security for the message, Apollonian cell encoders prove to have perfect secrecy on their key allowing unlimited key reuse. Conditions for the existence of Apollonian cell encoders are presented, as well as a constructive proof. Further, a compact representation of Apollonian cell encoders is presented, allowing for practical implementation.

[18] (C) High-security processes have to load confidential information into shared resources as part of their operation. This confidential information may be leaked (directly or indirectly) to low-security processes via the shared resource. This paper considers leakage from high-security to low-security processes from the perspective of scheduling. The workflow model is here extended to support preemption, security levels, and leakage. Formalization of leakage properties is then built upon this extended model, allowing formal reasoning about the security of schedulers. Several heuristics are presented in the form of compositional preprocessors and postprocessors as part of a more general scheduling approach. The effectiveness of such heuristics are evaluated experimentally, showing them to achieve significantly better schedulability than the state of the art. Modeling of leakage from cache attacks is presented as a case study.

[52] (C) Quantitative information flow measurement techniques have been proven to be successful in detecting leakage of confidential information from programs. Modern approaches are based on formal methods, relying on program analysis to produce a SAT formula representing the program's

behavior, and model counting to measure the possible information flow. However, while program analysis scales to large codebases like the OpenSSL project, the formulas produced are too complex for analysis with precise model counting. In this paper we use the approximate model counter ApproxMC2 to quantify information flow. We show that ApproxMC2 is able to provide a large performance increase for a very small loss of precision, allowing the analysis of SAT formulas produced from complex code. We call the resulting technique ApproxFlow and test it on a large set of benchmarks against the state of the art. Finally, we show that ApproxFlow can evaluate the leakage incurred by the Heartbleed OpenSSL bug, contrarily to the state of the art.

[20] (C) We present HyLeak, a tool for reasoning about the quantity of information leakage in programs. The tool takes as input the source code of a program and analyzes it to estimate the amount of leaked information measured by mutual information. The leakage estimation is mainly based on a hybrid method that combines precise program analysis with statistical analysis using stochastic program simulation. This way, the tool combines the best of both symbolic and randomized techniques to provide more accurate estimates with cheaper analysis, in comparison with the previous tools using one of the analysis methods alone. HyLeak is publicly available and is able to evaluate the information leakage of randomized programs, even when the secret domain is large. We demonstrate with examples that HyLeaks has the best performance among the tools that are able to analyze randomized programs with similarly high precision of estimates.

[54] (J; submitted) Analysis of a probabilistic system often requires to learn the joint probability distribution of its random variables. The computation of the exact distribution is usually an exhaustive precise analysis on all executions of the system. To avoid the high computational cost of such an exhaustive search, statistical analysis has been studied to efficiently obtain approximate estimates by analyzing only a small but representative subset of the system's behavior. In this paper we propose a hybrid statistical estimation method that combines precise and statistical analyses to estimate mutual information, Shannon entropy, and conditional entropy, together with their confidence intervals. We show how to combine the analyses on different components of the system with different accuracy to obtain an estimate for the whole system. The new method performs weighted statistical analysis with different sample sizes over different components and dynamically finds their optimal sample sizes. Moreover it can reduce sample sizes by using prior knowledge about systems and a new abstraction-then-sampling technique based on qualitative analysis. To apply the method to the source code of a system, we show how to decompose the code into components and to determine the analysis method for each component by overviewing the implementation of those techniques in HyLeak tool. We demonstrate with case studies that the new method outperforms the state of the art in quantifying information leakage.

7.4.2. *Security for therapeutical environments*

Participants: Axel Legay, Olivier Zendra, Thomas Given-Wilson, Sean Sedwards.

This work is done in the context of the ACANTO EU project. We aim at helping develop robotic assistants to aid mobility of mobility-impaired and elderly adults. These robotic assistants provide a variety of support to their users, including: navigational assistance, social networking, social activity planning, therapeutic regime support, and diagnostic support. In Tamis, we focus on navigational assistance and social activities, as together they yield an interesting challenge in human robot interaction. The goal is to help groups of users navigate in a potentially busy dynamic environment, while also maintaining social group cohesion.

A robotic assistant has been developed before in the DALi project, acting selfishly to ensure the safe navigation of a single user. This was achieved by using the social force model and statistical model checking in a reactive planner that frequently replanned and made immediate navigational suggestions to the user. The key operational loop of this solution was to: observe the environment, model the agents in the environment in the social force model, give safety constraints for the user, and then use statistical model checking to find the optimal next move for the user.

Generalising to groups of users poses several significant difficulties. Computationally, the challenge is exponential in the number of users, considering all their possible navigational choices. Incomplete information is normal, since sensors are distributed between robotic assistants and the environment, and communication may fail, leading to different robots having different knowledge of the environment. Maintaining group cohesion is non-trivial, since group composition and position are dynamic and, unlike swarm robotics, no group member can be abandoned. Frequent replanning is necessary since there is minimal control over the users' actions, which may include ignoring the advise of the robotic assistant

The solution we designed is to abstract away from individual users in favour of groups. This refines the prior solution for a single user. Sensor information is used to obtain traces that provide behavioural information about users and pedestrians in the environment. These traces are clustered into groups that capture both location and motion behaviour. The groups are used as the social particles in the social force model, with parameters adjusted to account for group dynamics. Statistical model checking is used to find the optimal next move for the group containing the user, and the navigation for the optimal next move is displayed to the user. The effectiveness of the group abstraction mechanisms use in this refined algorithm are validated on the BIWI walking pedestrians dataset. This shows they operate correctly and effectively, even improving over human annotations, on real world data of pedestrians in a chaotic environment.

[27] (C) People with impaired physical and mental ability often find it challenging to negotiate crowded or unfamiliar environments, leading to a vicious cycle of deteriorating mobility and sociability. To address this issue the ACANTO project is developing a robotic assistant that allows its users to engage in therapeutic group social activities, building on work done in the DALi project. Key components of the ACANTO technology are social networking and group motion planning, both of which entail the sharing and broadcasting of information. Given that the system may also make use of medical records, it is clear that the issues of security, privacy, and trust are of supreme importance to ACANTO.

[58] (C; submitted) The ACANTO project is developing robotic assistants to aid the mobility and recovery of mobility-impaired and older adults. One key feature of the project's robotic assistants is aiding with navigation in chaotic environments. Prior work has solved this for a single user with a single robot, however for therapeutic outcomes ACANTO supports social groups and group activities. Thus these robotic assistants must be able to efficiently support groups of users walking together. This requires an efficient navigation solution that can handle large numbers of users, maintain (de-facto) group cohesion despite unpredictable behaviours, and operate rapidly on embedded devices. We address these challenges by: using sensor information to develop behavioural traces, clustering traces to determine groups, modeling the groups using the social force model, and finding an optimal navigation solution using statistical model checking. The new components of this solution are validated on the ETH Zürich dataset of pedestrians in an open environment.

7.4.3. Mobile air pollution sensor platform for smart-cities

Participant: Laurent Morin.

This work is organized and coordinated by the Chaire “mobilité dans une ville durable” and financed by the Foundation of Rennes 1 (https://fondation.univ-rennes1.fr/)

The purpose of this work is to design and experiment a mobile pollution sensor platform for Smart-Cities in Rennes.

The platform is integrated in the project ROAD (Rennes Open Access to Data) proposing to development of mobile systems operating the collection and the management of open data in Rennes for a future development of a smart-city. The collaboration is part of an ecosystem developed by the Chair “mobilité dans une ville durable” via the production of multiple experimentations in the city.

In the ROAD project context, the air quality in the city has been identified as one of the major challenge. Air quality improvement can only be achieved with a citizen and political full cooperation and involvement. This experimentation aims at providing an end-to-end urban platform that extends current practices in air quality measurements and allows citizens and policy makers to obtain the data and make informed decisions.

The mobile air pollution sensor platform for smart-cities proposes a innovative IoT architecture introducing the deployment of a small set of advanced and cost-effective sensors around a balanced high-performance/low-power compute unit inside a mobile agent in the city. The compute unit will have to provide the necessary computation power needed to produce advanced analysis and the security management on-site (integrity, authentication, ...).

The mobile sensor platform developments partially started in July 2017, and accelerated in October for a real deployment in buses in 2018. During this period, the core system of the platform was designed, adapted, and partially implemented to offer an operational prototype. This year lead to the design of a suitcase containing a self-sufficient measurement system: a main compute unit, its power supply and power management, and a set of satellite pollution sensors. This achievement was disseminated to the Rennes ecosystem (Rennes Atalante, Rennes Métropole, Inria) through the participation to several meetings and exhibitions.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- CISCO (http://www.cisco.com) contract (2017–2022) to work on graph analysis of malware

8.2. Bilateral Grants with Industry

- CISCO (http://www.cisco.com) one grant (2016–2019) to work on semantical analysis of malware
- Thales (https://www.thalesgroup.com) one CIFRE (2016–2019) to work on verification of communication protocols, one grant (2018–2019) to work on learning algorithms
- Oberthur Technologies (http://www.oberthur.com/) one grant (2016–2020) to work on fuzzing and fault injection
- Secure IC (http://www.secure-ic.com/), one CIFRE (2017–2020) to work on post-quantum cryptography

9. Partnerships and Cooperations

9.1. Regional Initiatives

- ARED grant for Lamine Nouredine and Florian Dolt
- Postdocs grants for Najah Ben Said, Jeffrey Paul Burdges, Ronan Lashermes, Ludovic Claudepierre
- Starting Grant for hardware for Annelie Heuser from Rennes Metropole
- Software developer grant for Laurent Morin from "Chaire Mobilité dans une ville durable" (mobility in a sustainable city) by Fondation Université Rennes 1

9.2. National Initiatives

9.2.1. ANR

- ANR MALTHY, Méthodes ALgébriques pour la vérification de modèles Temporisés et HYbrides, Thao Dang, 4 years, Inria and VISEO and CEA and VERIMAG
- ANR COGITO, Runtime Code Generation to Secure Devices, 3 years, Inria and CEA and ENSMSE and XLIM.

9.2.2. DGA

- PhD grant for Nisrine Jafri (2016–2019),

- PhD grant for Aurélien Palisse (2016–2019),
- PhD grant for Alexandre Gonzalves (2016–2019),
- PhD grant for Olivier Decourbe (2017–2020),
- PhD grant for Alexandre Zdhanov (2017–2020)

9.2.3. Autres

- INS2I JCJC grant for Axel Legay, Annelie Heuser, Fabrizio Biondi.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. ACANTO

Title: ACANTO: A CyberphusicAl social NeTwOrk using robot friends

Program: H2020

Duration: February 2015 - July 2018

Coordinator: Universita di Trento

Partners:

> Atos Spain (Spain), Envitel Tecnologia Y Control S.A. (Spain), Foundation for Research and Technology Hellas (Greece), Servicio Madrileno Delud (Spain), Siemens Aktiengesellschaft Oesterreich (Austria), Telecom Italia S.P.A (Italy), Universita' Degli Studi di Siena (Italy), Universita Degli Studi di Trento (Italy), University of Northumbria At Newcastle. (United Kingdom)

Inria contact: Axel Legay

Despite its recognised benefits, most older adults do not engage in a regular physical activity. The ACANTO project proposes a friendly robot walker (the FriWalk) that will abate a some of the most important barriers to this healthy behaviour. The FriWalk revisits the notion of robotic walking assistants and evolves it towards an activity vehicle. The execution of a programme of physical training is embedded within familiar and compelling every-day activities. The FriWalk operates as a personal trainer triggering the user actions and monitoring their impact on the physical and mental well-being. It offers cognitive and emotional support for navigation pinpointing risk situations in the environment and understanding the social context. It supports coordinated motion with other FriWalks for group activities. The FriWalk combines low cost and advanced features, thanks to its reliance on a cloud of services that increase its computing power and interconnect it to other assisted living devices. Very innovative is its ability to collect observations on the user preferred behaviours, which are consolidated in a user profile and used for recommendation of future activities. In this way, the FriWalk operates as a gateway toward a CyberPhysical Social Network (CPSN), which is an important contribution of the project. The CPSN is at the basis of a recommendation system in which users' profiles are created, combined into 'circles' and matched with the opportunity offered by the environment to generate recommendations for activities to be executed with the FriWalk support. The permanent connection between users and CPSN is secured by the FriPad, a tablet with a specifically designed user interface. The CPSN creates a community of users, relatives and therapists, who can enter prescriptions on the user and receive information on her/his state. Users are involved in a large number in all the phases of the system development and an extensive validation is carried out at the end.

Axel Legay and Olivier Zendra are the permanent researchers of Tamis involved in this project. The project supports two postdocs in Tamis.

9.3.1.2. DIVIDEND

Title: DIVIDEND: Distributed Heterogeneous Vertically IntegrateD Energy Efficient Data centres

Program: CHIST-ERA 2013

Duration: 10/2014 - 10/2016 (extended 10/2017)

Coordinator: University of Edinburgh (UK)

Partners:

> École Normale Supérieure de Paris, Département d'Informatique (France); Inria (France); Ecole Polytechnique Fédérale de Lausanne, Computer & Communication Sciences (Switzerland); Queen's University of Belfast, School of Electronics, Electrical Engineering and Computer Science, Belfast (UK); University of Edinburgh, Scotland, (UK); University of Lancaster, School of Computing and Communications (UK); University Politehnica Timisoara, Department of Computer Engineering (Romania)

Inria contact: Albert Cohen

The DIVIDEND project (http://www.chistera.eu/projects/dividend) attacks the data centre energy efficiency bottleneck through vertical integration, specialization, and cross-layer optimization. Our vision is to present heterogeneous data centres, combining CPUs, GPUs, and task-specific accelerators, as a unified entity to the application developer and let the runtime optimize the utilization of the system resources during task execution. DIVIDEND embraces heterogeneity to dramatically lower the energy per task through extensive hardware specialization while maintaining the ease of programmability of a homogeneous architecture. To lower communication latency and energy, DIVIDEND refers a lean point-to-point messaging fabric over complex connection-oriented network protocols. DIVIDEND addresses the programmability challenge by adapting and extending the industry-led heterogeneous systems architecture programming language and runtime initiative to account for energy awareness and data movement. DIVIDEND provides for a cross-layer energy optimization framework via a set of APIs for energy accounting and feedback between hardware, compilation, runtime, and application layers. The DIVIDEND project will usher in a new class of vertically integrated data centres and will take a first stab at resolving the energy crisis by improving the power usage effectiveness of data centres.

Contributions of Inria in the project addresses the development of an energy aware distributed heterogeneous system (distributed HSA) between data center applications and HSA accelerators. It includes the design of a common API able to interface two tasks: the monitoring of the energy consumption, and the management of distributed heterogeneous hardware at a data center scale. The project ended by a project review the 23th March 2017, and the last contributions to the project ended the 30th September 2017.

One of the main contribution is the design of and energy-aware distributed heterogeneous system architecture framework (D-HSA) built using the combination of three major levels: the hardware platform based on an aggregation of HSA compliant devices, the system level based on device drivers and energy monitoring libraries, and finally the application layer using an extension of standard OpenCL programming model. This OpenCL extension is proposed as the main API for the energy-aware distributed HSA, and was made available for the tools and applications developed in the project.

A second contribution is the specification and the implementation of a distributed extension of the standard HSA Runtime API, and its functional validation on a basic system. The extension integrates the discovery, the management, and the execution of kernel computations on remote HSA agents in a distributed environment. The validation is based on an implementation using the Message Passing Interface (MPI) standard on an HSA compliant AMD machine. The Distributed HSA extension proposed offers a fully functional API for managing remote and distributed HSA agents, but at the cost of a limitation of the capability of the D-HSA system: the standard HSA memory model, based essentially on a coherent shared memory, is not supported for distributed HSA agents. As a primary implementation, focusing on a functional support of the new D-HSA verbs, this works tend to demonstrate that the extension is light and easy-to-use for a set of examples.

Laurent Morin from Tamis is involved in this project

9.3.1.3. EMC2

Title: Embedded Multi-Core Systems for Mixed Criticality Applications in Dynamic and Changeable Real-Time Environments

Program: FP7

Duration: April 2014 - March 2017

Coordinator: Infineon Technologies

Partners:

Aicas (Germany) Avl Software and Functions (Germany), Denso Automotive Deutschland (Germany), Elektrobit Automotive (Germany), Evision Systems (Germany), Nxp Semiconductors Germany (Germany), Tttech Computertechnik (Austria), "kompetenzzentrum - Das Virtuelle Fahrzeug, Forschungsgesellschaft Mbh" (Austria), Frequentis (Austria), Thales Austria (Austria), Blueice Bvba (Belgium), Freescale Polovodice Ceska Republika Sro (Czech Republic), Sysgo Sro (Czech Republic), Silkan Rt (France), "united Technologies Research Centre Ireland," (Ireland), Mbda Italia Spa (Italy), Fornebu Consulting As (Norway), Westerngeco As (Norway), Simula Research Laboratory As (Norway), Ixion Industry and Aerospace Sl (Spain), Visure Solutions Sl (Spain), Seven Solutions Sl (Spain), Telvent Energia (Spain), Instituto Tecnologico de Informatica (Spain), Ambar Telecomunicaciones Sl (Spain), Sics Swedish Ict (Sweden), Arcticus Systems (Sweden), Arccore (Sweden), Xdin Stockholm (Sweden), Systemite (Sweden), Stichting Imec Nederland (Netherlands), Tomtom International Bv (Netherlands), Infineon Technologies Uk Ltd (United Kingdom), Sundance Multiprocessor Technology Ltd (United Kingdom), Systonomy (United Kingdom), Ensilica Ltd (United Kingdom), Test and Verification Solutions Ltd (United Kingdom), Abb (Sweden), Ait Austrian Institute of Technology (Austria), Alenia Aermacchi Spa (Italy), Avl List (Austria), Airbus Defence and Space (Germany), Bayerische Motoren Werke Aktiengesellschaft (Germany), Vysoke Uceni Technicke V Brne (Czech Republic), Commissariat A L Energie Atomique et Aux Energies Alternatives (France), Consorzio Interuniversitario Nazionale Per l'Informatica (Italy), Centro Ricerche Fiat (Italy), Critical Software (Portugal), Chalmers Tekniska Hoegskola (Sweden), Danfoss Power Electronics As (Denmark), Danmarks Tekniske Universitet (Denmark), Ericsson (Sweden), Fraunhofer-Gesellschaft Zur Foerderung Der Angewandten Forschung E.V (Germany), Hi Iberia Ingenieria Y Proyectos Sl (Spain), Harokopio University (Greece), Infineon Technologies Austria (Austria), Institut Mikroelektronickych Aplikaci S.R.O. (Czech Republic), Inesc Id - Instituto de Engenharia de Sistemas E Computadores, Investigacao E Desenvolvimento Em Lisboa (Portugal), Infineon Technologies (Germany), Integrasys (Spain), Instituto Superior de Engenharia Do Porto (Portugal), Kungliga Tekniska Hoegskolan (Sweden), Lulea Tekniska Universitet (Sweden), Magillem Design Servicess (France), Nxp Semiconductors Netherlands Bv (Netherlands), Offis E.V. (Germany), Philips Medical Systems Nederland Bv (Netherlands), Politecnico di Torino (Italy), Quobis Networks Sl (Spain), Rockwell Collins France (France), Rigas Tehniska Universitate (Latvia), Selex Es Spa (Italy), Siemens Aktiengesellschaft (Germany), Systematic Paris Region Association (France), Sysgo (Germany), Thales Alenia Space Italia Spa (Italy), "thales Alenia Space Espana," (Spain), Technolution B.V. (Netherlands), Fundacion Tecnalia Research & Innovation (Spain), Thales Communications & Securitys (France), Thales Avionicss (France), Thales (France), Nederlandse Organisatie Voor Toegepast Natuurwetenschappelijk Onderzoek Tno (Netherlands), Technische Universitat Braunschweig (Germany), Technische Universiteit Delft (Netherlands), Technische Universitat Dortmund (Germany), Technische Universitaet Kaiserslautern (Germany), Technische Universitaet Wien (Austria), Technische Universiteit Eindhoven (Netherlands), Universita Degli Studi di l'aquila (Italy), Universita Degli Studi di Genova (Italy), The University of Manchester (United Kingdom), University of Bristol (United Kingdom), University of Limerick (Ireland), "ustav Teorie Informace A Automatizace Av Cr, V.V.I." (Czech

Republic), Universitetet I Oslo (Norway), Vector Fabrics Bv (Netherlands), Volvo Technology (Sweden)

Inria contact: Albert Cohen and Axel Legay

Embedded systems are the key innovation driver to improve almost all mechatronic products with cheaper and even new functionalities. Furthermore, they strongly support today's information society as inter-system communication enabler. Consequently boundaries of application domains are alleviated and ad-hoc connections and interoperability play an increasing role. At the same time, multi-core and many-core computing platforms are becoming available on the market and provide a breakthrough for system (and application) integration. A major industrial challenge arises facing (cost) efficient integration of different applications with different levels of safety and security on a single computing platform in an open context. The objective of the EMC2 project (Embedded multi-core systems for mixed criticality applications in dynamic and changeable real-time environments) is to foster these changes through an innovative and sustainable service-oriented architecture approach for mixed criticality applications in dynamic and changeable real-time environments. The EMC2 project focuses on the industrialization of European research outcomes and builds on the results of previous ARTEMIS, European and National projects. It provides the paradigm shift to a new and sustainable system architecture which is suitable to handle open dynamic systems. EMC2 is part of the European Embedded Systems industry strategy to maintain its leading edge position by providing solutions for: . Dynamic Adaptability in Open Systems . Utilization of expensive system features only as Service-on-Demand in order to reduce the overall system cost. . Handling of mixed criticality applications under real-time conditions . Scalability and utmost flexibility . Full scale deployment and management of integrated tool chains, through the entire lifecycle Approved by ARTEMIS-JU on 12/12/2013 for EoN. Minor mistakes and typos corrected by the Coordinator, finally approved by ARTEMIS-JU on 24/01/2014. Amendment 1 changes approved by ECSEL-JU on 31/03/2015.

The permanent members of Tamis who are involved are Axel Legay and Olivier Zendra. The project was initiated during the lifetime of the ESTASYS.Inria team.

9.3.1.4. ENABLE-S3

Title: ENABLE-S3: European Initiative to Enable Validation for Highly Automated Safe and Secure Systems

Program: H2020

Duration: 05/2016 - 04/2019

Coordinator: Avl List Gmbh (Austria)

Partners:

Aalborg Universitet (Denmark); Airbus Defence And Space Gmbh (Germany); Ait Austrian Institute Of Technology Gmbh (Austria); Avl Deutschland Gmbh (Germany); Avl Software And Functions Gmbh (Germany); Btc Embedded Systems Ag (Germany); Cavotec Germany Gmbh (Germany); Creanex Oy(Finland); Ceske Vysoke Uceni Technicke V Praze (Czech Republic); Deutsches Zentrum Fuer Luft - Und Raumfahrt Ev (Germany); Denso Automotive Deutschland Gmbh (Germany); Dr. Steffan Datentechnik Gmbh (Austria); Danmarks Tekniske Universitet (Denmark); Evidence Srl (Italy); Stiftung Fzi Forschungszentrum Informatik Am Karlsruher Institut Fur Technologie (Germany); Gmv Aerospace And Defence Sa (Spain); Gmvis Skysoft Sa (Portugal); Politechnika Gdanska (Poland); Hella Aglaia Mobile Vision Gmbh (Germany); Ibm Ireland Limited (Ireland); Interuniversitair Micro-Electronica Centrum (Belgium); Iminds (Belgium); Institut National De Recherche Eninformatique Et Automatique (France); Instituto Superior De Engenharia Do Porto (Portugal); Instituto Tecnologico De Informatica (Spain); Ixion Industry And Aerospace Sl (Spain); Universitat Linz (Austria); Linz Center Of Mechatronics Gmbh (Austria); Magillem Design Services Sas (France); Magneti Marelli S.P.A. (Italy); Microeletronica Maser Slspain); Mdal (France); Model Engineering Solutions Gmbhgermany); Magna Steyr Engineering Ag & Co Kg (Austria); Nabto

> Aps (Denmark); Navtor As (Norway); Nm Robotic Gmbh (Austria); Nxp Semiconductors Germany Gmbh(Germany); Offis E.V.(Germany); Philips Medical Systems Nederland Bvnetherlands); Rohde & Schwarz Gmbh&Co Kommanditgesellschaft(Germany); Reden B.V. (Netherlands); Renault Sas (France); Rugged Tooling Oyfinland); Serva Transport Systems Gmbh(Germany); Siemens Industry Software Nvbelgium); University Of Southampton (Uk); Safetrans E.V. (Germany); Thales Alenia Space Espana, Saspain); Fundacion Tecnalia Research & Innovationspain); Thales Austria Gmbh (Austria); The Motor Insurance Repair Researchcentre (Uk); Toyota Motor Europe (Belgium); Nederlandse Organisatie Voor Toegepast Natuurwetenschappelijk Onderzoek Tno (Netherlands); Ttcontrol Gmbh (Austria); Tttech Computertechnik Ag (Austria); Technische Universiteit Eindhoven (Netherlands); Technische Universitat Darmstadt (Germany); Technische Universitaet Graz (Austria); Twt Gmbh Science & Innovation (Germany); University College Dublin, National University Of Ireland, Dublin (Ireland); Universidad De Las Palmas De Gran Canaria (Spain); Universita Degli Studi Di Modena E Reggio Emilia (Italy); Universidad Politecnica De Madrid (Spain); Valeo Autoklimatizace K.S. (Czech Republic); Valeo Comfort And Driving Assistance (France); Valeo Schalter Und Sensoren Gmbh (Germany); Kompetenzzentrum - Das Virtuelle Fahrzeug, Forschungsgesellschaft Mbh (Austria); Vires Simulationstechnologie Gmbh (Germany); Teknologian Tutkimuskeskus Vtt Oy (Finland); Tieto Finland Support Services Oy (Finland); Zilinska Univerzita V Ziline (Slovakia);

Inria contact: Axel Legay

The objective of ENABLE-S3 (http://www.enable-s3.eu) is to establish cost-efficient cross-domain virtual and semi-virtual V&V platforms and methods for ACPS. Advanced functional, safety and security test methods will be developed in order to significantly reduce the verification and validation time but preserve the validity of the tests for the requested high operation range. ENABLE-S3 aspires to substitute today's physical validation and verification efforts by virtual testing and verification, coverage-oriented test selection methods and standardization. ENABLE-S3 is use-case driven; these use cases represent relevant environments and scenarios. Each of the models, methods and tools integrated into the validation platform will be applied to at least one use case (under the guidance of the V&V methodology), where they will be validated (TRL 5) and their usability demonstrated (TRL6). Representative use cases and according applications provide the base for the requirements of methods and tools, as well as for the evaluation of automated systems and respective safety. This project is industry driven and has the objective of designing new technologies for autonomous transportation, including to secure them. Tamis tests its results on the case studies of the project.

Axel Legay and Jean-Louis Lanet are involved in this project. The project supports one postdoc in Tamis starting in 2017.

9.3.1.5. SUCCESS

Title: SUCCESS: SecUre aCCESSibility for the internet of things

Program: CHIST-ERA 2015

Duration: 10/2016 - 10/2018

Coordinator: Middlesex University (UK)

Partners:

> Middlesex University, School of Science and Technology (France); Inria (France); Université Grenoble Alpes, Verimag (FRANCE); Univesity of TWENTE, (Netherlands)

Inria contact: Axel Legay

The SUCCESS project ...The core idea of SUCCESS is to use formal methods and verification tools with a proven track record to provide more transparency of security risks for people in given IoT scenarios. Our core scientific innovation will consist on the extension of well-known industry-strength methods Our technological innovation will provide adequate tools to address risk assessment and adaptivity within IoT in healthcare environments and an open source repository to foster future reuse, extension and progress in this area. Our project will validate the scientific and technological innovation through pilots, one of which will be in collaboration with a hospital and will allow all stakeholders (e.g. physicians, hospital technicians, patients and relatives) to enjoy a safer system capable to appropriately handle highly sensitive information on vulnerable people while making security and privacy risks understandable and secure solutions accessible.

Within SUCCESS, the contribution of the TAMIS team consists in a framework for analyzing the security of a given IOT system, and notably whether it resists to attack. Our approach is to build a high-level model of the system, including vulnerabilities, as well as an attacker. We represent the set of possible attacks using an attack tree. Finally, we evaluate the probability that an attack succeeds using Statistical Model Checking.

In the TAMIS team, Axel Legay, Delphine Beaulaton, Najah Ben-Saïd and Jean Quilbeuf are involved in this project.

9.3.1.6. TeamPlay

Title: TeamPlay: Time, Energy and security Analysis for Multi/Many-core heterogeneous PLAtforms

Program: H2020

Duration: 01/2018 - 12/2020

Coordinator: Inria

Partners:

> Absint Angewandte Informatik Gmbh (Germany), Institut National De Recherche en Informatique et Automatique (France), Secure-Ic Sas (France), Sky-Watch A/S (Danemark), Syddansk Universitet (Danemark), Systhmata Ypologistikis Orashs Irida Labs Ae (Greece), Technische Universität Hamburg-Harburg (Germany), Thales Alenia Space Espana (Spain), Universiteit Van Amsterdam (Netherlands), University Of Bristol (UK), University Of St Andrews (UK)

Inria contact: Olivier Zendra and Axel Legay

The TeamPlay (Time, Energy and security Analysis for Multi/Many-core heterogeneous PLAtforms) project federates 6 academic and 5 industrial partners and aims to develop new, formally-motivated, techniques that will allow execution time, energy usage, security, and other important non-functional properties of parallel software to be treated effectively, and as first- class citizens. We will build this into a toolbox for developing highly parallel software for low-energy systems, as required by the internet of things, cyber-physical systems etc. The TeamPlay approach will allow programs to reflect directly on their own time, energy consumption, security, etc., as well as enabling the developer to reason about both the functional and the non-functional properties of their software at the source code level. Our success will ensure significant progress on a pressing problem of major industrial importance: how to effectively manage energy consumption for parallel systems while maintaining the right balance with other important software metrics, including time, security etc. The project brings together leading industrial and academic experts in parallelism, energy modeling/ transparency, worst-case execution time analysis, non-functional property analysis, compilation, security, and task coordination. Results will be evaluated using industrial use cases taken from the computer vision, satellites, flying drones, medical and cyber security domains. Within TeamPlay, Inria and TAMIS coordinate the whole project, while being also in charge of aspects related more specifically to security.

The permanent members of Tamis who are involved are Axel Legay, Olivier Zendra and Annelie Heuser.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Axel Legay was General Chair for KimFest, an event organized for the 60th Birthday of Kim. G. Larsen
- Axel Legay was General Chair for the 1st ACM SAC Conference Track on Software-intensive Systems-of-Systems
- Jean-Louis Lanet was General Chair of Crisis 2017,
- Olivier Zendra was General co-Chair for ARCHI'17, the 9th Summer school on « Architecture des systèmes matériels et logiciels embarqués, et méthodes de conception associées »

10.1.1.2. Member of the Organizing Committees

- Axel Legay coordinated KimFest, an event organized for the 60th Birthday of Kim. G. Larsen

10.1.2. Scientific Events Selection

10.1.2.1. Member of Conference Steering Committees

- Axel Legay is a member of the Steering Committee of the Security summer school organized jointly by pre-GDR security and PEC (Pole d'Excellence Cyber).
- Jean-Louis Lanet has been member of the Steering Committee of Cardis 2017
- Olivier Zendra is a founder and a member of the Steering Committee of ICOOOLPS (International Workshop on Implementation, Compilation, Optimization of OO Languages, Programs and Systems)

10.1.2.2. Chair of Conference Program Committees

- Axel Legay was Scientific chair of the 23rd International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)
- Axel Legay was the Scientific chair for the 11th International Conference on Risks and Security of Internet and Systems
- Axel Legay was the Scientific chair for the 17th International Conference on Application of Concurrency to System Design (ACSD 2017)
- Olivier Zendra was co-chair with Mario Wolzco of the Program Committee and the Organizing Committee of the 12th Workshop on Implementation, Compilation, Optimization of Object-Oriented Languages, Programs and Systems (ICOOOLPS 2017)

10.1.2.3. Member of the Conference Program Committees

- Axel Legay was a PC member of RV'17, ACSD'17, TACAS'17, CRISIS'17, CMSB'17, SETTA'17, FORMALIZE'17
- Fabrizio Biondi was a PC member of CRISIS'17, MCETECH'17, SAC'17
- Jean-Louis Lanet was PC member of Gramsec'17, Secitc'17, C2SI'17, Mcetech'17 and Afadl'17
- Olivier Zendra was PC member of ICOOOLPS'2017, ARCHI'17 and PEC 2017.

10.1.2.4. Reviewer

- Olivier Zendra was reviewer for MFCS.
- Fabrizio Biondi was a reviewer for CRISIS'17, ESORICS'17, KimFest, LATA'17, MFCS'17, RV'17, MCETECH'17

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Axel Legay is a funder and member of the editorial board of “Foundations for Mastering Changes” journal.
- Annelie Heuser was PC Member/Editorial Board for IACR Transactions On Cryptographic Hardware And Embedded Systems

10.1.3.2. Reviewer - Reviewing Activities

- Axel Legay was reviewer for TCS, TSE, Information and Computation.
- Annelie Heuser was a reviewer for Transactions on Information Forensics & Security, Journal of Cryptographic Engineering, Transactions on Embedded Computing Systems, IEEE Transactions on Very Large Scale Integration Systems
- Jean-Louis was reviewer of Computer and Security journal

10.1.4. Invited Talks

- Axel Legay was an invited speaker for the 11th International Workshop on Reachability Problems.
- Axel Legay was invited speaker for the 43rd International Conference on Current Trends in Theory and Practice of Computer Science
- Fabrizio Biondi was invited speaker for The 12th International Conference on Risks and Security of Internet and Systems
- Florian Dold was invited to the "Re-Imagining Finance" workshop at Columbia Law School in New York City in September 2017.
- Annelie Heuser was invited to a panel discussion for Malicious Software and Hardware in Internet of Things (ACM International Conference on Computing Frontiers
- Jean-Louis Lanet was invited speaker for the INS3PECT workshop, the ROOTS conference, the Conference on Operational Planning, Technological Innovations and Mathematical Applications, the Journée AFSEC. 2017)

10.1.5. Scientific Expertise

- Axel Legay was an expert for the Wallonie Government.
- Axel Legay participated to the CR2 jury for Inria Nice Center as a member of Inria’s evaluation committee.
- Olivier Zendra is a CIR expert for the MENESR.
- Olviier Zendra participated to the CR2 jury for Inria Paris Center as a member of Inria’s evaluation committee.
- Olivier Zendra is a member of the editorial board and co-author of the “HiPEAC 2017 Vision” [47], as well as the HiPEAC 2019 Vision.

10.1.6. Research Administration

- Axel Legay is a member of Inria’s evaluation committee.
- Axel Legay is the Representative for non-permanent staff committees (in charge of postdocs).
- Axel Legay is a member of "club équipe Française de la cyber sécurité"
- Axel Legay is the Britany Region representative in the ECSO organization
- Olivier Zendra is a member of Inria’s evaluation committee.
- Olivier Zendra is a member of Inria’s workgroup on Inria’s social barometer.
- Olivier Zendra was a member of Inria’s CNHSCT.
- Olivier Zendra was Head of Inria Nancy’s IES Committee (formerly IST).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Master : Axel Legay, Introduction au Model Checking, 36, M2, Université de Bretagne Sud, France
- Master : Axel Legay, Introduction à l'analyse de risques, M2, Université de Bretagne Sud, France
- Licence : Nisrine Jafri , Programmation Java, L3, l'ISTIC, Université Rennes 1, France

10.2.2. Supervision

- PhD in progress : Kevin Bukasa, Démarrage sécurisé, 2015, Jean-Louis Lanet and Axel Legay
- PhD in progress : Mounir Chadli (Rennes 1), On Scheduling and SMC, December 2014, Axel Legay and Saddek Bensalem.
- PhD in progress : Olivier Descourbe, On Code Obfuscation, October 2016, Axel Legay and Fabrizio Biondi.
- PhD in progress : Mike Enescu, On Symbolic Execution for Malware Detection, October 2016, Axel Legay, Flavio Oquendo and Fabrizio Biondi. Terminated on October 2017.
- PhD in progress : Alexandre Gonsalvez, On Obfuscation via crypto primitives, April 2016, Axel Legay and Caroline Fontaine.
- PhD in progress : Nisrine Jafri (Rennes1), On fault Injection detection with MC of Binary code, December 2015, Axel Legay and Jean-Louis Lanet.
- PhD in progress : Razika Lounas, Validation des spécifications formelles de la mise à jour dynamique des applications Java Card, 2010, Mohamed Mezghiche and Jean-Louis Lanet
- PhD in progress: Martin Moreau (Rennes1); On the study of post-quantum cryptography mechanisms (provisory), Axel Legay, Annelie Heuser and Sylvain Guilley
- PhD in progress : Routa Moussaileb, From Data Signature to Behavior Analysis, 2017, Nora Cuppens and Jean-Louis Lanet
- PhD in progress : Tristan Ninet (Rennes 1), Vérification formelle d'une implémentation de la pile protocolaire IKEv2, December 2016, Axel Legay, Romaric Maillard and Olivier Zendra
- PhD in progress: Lamine Nouredine (Rennes1); Developing new packing detection techniques to stop malware propagation, November 2017, Axel Legay and Annelie Heuser.
- PhD in progress : Aurélien Palisse, Observabilité de codes hostiles, 2015, Jean-Louis Lanet
- PhD in progress: Emmanuel Tacheau (Rennes1); Analyse et détection de malwares au moyen de méthodes d'analyse symbolique, September 2017, Axel Legay, Fabrizio Biondi, Alain Fiocco.
- PhD in progress : Aurélien Trulla, Caractérisation de malware Android par suivi de flux d'information et nouvelles techniques d'évasion, 2016, Valerie Viet Triem Tong and Jean-Louis Lanet
- PhD in progress: Alexander Zhdanov (Rennes 1): Modular Automated Syntactic Signature Extraction (MASSE), December 2017, Axel Legay, Fabrizio Biondi, François Déchelle and Olivier Zendra.

10.2.3. Juries

- Axel Legay was a referee for the PhD defense of Xavier Devroye (University of Namur Belgium)
- Axel Legay was a referee for the PhD defense of Quentin Cappart (University of Louvain Belgium)
- Axel Legay was a referee for the PhD defense of Stefan Naujokat (University of Dortmund, Germany)

10.3. Popularization

- Axel Legay participated to the "Forum Cyberstrategia" organized by the ministry of defense, September 2017
- Axel Legay participated to the "Inria Industry days" organized by Inria, October 2017
- Axel Legay participated to the "table ronde sur l'intelligence économique", Rennes November 2017

- Fabrizio Biondi participated to the "Forum International de la Cybersécurité", January 2017
- Fabrizio Biondi participated to the "Forum Cyberstrategia" organized by the ministry of defense, September 2017
- Fabrizio Biondi participated to the "Inria Industry days" organized by Inria, October 2017
- Fabrizio Biondi participated to the "European Cyber Week" organized by IRISA and Bretagne Development Innovation, November 2017

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] Á. García-Recuero. *Discouraging abusive behavior in privacy-preserving decentralized online social networks*, Université Rennes 1, May 2017, https://tel.archives-ouvertes.fr/tel-01548658

[2] C. Grothoff. *The GNUnet System*, Université de Rennes 1, October 2017, Habilitation à diriger des recherches, https://hal.inria.fr/tel-01654244

Articles in International Peer-Reviewed Journals

[3] A. Arnold, M. Baleani, A. Ferrari, M. Marazza, V. Senni, A. Legay, J. Quilbeuf, C. Etzien. *An Application of SMC to continuous validation of heterogeneous systems*, in "EAI Endorsed Transactions on Industrial Networks and Intelligent Systems", February 2017, vol. 4, n^{o} 10, pp. 1-19 [*DOI :* 10.4108/eai.1-2-2017.152154], https://hal.inria.fr/hal-01630523

[4] F. Biondi, S. Josse, A. Legay, T. Sirvent. *Effectiveness of Synthesis in Concolic Deobfuscation*, in "Computers and Security", September 2017, vol. 70, pp. 500-515 [*DOI :* 10.1016/j.cose.2017.07.006], https://hal.inria.fr/hal-01241356

[5] Y. Cheherazed, J.-L. Lanet, M. Mezghiche, K. Tamine. *Machine Learning Techniques to Predict Sensitive Patterns to Fault Attack in the Java Card Application*, in "Journal of Experimental and Theoretical Artificial Intelligence", 2017, forthcoming, https://hal.inria.fr/hal-01645392

[6] X. Devroey, G. Perrouin, M. Cordy, H. Samih, A. Legay, P.-Y. Schobbens, P. Heymans. *Statistical Prioritization for Software Product Line Testing: an Experience Report*, in "Software and Systems Modeling", February 2017, https://hal.inria.fr/hal-01642289

[7] F. Dold, C. Grothoff. *Byzantine set-union consensus using efficient set reconciliation*, in "EURASIP Journal on Information Security", December 2017, vol. 2017, n^{o} 1, pp. 1-18 [*DOI :* 10.1186/s13635-017-0066-3], https://hal.inria.fr/hal-01657397

[8] J. Duchêne, C. Le Guernic, E. Alata, V. Nicomette, M. Kaâniche. *State of the art of network protocol reverse engineering tools*, in "Journal of Computer Virology and Hacking Techniques", 2017, 27 p. [*DOI :* 10.1007/s11416-016-0289-8], https://hal.inria.fr/hal-01496958

[9] N. El Mrabet, A. Mrabet, R. Lashermes, J.-B. Rigaud, B. Bouallegue, S. Mesnager, M. Machhout. *A scalable and systolic architectures of montgomery modular multiplication for public key cryptosystems based on dsps*, in "Journal Hardware and Systems Security", 2017, https://hal.archives-ouvertes.fr/hal-01579811

[10] A. Heuser, S. Picek, S. Guilley, N. Mentens. *Lightweight Ciphers and their Side-channel Resilience*, in "IEEE Transactions on Computers", August 2017, pp. 1-16 [*DOI :* 10.1109/TC.2017.2757921], https://hal.archives-ouvertes.fr/hal-01629886

[11] J.-L. Lanet, H. Le Bouder, M. Benattou, A. Legay. *When time meets test*, in "International Journal of Information Security", 2017 [*DOI :* 10.1007/s10207-017-0371-3], https://hal.inria.fr/hal-01645395

[12] J.-L. Lanet, A. Mesbah, M. Mezghiche. *Reverse engineering a Java Card memory management algorithm*, in "Computers & Security", May 2017, vol. 66, pp. 97 - 114 [*DOI :* 10.1016/j.cose.2017.01.005], https://hal.inria.fr/hal-01645393

[13] R. Lounas, J.-L. Lanet, M. Mezghiche. *A Formal Verification of Dynamic Updating in a Java-based embedded System*, in "International Journal of Critical Computer-Based Systems", 2017, forthcoming, https://hal.inria.fr/hal-01645401

[14] B. Nicolas, C. Carlet, S. Guilley, A. Heuser, E. Prouff, O. Rioul. *Stochastic Collision Attack*, in "IEEE Transactions on Information Forensics and Security", April 2017, vol. 12, n^{o} 9, pp. 2090 - 2104 [*DOI :* 10.1109/TIFS.2017.2697401], https://hal.archives-ouvertes.fr/hal-01629880

[15] B. Nicolas, S. Guilley, A. Heuser, M. Damien, O. Rioul. *Optimal side-channel attacks for multivariate leakages and multiple models*, in "Journal of Cryptographic Engineering", August 2017, vol. 7, n^{o} 4, pp. 331–341 [*DOI :* 10.1007/s13389-017-0170-9], https://hal.archives-ouvertes.fr/hal-01629885

[16] S. Picek, A. Heuser, S. Guilley. *Template attack versus Bayes classifier*, in "Journal of Cryptographic Engineering", September 2017, vol. 7, n^{o} 4, pp. 343–351 [*DOI :* 10.1007/s13389-017-0172-7], https://hal.archives-ouvertes.fr/hal-01629884

[17] È. de Chèrisey, S. Guilley, A. Heuser, O. Rioul. *On the optimality and practicability of mutual information analysis in some scenarios*, in "Cryptography and Communications - Discrete Structures, Boolean Functions and Sequences ", June 2017, https://hal.archives-ouvertes.fr/hal-01645127

International Conferences with Proceedings

[18] F. Biondi, M. Chadli, T. Given-Wilson, A. Legay. *Information Leakage as a Scheduling Resource*, in "International Workshop on Formal Methods for Industrial Critical Systems and Automated Verification of Critical Systems", Turin, Italy, September 2017, https://hal.inria.fr/hal-01382052

[19] F. Biondi, F. Déchelle, A. Legay. *MASSE: Modular Automated Syntactic Signature Extraction*, in "ISSRE 2017 - The 28th International Symposium on Software Reliability Engineering - IEEE", Toulouse, France, IEEE, October 2017, pp. 1-2, https://hal.inria.fr/hal-01629035

[20] F. Biondi, Y. Kawamoto, A. Legay, L.-M. Traonouez. *HyLeak: Hybrid Analysis Tool for Information Leakage*, in "ATVA 2017 - Fifteenth International Symposium on Automated Technology for Verification and Analysis", Pune, India, October 2017, 14 p. , https://hal.inria.fr/hal-01546817

[21] A. Bkakria, M. Graa, N. Cuppens-Boulahia, F. Cuppens, J.-L. Lanet. *Real-time detection and reaction to Activity hijacking attacks in Android smartphones*, in "PST 2017, Privacy, Security, and Trust - 15th International Conference", Calgary, Canada, August 2017, https://hal.inria.fr/hal-01645399

[22] Q. Cappart, C. Limbrée, P. Schaus, J. Quilbeuf, L.-M. Traonouez, A. Legay. *Verification of Interlocking Systems Using Statistical Model Checking*, in "18th IEEE International Symposium on High Assurance Systems Engineering (HASE)", Singapore, Singapore, 18th IEEE International Symposium on High Assurance Systems Engineering (HASE), January 2017, pp. 61 - 68 [*DOI :* 10.1109/HASE.2017.10], https://hal.archives-ouvertes.fr/hal-01591338

[23] C. Carlet, A. Heuser, S. Picek. *Trade-Offs for S-Boxes: Cryptographic Properties and Side-Channel Resilience*, in "ACNS 2017 - International Conference on Applied Cryptography and Network Security", Kanazawa, Japan, LNCS, Springer, July 2017, vol. 10355, pp. 393-414 [*DOI :* 10.1007/978-3-319-61204-1_20], https://hal.archives-ouvertes.fr/hal-01629879

[24] M. Contini, G. Di Natale, A. Heuser, T. Poppelmann, N. Mentens. *Do we need a holistic approach for the design of secure IoT systems?*, in "Computing Frontiers Conference", Siena, Italy, May 2017, https://hal.archives-ouvertes.fr/hal-01628683

[25] X. Devroey, G. Perrouin, M. Papadakis, A. Legay, P.-Y. Schobbens, P. Heymans. *Automata Language Equivalence vs. Simulations for Model-based Mutant Equivalence: An Empirical Evaluation*, in "ICST 2017 - International Conference on Software Testing, Verification and Validation", tokyo, Japan, March 2017, https://hal.inria.fr/hal-01640101

[26] T. Given-Wilson, N. JAFRI, J.-L. Lanet, A. Legay. *An Automated Formal Process for Detecting Fault Injection Vulnerabilities in Binaries and Case Study on PRESENT*, in "2017 IEEE Trustcom/BigDataSE/ICESS", Sydney, Australia, August 2017, pp. 293 - 300 [*DOI :* 10.1109/Trustcom/BigDataSE/ICESS.2017.250], https://hal.inria.fr/hal-01629098

[27] T. Given-Wilson, A. Legay, S. Sedwards. *Information Security, Privacy, and Trust in Social Robotic Assistants for Older Adults*, in "HAS 2017 - International Conference on Human Aspects of Information Security, Privacy, and Trust", Vancouver, Canada, T. Tryfonas (editor), LNCS, Springer, July 2017, vol. 10292, pp. 90-109 [*DOI :* 10.1007/978-3-319-58460-7_7], https://hal.inria.fr/hal-01629094

[28] P. Gjøl Jensen, K. G. Larsen, A. Legay, D. B. Poulsen. *Quantitative Evaluation of Attack Defense Trees using Stochastic Timed Automata*, in "GraMSec 2017 - The Fourth International Workshop on Graphical Models for Security", Santa Barbara , United States, August 2017, https://hal.inria.fr/hal-01640091

[29] M. Graa, N. Cuppens-Boulahia, F. Cuppens, J.-L. Lanet, R. Moussaileb. *Detection of Side Channel Attacks Based on Data Tainting in Android Systems*, in "SEC 2017 - 32th IFIP International Conference on ICT Systems Security and Privacy Protection", Rome, Italy, S. D. C. di Vimercati, F. Martinelli (editors), ICT Systems Security and Privacy Protection, Springer International Publishing, May 2017, vol. AICT-502, pp. 205-218, Part 4: Operating System and Firmware Security [*DOI :* 10.1007/978-3-319-58469-0_14], https://hal.inria.fr/hal-01648994

[30] S. Guilley, A. Heuser, O. Rioul. *Codes for Side-Channel Attacks and Protections*, in "C2SI 2017 - International Conference on Codes, Cryptology, and Information Security", Rabat, Morocco, LNCS, April 2017, vol. 10194, pp. 35-55 [*DOI :* 10.1007/978-3-319-55589-8_3], https://hal.archives-ouvertes.fr/hal-01629876

[31] B. Kevin, R. Lashermes, J.-L. Lanet, H. Le Bouder, A. Legay. *How TrustZone could be bypassed: Side-Channel Attacks on a modern System-on-Chip*, in "Wistp'17, International Conference on Information Security Theory and Practice", Heraklion, Greece, September 2017, https://hal.inria.fr/hal-01645398

[32] A. KUNNAPPILLY, A. LEGAY, T. MARGARIA, C. SECELEANU, B. STEFFEN, L.-M. TRAONOUEZ. *Analyzing Ambient Assisted Living Solutions: A Research Perspective*, in "12th International Conference on Desig &Technology of Integrated Systems In Nanoscale Era (DTIS)", Palma de Mallorca, Spain, 12th International Conference on Desig &Technology of Integrated Systems In Nanoscale Era (DTIS), April 2017, pp. 1 - 7 [*DOI :* 10.1109/DTIS.2017.7930168], https://hal.archives-ouvertes.fr/hal-01591347

[33] J.-L. LANET. *Experimenting similarity-based hijacking attacks detection and response in Android Systems*, in "ICISS, International Conference on Information Systems Security", Bombay, India, December 2017, https://hal.inria.fr/hal-01645400

[34] J.-L. LANET, A. PALISSE. *When data reveals ransomware activity*, in "4th International Conference on Operational Planning, Technological Innovations and Mathematical Applications,", Athen, Greece, May 2017, https://hal.inria.fr/hal-01645396

[35] C. LE GUERNIC. *Toward a Sound Analysis of Guarded LTI Loops with Inputs by Abstract Acceleration (extended version)*, in "Static Analysis Symposium", New York, United States, Lecture Notes in Computer Science, August 2017, vol. 10422, https://hal.inria.fr/hal-01550767

[36] G. LI, P. M. JENSEN, K. G. LARSEN, A. LEGAY, D. B. POULSEN. *Practical Controller Synthesis for MTL*$0, \infty$, in "International SPIN Symposium on Model Checking of Software", Santa barbara , United States, July 2017, https://hal.inria.fr/hal-01642162

[37] G. PERROUIN, P. HEYMANS, A. LEGAY, X. DEVROEY, M. CORDY, P.-Y. SCHOBBENS. *On Featured Transition Systems*, in "SOFSEM 2017 - 43rd International Conference on Current Trends in Theory and Practice of Informatics", Limerick , Ireland, January 2017, https://hal.inria.fr/hal-01640267

[38] S. PICEK, A. HEUSER, A. JOVIC, A. LEGAY. *Climbing Down the Hierarchy: Hierarchical Classification for Machine Learning Side-Channel Attacks*, in "AFRICACRYPT 2017 - International Conference on Cryptology in Africa", Dakar, Senegal, LNCS, Springer, May 2017, vol. 10239, pp. 61-78 [*DOI :* 10.1007/978-3-319-57339-7_4], https://hal.archives-ouvertes.fr/hal-01629878

Conferences without Proceedings

[39] S. GUILLEY, A. HEUSER, M. TANG, O. RIOUL. *Stochastic Side-Channel Leakage Analysis via Orthonormal Decomposition*, in "SecITC 2017", Bucharest, Romania, June 2017, https://hal.archives-ouvertes.fr/hal-01628679

[40] J.-L. LANET. *Formal Methods and the Dark Side of the Force*, in "Journée Approches Formelles des Systèmes Embarqués Communicants", Paris, France, June 2017, https://hal.inria.fr/hal-01645403

[41] J.-L. LANET, A. MESBAH. *The Express Laundry - from black box to white box*, in "Workshop InS3PECT: Ingénierie Système de Services Sécurisés Pour objEts ConnecTé", Nice, France, December 2017, https://hal.inria.fr/hal-01645404

[42] A. LEGAY, L.-M. TRAONOUEZ. *Plasma Lab Statistical Model Checker: Architecture, Usage and Extension*, in "SOFSEM 2017 - 43rd International Conference on Current Trends in Theory and Practice of Computer Science", Limerick, Ireland, January 2017, https://hal.archives-ouvertes.fr/hal-01613581

[43] A. MESBAH, J.-L. LANET, M. MEZGHICHE. *Reverse Engineering a Code without the Code*, in "1st Reversing and Offensive-oriented Trends Symposium 2017", Vienna, Austria, Sergey Bratus (Dartmouth College), November 2017, https://hal.inria.fr/hal-01591926

[44] A. MESBAH, M. MEZGHICHE, J.-L. LANET. *Persistant Fault Injection Attack, From White-box to Black-box*, in "The 5th International Conference on Electrical Engineering - ICEE 2017", Boumedrès, Algeria, October 2017, https://hal.inria.fr/hal-01591931

[45] A. PALISSE, A. DURAND, H. LE BOUDER, C. LE GUERNIC, J.-L. LANET. *Data Aware Defense (DaD): Towards a Generic and Practical Ransomware Countermeasure,*, in "NordSec 2017", Tartu, Estonia, November 2017, https://hal.inria.fr/hal-01591939

[46] S. PICEK, A. HEUSER, A. JOVIC, S. LUDWIG, S. GUILLEY, D. JAKOBOVIĆ, N. MENTENS. *Side-channel analysis and machine learning: A practical perspective*, in "International Joint Conference on Neural Networks (IJCNN)", Anchorage, United States, May 2017, https://hal.archives-ouvertes.fr/hal-01628681

Books or Proceedings Editing

[47] M. DURANTON, K. DE BOSSCHERE, C. GAMRAT, J. MAEBE, H. MUNK, O. ZENDRA (editors). *The HiPEAC Vision 2017*, HiPEAC network of excellence, January 2017, 138 p. , https://hal.inria.fr/hal-01491758

[48] F. OQUENDO, K. DRIRA, A. LEGAY, T. BATISTA (editors). *Proceedings of the 1st ACM SAC Conference Track on Software-intensive Systems-of-Systems (SiSoS 2017): 32nd ACM SIGAPP Symposium On Applied Computing*, ACM, Marrakesh, Morocco, April 2017, https://hal.archives-ouvertes.fr/hal-01445350

Scientific Popularization

[49] A. PALISSE, J. JONGWANE. *Quelles solutions pour se protéger des logiciels de rançon ?*, in "Interstices", October 2017, https://hal.inria.fr/hal-01688779

Other Publications

[50] N. BEN SAID, T. ABDELLATIF, M. BOZGA, S. BEN SALEM, A. LEGAY. *Orchestration for Secure Multi-party Communications in Web-Services ☆*, November 2017, working paper or preprint, https://hal.inria.fr/hal-01629427

[51] N. BEN SAID, F. BIONDI, V. BONTCHEV, O. DECOURBE, T. GIVEN-WILSON, A. LEGAY, J. QUILBEUF. *Detection of Mirai by Syntactic and Semantic Analysis*, November 2017, working paper or preprint, https://hal.inria.fr/hal-01629040

[52] F. BIONDI, M. A. ENESCU, A. HEUSER, A. LEGAY, K. S. MEEL, J. QUILBEUF. *Scalable Approximation of Quantitative Information Flow in Programs*, November 2017, working paper or preprint, https://hal.inria.fr/hal-01629131

[53] F. BIONDI, T. GIVEN-WILSON, A. LEGAY. *Universal Optimality of Apollonian Cell Encoders*, October 2017, working paper or preprint, https://hal.inria.fr/hal-01571226

[54] F. BIONDI, Y. KAWAMOTO, A. LEGAY, L.-M. TRAONOUEZ. *Hybrid Statistical Estimation of Mutual Information and its Application to Information Flow*, November 2017, working paper or preprint, https://hal.inria.fr/hal-01629033

[55] M. Chadli, J. H. Kim, K. G. Larsen, A. Legay, S. Naujokat, B. Steffen, L.-M. Traonouez. *High-level Frameworks for the Specification and Verification of Scheduling Problems*, October 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01613576

[56] T. Given-Wilson, A. Heuser, N. JAFRI, J.-L. Lanet, A. Legay. *An Automated and Scalable Formal Process for Detecting Fault Injection Vulnerabilities in Binaries*, November 2017, working paper or preprint, https://hal.inria.fr/hal-01629135

[57] T. Given-Wilson, N. JAFRI, J.-L. Lanet, A. Legay. *An Automated Formal Process for Detecting Fault Injection Vulnerabilities in Binaries and Case Study on PRESENT – Extended Version*, April 2017, working paper or preprint, https://hal.inria.fr/hal-01400283

[58] T. Given-Wilson, A. Legay, S. Sedwards, O. Zendra. *Modelling of Machine-Aided Human Group Motion*, November 2017, working paper or preprint, https://hal.inria.fr/hal-01629137

[59] A. Heuser, S. Picek, A. Jovic, A. Legay. *On the Relevance of Feature Selection for Profiled Side-channel Attacks*, November 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01645128

[60] A. Heuser, S. Picek, A. Legay, K. Knezevic. *Profiled SCA with a New Twist: Semi-supervised Learning*, November 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01645130

[61] L.-M. Traonouez, A. Legay, D. Nowotka, D. B. Poulsen. *Statistical Model Checking of LLVM Code*, November 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01640097

References in notes

[62] M. Chadli, J. H. Kim, A. Legay, L.-M. Traonouez, S. Naujokat, B. Steffen, K. G. Larsen. *A Model-Based Framework for the Specification and Analysis of Hierarchical Scheduling Systems*, in "FMICS-AVoCS", Pise, Italy, Critical Systems: Formal Methods and Automated Verification, Springer, September 2016, vol. 9933, pp. 133 - 141 [*DOI :* 10.1007/978-3-319-45943-1_9], https://hal.archives-ouvertes.fr/hal-01241681

[63] Á. García-Recuero, J. Burdges, C. Grothoff. *Privacy-Preserving Abuse Detection in Future Decentralised Online Social Networks*, in "11th International ESORICS Workshop in Data Privacy Management, DPM 2016", Heraklion, Crete, Greece, G. Livraga, V. Torra, A. Aldini, F. Martinelli, N. Suri (editors), Springer Lecture Notes in Computer Science (LNCS) series, Springer, September 2016, vol. 9963, pp. 78-93 [*DOI :* 10.1007/978-3-319-47072-6_6], https://hal.inria.fr/hal-01355951

[64] A. Heuser, S. Picek, S. Guilley, N. Mentens. *Side-channel Analysis of Lightweight Ciphers: Current Status and Future Directions*, October 2016, Lightweight Cryptography Workshop 2016, https://hal.inria.fr/hal-01407264

[65] A. Heuser, S. Picek, S. Guilley, N. Mentens. *Side-channel Analysis of Lightweight Ciphers: Does Lightweight Equal Easy?*, in "RFIDSec 2016: 12th Workshop on RFID and IoT Security", Hong Kong, Hong Kong SAR China, November 2016, https://hal.inria.fr/hal-01402238

[66] A. Savary, M. Frappier, M. Leuschel, J. Lanet. *Model-Based Robustness Testing in Event-B Using Mutation*, in "Software Engineering and Formal Methods - 13th International Conference, SEFM 2015, York,

UK, September 7-11, 2015. Proceedings", R. CALINESCU, B. RUMPE (editors), Lecture Notes in Computer Science, Springer, 2015, vol. 9276, pp. 132–147, http://dx.doi.org/10.1007/978-3-319-22969-0_10

[67] N. H. WALFIELD, J. L. GRIFFIN, C. GROTHOFF. *A Quantitative Analysis of Cell Tower Trace Data for Understanding Human Mobility and Mobile Networks*, in "6th International Workshop on Mobile Entity Localization, Tracking and Analysis (MELT)", San Francisco, United States, October 2016 [*DOI :* 10.1145/1235], https://hal.inria.fr/hal-01378622