Activity Report 2015

Team ESTASYS

Efficient STAtistical methods in SYstems of systems

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
Proofs and Verification

Table of contents

1. Members **1**
2. Overall Objectives **2**
3. Research Program **2**
3.1.1. Systems of Systems (SoS) 2
3.1.2. Grand Challenge and Breakthroughs of ESTASYS 3
3.1.3. Methodology and Organization 3
4. Highlights of the Year **5**
5. New Software and Platforms **5**
5.1. PLASMA Lab 5
5.2. PyECDAR 5
5.3. Quail 6
6. New Results **6**
6.1. Heterogeneous Systems 6
6.2. Statistical Model Checking 7
6.3. Formal Models for Variability 10
6.4. Privacy and Security 11
6.4.1. Information-Theoretical Quantification of Security Properties 11
6.4.2. Equivocation-based Security Measures for Shared-Key Cryptosystems 13
6.4.3. Malware Classification via Deobfuscation and Behavioral Fingerprinting 14
6.5. Energy-Centric Systems 15
6.6. Languages for composition 16
7. Partnerships and Cooperations **17**
7.1. Regional Initiatives 17
7.1.1. Privacy 17
7.1.2. Variability 17
7.2. National Initiatives 17
7.2.1. ANR Malthy 17
7.2.2. BGLE SyS2Soft 17
7.3. European Initiatives 17
7.3.1. FP7 & H2020 Projects 17
7.3.2. Danse 18
7.3.3. Meals 18
7.3.4. Sensation 18
7.3.5. EMC2 19
7.3.6. Collaborations with Major European Organizations 19
7.4. International Initiatives 19
8. Dissemination **19**
8.1. Promoting Scientific Activities 19
8.1.1. Scientific events selection 19
8.1.2. Journal 19
8.1.2.1. Member of the editorial boards 19
8.1.2.2. Reviewer - Reviewing activities 20
8.1.3. Invited talks 20
8.1.4. Leadership within the scientific community 20
8.1.5. Scientific expertise 20
8.1.6. Research administration 20
8.2. Teaching - Supervision - Juries 20
8.2.1. Teaching 20
8.2.2. Supervision 20

8.2.3. Juries 20
8.3. Popularization 20
9. Bibliography **20**

Team ESTASYS

Creation of the Team: 2014 January 01, end of the Team: 2015 December 31

Keywords:

Computer Science and Digital Science:

4. - Security and privacy
4.5. - Formal methods for security
6. - Modeling, simulation and control
6.1. - Mathematical Modeling
6.3. - Computation-data interaction
8.2. - Machine learning

Other Research Topics and Application Domains:

4. - Energy
4.2. - Renewable energy production
4.4. - Energy consumption
6. - IT and telecom
6.4. - Internet of things
6.5. - Information systems
6.6. - Embedded systems

1. Members

Research Scientist

Axel Legay [Team leader, Inria, Researcher, HdR]

Engineers

Fabrizio Biondi [Inria]
Rudolf Fahrenberg [Inria]
Thomas Given-Wilson [Inria]
Cyrille Jegourel [Inria, until Apr 2015]
Laurent Morin [Inria, from Jun 2015]
Van-Chan Ngo [Inria]
Sean Sedwards [Inria]
Louis Marie Traonouez [Inria]

PhD Students

Mounir Chadli [Inria]
Nisrine Jafri [Inria]

Post-Doctoral Fellows

Jin Hyun Kim [Inria, until Nov 2015]
Jean Quilbeuf [Inria, until Dec 2015 now Bretagne Sud and Inria]

Administrative Assistant

Stephanie Lemaile [Inria]

Other

Rafael Olaechea Velazco [Inria,from Jun 2015 until Sep 2015]

2. Overall Objectives

2.1. Overall Objectives

Computer systems play a central role in modern societies and their errors can have dramatic consequences. Industry and academics thus invest a considerable amount of effort developing techniques to prove the correctness of these systems. Among such techniques, one finds (1) *testing*, the traditional approach to detect bugs with test cases, and (2) *formal methods*, e.g., model checking (Turing award), that can *guarantee* the absence of bugs. Both approaches have been largely deployed on static systems, whose behaviour is entirely known. **ESTASYS focuses on developping brand new formal methods for Systems of Systems.**

3. Research Program

3.1. Systems of Systems, Heterogeneous Systems, Dynamicity, Statistical Model Checking

Formal methods rely on the notion of *transition system* (TS): an abstract machine that characterises a system's *complete* behaviour. This machine consists of a complete set of states (each representing full knowledge of the system at a given moment) and transitions between states, which may be labelled with labels chosen from some set of actions. This definition makes it necessary to have advanced knowledge of all the possible states of the system – to have a statically configured system. The algorithms used by formal methods perform an exhaustive exploration of the state space of the TS, so such methods suffer from the so-called *state-space explosion problem*. As a consequence, there are many real systems that are beyond the scope of such techniques. Despite this, over the last thirty years it has been shown that, when combined with heuristics such as partial order reductions or abstraction, **formal approaches are powerful enough to verify industrial-scale systems**.

The first wave of techniques was deployed to verify whether a certain set of (problem) states can be reached ('reachability'). Later, extensions of TS, such as *hybrid systems* and *stochastic automata*, were proposed to cope with new problems (e.g., energy consumption) or to reason on distributed real-time embedded components (possibly heterogeneous). It was quickly observed that the complexity of assessing correctness of such extended models arises not exclusively from the fact that they are large, but also because they introduce *undecidability*. As a concrete example, the reachability problem is already undecidable for any real-time system whose time evolution is described by a non-constant derivative equation.

This motivated the development of more efficient techniques that approximate the answer to the original problem. Of these, perhaps the most successful quantitative technique is *Statistical Model Checking* , that can be seen as a trade-off between testing and formal verification. The core idea of SMC is to generate a number of *simulations* of the system and verify whether they satisfy a given property expressed in temporal logics, which can be done by using *runtime verification approaches*. The results are then used together with algorithms from the statistical area in order to decide whether the system satisfies the property with some probability. SMC resembles classical simulation-based techniques used in industry, but uses a formal model of systems and requirements. This not only gives a rigorous meaning to industrial practices, but also makes available more than twenty years of research in the area of *runtime verification*. Last but not least, **the use of statistical algorithms allows us to approximate undecidable problems**. Recent successful applications of SMC can be found in systems biology, security protocols and avionics. In particular, SMC was used to discover inconsistent requirements of an EADS airplane communication system.

3.1.1. Systems of Systems (SoS)

The advent of service-oriented and cloud architectures is leading to generations of computer systems that exhibit a new type of complexity: such systems are no longer statically configured, but comprise components that are systems in their own right, able to discover, select and bind on-the-fly to other components that can deliver services that they require. These complex systems, referred to as *Systems of Systems* (SoS), can change over time as each component creates and modifies the network over which it needs to operate: as they execute, the components create a network of their own and use it to fulfil their goals.

The Internet, made up of an unsupervised and rapidly growing, dynamically configured set of computers and physical connections, is an obvious illustration of the potential complexity of dynamic networks of interactions. Another example is the so-called "Flash Crash" in the U.S. equity market: on May 6, 2010, a block sale of 4.1 billion dollars of futures contracts executed on behalf of a fund-management company triggered a complex pattern of interactions between the high-frequency algorithmic trading systems that buy and sell blocks of financial instruments and made the Dow Jones Industrial Average drop more than 600 points, representing the disappearance of 800 billion dollars of market value. This example is an illustration of the faulty divergence of SoS behaviour, where the system starts to misbehave and dynamically creates new components that follow the same pattern and make the problem worse. Examples of this include when a SoS detects high energy use and invokes a new component to reduce the energy, thus consuming *more* energy. **Until now, such divergence has been mostly handled by humans that eventually observe the faulty behaviour and manually intervene to stop it. This human-based solution is not always successful and clearly unsatisfactory, since it acts retrospectively, when the system has already failed.**

3.1.2. Grand Challenge and Breakthroughs of ESTASYS

SoS are an efficient means of achieving high performance and are thus becoming ubiquitous. Society's increasing reliance on SoS demands that they are reliable, but tools to guarantee this at the design stage do not exist. Most conventional formal analysis techniques, even those dedicated to adaptive systems, fail when applied to SoS because they are designed to reason on systems whose state space can be predicted in advance. **The grand challenge addressed by ESTASYS is the fundamental overhaul of formal methods techniques in the design of SoS life cycle**.

It is clear that SMC can be applied to the verification of complex systems. Unfortunately, SMC cannot yet be applied to SoS, because existing techniques are designed to capture the behaviour of statically configured systems, or systems whose dynamical configuration arises from permutations of known components. ESTASYS defines new abstract computational models and extend the state of the art of SMC to include SoS.

ESTASYS proposes a new formal methodology to support an evolutionary adaptive and iterative SoS life cycle. *We foresee the following breakthroughs:*

1. Our ground-breaking computational model addresses the complex dynamic nature of SoS. The model is based on new interface theories that take into account behaviours of possibly unknown components and thus abstract what is unknown.
2. Cutting edge algorithms coming from the area of statistics and learning are exploited to make predictions about autonomous systems making local decisions. For example, **statistical abstraction** abstracts the behaviour of unknown environments by interleaving analysis and runtime monitoring of deployed systems to continuously update distributions embedded in the interfaces.
3. New statistical algorithms for SMC that scale efficiently and handle undecidability impacts the formal analysis of complex systems.
4. Our results are implemented in a professional toolset, ESTASYS-PLASMA, that is constructed in close collaboration with our industrial partners. This ensures relevance to industry and potentially high impact in the marketplace.

3.1.3. Methodology and Organization

ESTASYS's main challenge is to lay the foundation of a novel rigorous software construction methodology for SoS, based on simulation, statistics and industrial practices. ESTASYS establishes theories and empirical evidence for the introduction of formal verification-based approaches in the rigorous design of SoS.

ESTASYS addresses essential research questions for the introduction of formal techniques to support the SoS life-cycle. SoS occur in multiple disciplines and therefore there is a need for a common language. In particular, notions such as **autonomous decisions and dynamicity** must be standardized and well understood by those that will apply our methodology. Additionally, **characterizing the topological structure** of a SoS is essential for the study of component interactions and data exchanges. The complexity of SoS requires the development of a **sound formal semantic foundation** to support deployment of formal methods. We thus

identify a minimal computational model that characterize SoS, on which classes of properties of interest can be defined. The project investigates new simulation-based approaches, combined with other domains (statistics, learning, ...), to verify such properties on the new computational model. Finally, ESTASYS identifies under which conditions the new techniques can be used, to take decisions during design and evolution time, leading to a fully integrated development cycle.

ESTASYS focuses on both the static and dynamic properties of SoS. ESTASYS establishes models for each component and investigates the connection and dynamical interactions between them. ESTASYS's activities are organize in six main tasks: tasks 1, 2 and 3 are responsible for breakthrough 1; task 4 is responsible for breakthrough 2; task 5 is responsible for breakthrough 3; task 6 is responsible for breakthrough 4.

Task 1. Characterizing SoS. Examples of SoS found in various areas, such as health care, smart buildings and energy grids, are analysed and used to standardize notions of autonomous decisions and dynamicity. We also study and classify SoS-related problems, such as faulty behaviour divergence. Our objective is to derive in Task 2 formal models that abstract the above classification.

Task 2. Formal Modeling of SoS. Classical theories do not provide for SoS, hence we require new formal models for SoS that take into account (i) dynamicity and emergent behaviours, (ii) autonomous decisions of components, and (iii) architectural constraints, including information regarding the viability of the hardware. In particular, we devise new logics tailored to the specific needs of SoS. Such logics, dynamic by nature, includes extended notions of quantification, such as energy, and considers hardware constraints and distributions of system configurations. Task 2 includes modelling the various components running within the SoS and their (dynamical) interactions. This requires the definition of a new type of interface able to work with heterogeneous components and to abstract the behaviour of unknown resources. Interfaces act as an abstraction for the internal behaviour of each component and encodes the dynamical constraints of the SoS. They are used to (i) model and define the authorised interactions between the components, (ii) reason on dynamical aspects and (iii) abstract unknown behaviour.

Task 3. Statistical abstraction interleaving design and deployment. Abstraction techniques are necessary to reduce the complexity of SoS and to model uncertainty. Specifically, **statistical abstractions** of the observed runtime behaviour of components is used to quantify, e.g., the probability that a number of new components satisfying some constraints is started at a given execution point. Runtime verification monitors the executions of the deployed system to create distributions embedded in the interfaces developed in Task 1. When a deployed system is available, ESTASYS interleaves simulation, analysis and runtime monitoring, using real behaviour to update the statistical abstractions, and eventually replace some of those abstractions by concrete ESTASYS-Interface models. The ESTASYS methodology adopts a Bayesian approach: (i) an initial, plausible distribution is 'guessed', based on whatever is known; (ii) the system is simulated using the current approximated distribution; (iii) the behaviour of the simulated system becomes the new approximation; (iv) the process is iterated as necessary. While learning-based simulation approaches, such as model fitting, can be used to learn the abstraction by conducting simulations from a finite set of initial components, we have to provide clear evidence that a global property holds on the system if it holds on its corresponding statistical abstraction. The task requires strong competences in statistics.

Task 4. Developing Efficient Simulation and Monitoring Algorithms for SoS. The ground-breaking models developed in Task 2 requires efficient simulation and monitoring techniques. This necessitates the study of new algorithms for dynamically configured systems and monitoring approaches to reason on heterogeneous components and the new quantitative logics and interface paradigms developed in Task 2.
A major difficulty in developing monitoring techniques for SoS is that the components have their own goals and behave differently in different environments. Unnecessary high-level hypotheses on properties may drastically increase simulation time and should be avoided.

Task 5. Developing Efficient Statistical Techniques for SoS. SoS pose new challenges for statistical techniques, requiring the study of new SMC algorithms dedicated to SoS goals. In contrast to existing SMC algorithms that can only be applied to pure stochastic systems, SMC algorithms for SoS have to take into account the non-deterministic aspects of autonomous decisions made by neighbour components. We postulate that this can be done by extending very recent advances in reinforcement learning algorithms. Rare events play an important role in system reliability, so we include rare-event simulation algorithms, such as importance sampling and importance splitting, which can reduce variance and significantly increase simulation efficiency.

Task 6. Evaluating the impact of statistical and simulation-based techniques. Evidence of the success of ESTASYS is provided by the publishing of a complete experimental environment, ESTASYS-PLASMA, that supports the empirical validation of ESTASYS's theories. ESTASYS-PLASMA contains efficient implementations of the results discovered in Tasks 2-5, and will provide intuitive feedback mechanisms so that the engineer can use the results of the verification process to improve SoS design.

4. Highlights of the Year

4.1. Highlights of the Year

The ESTASYS team has developped a full tool chain for the rigorous design of Systems of Systems and has achieved its two years objectives. The team has also prepared its reconfiguration into a new team where security issues will become fundamental.

4.1.1. Awards

Axel Legay has received a Villumn award from Aalborg University.

5. New Software and Platforms

5.1. PLASMA Lab

Platform for Learning and Advanced Statistical Model checking Algorithms
KEYWORDS: Model Checking - Statistical - Model Checker - Runtime Analysis - Statistics
SCIENTIFIC DESCRIPTION

Statistical model checking (SMC) is a fast emerging technology for industrial scale verification and optimisation problems. Plasma was conceived to have high performance and be extensible, using a proprietary virtual machine. Since SMC requires only an executable semantics and is not constrained by decidability, we can easily implement different modelling languages and logics.
FUNCTIONAL DESCRIPTION

Plasma-Lab is a formal verification tool for complex embeded systems. It uses statistical model checking, and applies to complex problems coming from the area of security, cyber physical systems, or privacy.

- Participants: Axel Legay, Sean Sedwards, Louis-Marie Traonouez, Jean Quilbeuf
- Contact: Axel Legay
- URL: https://project.inria.fr/plasma-lab

5.2. PyECDAR

KEYWORDS: Timed input - Output automata
SCIENTIFIC DESCRIPTION

The tool has been originally developed to analyze the robustness of timed specifications, in extension of the tool Ecdar. As Ecdar , it allows to compose components specifications based on Timed I/O Automata (TIOA), and it implements timed game algorithms for checking consistency and compatibility. Additionally, it features original methods for checking the robustness of these specifications.

The tool has been later extended to analyse adaptive systems. It therefore implements original algorithms for checking featured timed games against requirements expressed in the timed AdaCTL logic.

The tool is written in Python with around 3'000 lines of code. It uses a Python console as user interface, from which it can load TIOA components from XML files written in the UPPAAL format, and design complex systems by combining the components using a simple algebra. Then, it can analyze these systems, transform them and save them in a new XML file.

FUNCTIONAL DESCRIPTION

PyEcdar is a free software that analyses timed games and timed specifications. The goal of the tool is to allow a fast prototyping of new analysis techniques. It currently allows to solve timed games based on timed automata models. These can be extended with adaptive features to represent dynamicity and to model software product lines.

- Participants: Louis-Marie Traonouez and Axel Legay
- Contact: Louis-Marie Traonouez
- URL: https://project.inria.fr/pyecdar/

5.3. Quail

FUNCTIONAL DESCRIPTION

Privacy is a central issue for Systems of Systems and interconnected objects. We propose QUAIL, a tool that can be used to quantify privacy of components. QUAIL is the only tool able to perform an arbitrary-precision quantitative analysis of the security of a system depending on private information. Thanks to its Markovian semantics model, QUAIL computes the correlation between the system's observable output and the private information, obtaining the amount of bits of the secret that the attacker will infer by observing the output.

- Participants: Fabrizio Biondi, Axel Legay, Louis-Marie Traonouez and Andrzej Wasowski
- Contact: Axel Legay
- URL: https://project.inria.fr/quail/

6. New Results

6.1. Heterogeneous Systems

Participants: Axel Legay, Jean Quilbeuf.

> This part concerns Tasks 1, 2 and 4 of the action. We characterize and formalize heterogeneous aspects of SoS and then we define efficient monitoring algorithms and representations for their requirements. We then combine the results with Statistical Model Checking (Task 5).

Systems of Systems (SoS) are very large scale systems with particular characteristics. SoS are not directly built from scratch by a single designer or a single team but are obtained as the composition of simpler systems. SoS have strong reliability and dependability requirements, as they aim to provide a service over a long running period. SoS may dynamically modify themselves by connecting to new systems, updating or disconnecting faulty ones, making it impossible to statically know the set of subsystems that are part of the SoS before runtime.

One of the main difficulty arising when developing SoS is the fact that subsystems may have been designed with a different goal in mind. In particular, some subsystems may have their own goal which differs from the global goal of the SoS. Furthermore, each subsystem may be developed in a particular computation model, making it difficult to find a common unifying semantics for the whole SoS. Finally, SoS may exhibit some emergent behaviors that are hardly predictable at design time.

One of the solutions to allow simulation of an SoS is to rely on a common interface for interconnecting the subsystems. The Functional Mockup Interface (FMI) standard is a natural candidate for such an interface. The different components of an SoS developed in different models of computation can be translated to Functional Mockup Units (FMU). Then a so-called master algorithm coordinates the FMUs composing the system. The execution of each FMU is either directly handled by the master algorithm or relies on an external tool for its execution.

Because the subsystems composing an SoS are of heterogeneous nature, it is difficult to find a common semantics model for the whole system. Furthermore, building such a transition system is not tractable due to the complexity of the system. Thus verification through traditional model checking is not possible for SoS. However, since the FMI/FMU framework enables simulation of such systems, the statistical model checking approach can be used.

The DANSE EU project aims to provide a complete tool chain from the modeling to the verification of SoS. At the higher level, the modeling is done in UPDM using the RHAPSODY tool. At the same level, the designer can express requirements over the model using some patterns written in GCSL. The UPDM model can then be translated into a FMI/FMU format that can be simulated by a dedicated tool, named DESYRE. Similarly, the GCSL requirements are transformed into BLTL formulas. Finally, the PLASMA statistical model checker has been integrated with the DESYRE tool chain in order to check the BLTL formulas based on the simulations provided by DESYRE.

6.1.1. Papers:

papier DANSE(en cours) Ensuring a correct behaviour of SoS has a significant social impact. Their complexity and inherent dynamicity pose a serious challenge to traditional design methodologies. We propose a methodology and a tool-chain supporting design and validation of SoSs. We integrate SMC with existing industrial practice, by addressing both methodological and technological issues. Our contribution is summarized as follows: (1) a methodology for continuous and scalable validation of SoS formal requirements; (2) a natural-language based formal specification language able to express complex SoS requirements; (3) adoption of widely used industry standards for simulation and heterogeneous systems integration (FMI and UPDM); (4) development of a robust SMC tool-chain integrated with system design tools used in practice. We illustrate the application of our SMC tool-chain and the obtained results on an industrial case study from the DANSE project.

6.2. Statistical Model Checking

Participants: Axel Legay, Sean Sedwards, Jean Quilbeuf, Louis-Marie Traonouez, Chan Ngo, Cyrille Jegourel.

> This section covers Tasks 4 and 5 of the action. It consists in developping Simulation based techniques and efficient statistical algorithms for SoS.

The use of test cases remains the default means of ensuring the correct behaviour of systems in industry, but this technique is limited by the need to hypothesise scenarios that cause interesting behaviour and the fact that a reasonable set of test cases is unlikely to cover all possible eventualities. Static analysis is more thorough and has been successful in debugging very large systems, but its ability to analyse complex dynamical properties is limited. In contrast, model checking is an exhaustive technique that verifies whether a system satisfies a dynamical temporal logic property under all possible scenarios. For nondeterministic and probabilistic systems, numerical model checking quantifies the probability that a system satisfies a property. It can also be used to quantify the expected cost or reward of sets of executions.

Numerical model checking gives precise, accurate and certain results by exhaustively exploring the state space of the model, however the exponential growth of the state space with system size (the 'state explosion problem) typically limits its applicability to "toy" systems. Symbolic model checking using efficient data structures can make certain very large models tractable. It may also be possible to construct simpler but behaviourally equivalent models using various symmetry reduction techniques, such as partial order reduction, bisimulation and lumping. If a new system is being constructed, it may be possible to guarantee the overall behaviour by verifying the behaviour of its subcomponents and limiting the way they interact. Despite these techniques, however, the size, unpredictability and heterogeneity of real systems usually make numerical techniques infeasible. Moreover, even if a system has been specified not to misbehave, it is nevertheless necessary to check that it meets its specification.

Simulation-based approaches are becoming increasingly tractable due to the availability of high performance parallel hardware and algorithms. In particular, statistical model checking (SMC) combines the simplicity of testing with the formality of numerical model checking. The core idea of SMC is to create multiple independent execution traces of a system and count how many satisfy a property specified in temporal logic. The proportion of satisfying traces is an estimate of the probability that the system satisfies the property. By thus modelling the executions of a system as a Bernoulli random variable, the absolute error of the estimate can be bounded using, for example, a confidence interval or a Chernoff bound. It is also possible to use efficient sequential hypothesis testing, to decide with specified statistical confidence whether the probability of a property is above or below a given threshold. Since SMC requires multiple independent simulations, it may be efficiently divided on parallel computer architectures, such as grids, clusters, clouds and general purpose computing on graphics processors (GPGPU).

Knowing a result with less than 100% confidence is often sufficient in real applications, since the confidence bounds may be made arbitrarily tight. Moreover, a swiftly achieved approximation may prevent a lot of wasted time during model design. For many complex systems, SMC offers the only feasible means of quantifying performance. Historically relevant SMC tools include APMC, YMER and VESTA. Well-established numerical model checkers, such as PRISM and UPPAAL, are now also including SMC engines. Dedicated SMC tools under active development include COSMOS and our own tool PLASMA. Recognising that SMC may be applied to any discrete event trace obtained by stochastic simulation, we have devised PLASMA-lab, a modular library of SMC algorithms that may be used to construct domain-specific SMC tools. PLASMA-lab has become the main vehicle of our ongoing development of SMC algorithms.

Statistical model checking (SMC) addresses the state explosion problem of numerical model checking by estimating quantitative properties using simulation. To advance the state of the art of SMC we address the ongoing challenges of rare events and nondeterminsm. We also make novel use of SMC by applying it to motion planning in the context of assisted living. Rare events are often of critical importance and are challenging to SMC because they appear infrequently in simulations. Nondeterministic models are useful to model unspecified interactions, but simulation requires that nondeterminism is resolved.

We also applied SMC in the context of Systems of Systems (SoS). In the frame of the DANSE project, Plasma-Lab was used to verify SoS, and completely integrated with the DANSE tool-chain. We are currently working on verification of dynamic SoS, where systems can appear and disappear during execution. This work is done in collaboration with the ArchWare team from IRISA. We will interface Plasma-Lab with a simulator for the Pi-ADL language that enables simulation of dynamic systems.

Our group is devising cutting edge techniques for SMC. In particular, we are developing new algorithms for non-deterministic systems as well as for dynamic systems. Rare event systems are also addressed. Finally, we also devote a large amount of time to applying our technology to realistic case studies described in high-level languages such as Simulink or System C, or even a robot moving an eldery person in a commercial center.

6.2.1. Papers:

[2] (J) 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 we present a novel motion planning algorithm that is able to intelligently deal

with crowded areas, permanent or temporary anomalies in the environment (e.g., road blocks, wet floors) as well as hard and soft constraints (e.g., "keep a toilet within reach of 10 meters during the journey", "always avoid stairs"). Constraints can be assigned a priority tailored on the user's needs. The planner has been validated by means of simulations and experiments with elderly people within the context of the DALi FP7 EU project.

[3] (J) Markov decision processes (MDP) are useful to model optimisation problems in concurrent systems. To verify MDPs with efficient Monte Carlo techniques requires that their nondeterminism be resolved by a scheduler. Recent work has introduced the elements of lightweight techniques to sample directly from scheduler space, but finding optimal schedulers by simple sampling may be inefficient. Here we describe "smart" sampling algorithms that can make substantial improvements in performance.

[21] (C) Rare properties remain a challenge for statistical model checking (SMC) due to the quadratic scaling of variance with rarity. We address this with a variance reduction framework based on lightweight importance splitting observers. These expose the model-property automaton to allow the construction of score functions for high performance algorithms. The confidence intervals defined for importance splitting make it appealing for SMC, but optimising its performance in the standard way makes distribution inefficient. We show how it is possible to achieve equivalently good results in less time by distributing simpler algorithms. We first explore the challenges posed by importance splitting and present an algorithm optimised for distribution. We then define a specific bounded time logic that is compiled into memory-efficient observers to monitor executions. Finally, we demonstrate our framework on a number of challenging case studies.

[23] (C) Exhaustive verification can quantify critical behaviour arising from concurrency in nondeterministic models. Rare events typically entail no additional challenge, but complex systems are generally untractable. Recent work on Markov decision processes allows the extremal probabilities of a property to be estimated using Monte Carlo techniques, offering the potential to handle much larger models. Here we present algorithms to estimate extremal rewards and consider the challenges posed by rarity. We find that rewards require a different interpretation of confidence and that reachability rewards require the introduction of an auxiliary hypothesis test. We show how importance sampling can significantly improve estimation when probabilities are low, but find it is not a panacea for rare schedulers.

[36] (J; accepted) We propose a new SMC technique based on CUSUM, an algorithm originally used in signal processing, that detects probability change at runtime on a single execution of a system. The principle is to monitor the execution at regular time intervals, and to perform Monte Carlo checks over the samples of the execution. The results of these checks are used to compute the CUSUM ratio, whose variation allows to detect a change of the probability measure of the system. We demonstrate the algorithm to detect failures in a Simulink model of a temperature controller. Computing the exact time at which failures may happen is then useful to schedule maintenance operations.

[42] (W) Many embedded and real-time systems have a inherent probabilistic behaviour (sensors data, unreliable hardware,...). In that context, it is crucial to evaluate system properties such as "the probability that a particular hardware fails". Such properties can be evaluated by using probabilistic model checking. However, this technique fails on models representing realistic embedded and real-time systems because of the state space explosion. To overcome this problem, we propose a verification framework based on *Statistical Model Checking*. Our framework is able to evaluate probabilistic and temporal properties on large systems modelled in SystemC, a standard system-level modelling language. It is fully implemented as an extension of the Plasma-lab statistical model checker. We illustrate our approach on a multi-lift system case study.

[27] (W) Stochastic Petri nets are commonly used for modeling distributed systems in order to study their performance and dependability. This report proposes a realization of stochastic Petri nets in SystemC for modeling large embedded control systems. Then statistical model checking is used to analyze the dependability of the constructed model. Our verification framework allows users to express a wide range of useful properties to be verified which is illustrated through a case study.

[25] (C: accepted) Transaction-level modeling with SystemC has been very successful in describing the behavior of embedded systems by providing high-level executable models, in which many of them have an inherent probabilistic behavior, i.e., random data, unreliable components. It is crucial to evaluate the quantitative and qualitative analysis of the probability of the system properties. Such analysis can be conducted by constructing a formal model of the system and using probabilistic model checking. However, this method is infeasible for large and complex systems due to the state space explosion. In this work, we demonstrate the successful use of *Statistical Model Checking* to carry out such analysis directly from large SystemC models and allows designers to express a wide range of useful properties. This work is going to presented at 17th IEEE High Assurance Systems Engineering Symposium in January, 2016.

6.3. Formal Models for Variability

Participants: Axel Legay, Rudolf Fahrenberg, Jin Hyun Kim.

This part of the report is more concerned with task 2. It studies variability aspects in the broad scope. As in the first year, we have decided to use the concept of product lines as a general framework to reason on the problematic.

The behaviour of a software system is often described in terms of its features, where each *feature* is a unit of functionality that adds value to the system. *Feature-oriented software development (FOSD)* is a software-development strategy that is based on feature decomposition and modularity. Features can be separate modules that are developed in isolation, allowing for parallel, incremental, or multi-vendor development of features. Feature orientation is particularly important in *software product lines*, where a family of related products is managed and evolved in terms of its features: a product line comprises a collection of mandatory and optional features, and individual products are derived by selecting among and integrating features from this feature set. A product line can be expressed as a single model, in which feature-specific behaviour is conditional on the presence of the feature in a product.

The downside of FOSD is that, although features are conceptualized, developed, managed, and evolved as separate concerns, they are not truly separate. They can interfere with each other, for example by trying to control the same variables, by issuing events that trigger other features, or by imposing conditions that suppress other features. Most of the early work on feature interactions focused on interactions that manifest themselves as logical inconsistencies, such as conflicting actions, nondeterminism, deadlock, invariant violation, or unsatisfiability. More recently, a more general definition of feature interaction has been introduced, in terms of a feature that is developed and verified to be correct in isolation but is found to behave differently when combined with other features, and it was shown how such *behaviour interactions* could be detected as a violation of bisimulation.

Another problem is that FTS models are monolithic models of full product lines. There is no means of modelling individual features and composing them into products or product-line models, or of specifying feature increments to an existing product-line model. As such, FTSs cannot be the mathematical basis for modelling technologies that support feature decomposition, composition, or incremental evolution of a product line.

6.3.1. Papers:

[11] (C) Featured Transition Systems (FTSs) is a popular representation for software product lines: an entire product line is compactly represented as a single transition-machine model, in which feature-specific behaviour is guarded by feature expressions that are satisfied (or not) by the presence or absence of individual features. In previous work, FTS models were monolithic in the sense that the modeller had to construct the full FTS model of the product line in its entirety. To allow for modularity of FTS models, we propose here a language for extending an existing FTS model with new features. We demonstrate the language using a running example and present results about the language's expressivity, commutativity of feature extensions, feature interactions, and resolution of such interactions.

[12] (C) We suggest a method for measuring the degree to which features interact in feature-oriented software development. To this end, we extend the notion of simulation between transition systems to a similarity measure and lift it to compute a behaviour interaction score in featured transition systems. We then develop an algorithm which can compute the degree of feature interactions in a featured transition system in an efficient manner.

6.4. Privacy and Security

Participants: Axel Legay, Fabrizio Biondi, Jean Quilbeuf, Thomas Given-Wilson, Sébastien Josse.

6.4.1. Information-Theoretical Quantification of Security Properties

This part of the work was not foreseen at the beginning of the action. It concerns security aspects, and more precisely quantifying privacy of data. This aspect is in fact central for SoS and all our algorithms developed for Tasks 4 and 5 should be adapted to solve a series of problems linked to privacy in interconnected object and dynamical environment. For now, we only studied the foundations.

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 informations 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 by combining this information with its knowledge of the system.

Armed with the produced output and the source code of the system, the attacker tries to infer the value of the secret. The quantitative analysis we implement computes with arbitrary precision the number of bits of the secret that the attacker will expectedly infer. This expected number of bits is the information leakage of the system.

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 who dont it is imperative to be able to distinguish between the ones leaking a very small amount of bits and the ones leaking a significant amount of bits, since only the latter are considered to pose a security vulnerability to the system.

Since black box security analyzes are immediately invalidated whenever an attacker gains information about the source code of the system, we assume that the attacker has a white box view of the system, meaning that it has access to the systems source code. This approach is also consistent with the fact that many security protocol implementations are in fact open source.

The scope of modern software projects is too large to be analyzed manually. For this reason we provide tools that can support the analyst and locate security vulnerabilities in large codebases and projects. We work with a variety of tools, including commercial software analysis tools being adapted with our techniques, and tools such as QUAIL developed here by our team.

We applied the leakage analysis provided by QUAIL to several case studies. Our case studies (voting protocol and smart grid coordination) have in common that a publicly disclosed information is computed from the secret of every participant in the model. In the voting example, the vote of a given voter is secret, but the number of votes for each candidates is public. Similarly, in the smart grid example, the consumption of one of the houses is secret, but the consumption of a whole quarter can be deduced. Qualitative analyses are either too restrictive or too permissive on these types of systems. For instance, non-interference will reject them as the public information depends on the secret. Declassification approaches will accept them, even if the number of voters or consumers is 2, in which case the secret can be deduced.

The development of better tools for quantitative security builds upon both theoretical developments in information theory, and development of the tools themselves. These often progress in parallel with each supporting the findings of the other, and increasing the demands and understanding upon each other.

6.4.1.1. Papers:

[34] (C; submitted) Systems dealing with confidential data may leak some information by their observable outputs. Quantitative information flow analysis provides a method for quantifying the amount of such information leakage. To avoid the high computational cost of exhaustive search, statistical analysis has been studied to estimate information leakage by analyzing only a small but representative subset of the system's behavior. In this paper we propose a new compositional statistical analysis method for quantitative information flow that combines multiple statistical analyses with static trace analysis. We use partial knowledge of the system's source code or specification, therefore improving both quality and cost of the analysis. The new method can optimize the use of weighted statistical analysis by performing it on components of the system and appropriately adapting their weights. We show this approach combined with the precision of trace analysis produces better estimates and narrower confidence intervals than the state of the art.

[38] (J) The quantification of information leakage provides a quantitative evaluation of the security of a system. We propose the usage of Markovian processes to model deterministic and probabilistic systems. By using a methodology generalizing the lattice of information approach we model refined attackers capable to observe the internal behavior of the system, and quantify the information leakage of such systems. We also use our method to obtain an algorithm for the computation of channel capacity from our Markovian models. Finally, we show how to use the method to analyze timed and non-timed attacks on the Onion Routing protocol.

[40] (C) Quantitative security analysis evaluates and compares how effectively a system protects its secret data. We introduce QUAIL, the first tool able to perform an arbitrary-precision quantitative analysis of the security of a system depending on private information. QUAIL builds a Markov Chain model of the system's behavior as observed by an attacker, and computes the correlation between the system's observable output and the behavior depending on the private information, obtaining the expected amount of bits of the secret that the attacker will infer by observing the system. QUAIL is able to evaluate the safety of randomized protocols depending on secret data, allowing to verify a security protocol's effectiveness. We experiment with a few examples and show that QUAIL's security analysis is more accurate and revealing than results of other tools.

[41] (C) Quantitative security techniques have been proven effective to measure the security of systems against various types of attackers. However, such techniques are based on computing exponentially large channel matrices or Markov chains, making them impractical for large programs. We propose a different approach based on abstract trace analysis. By analyzing directly sets of execution traces of the program and computing security measures on the results, we are able to scale down the exponential cost of the problem. Also, we are able to appy statistical simulation techniques, allowing us to obtain significant results even without exploring the full space of traces. We have implemented the resulting algorithms in the QUAIL tool. We compare their effectiveness against the state of the art LeakWatch tool on two case studies: privacy of user consumption in smart grid systems and anonymity of voters in different voting schemes.

[37] (C) In an election, it is imperative that the vote of the single voters remain anonymous and undisclosed. Alas, modern anonymity approaches acknowledge that there is an unavoidable leak of anonymity just by publishing data related to the secret, like the election's result. Information theory is applied to quantify this leak and ascertain that it remains below an acceptable threshold. We apply modern quantitative anonymity analysis techniques via the state-of-the-art QUAIL tool to the voting scenario. We consider different voting typologies and establish which are more effective in protecting the voter's privacy. We further demonstrate the effectiveness of the protocols in protecting the privacy of the single voters, deriving an important desirable property of protocols depending on composite secrets.

[39] (C) In recent years, quantitative security techniques have been providing effective measures of the security of a system against an attacker. Such techniques usually assume that the system produces a finite amount of observations based on a finite amount of secret bits and terminates, and the attack is based on these observations. By modeling systems with Markov chains, we are able to measure the effectiveness of attacks on non-terminating systems. Such systems do not necessarily produce a finite amount of output and are not necessarily based on a finite amount of secret bits. We provide characterizations and algorithms to define meaningful measures of security for non-terminating systems, and to compute them when possible. We also study the bounded versions of the problems, and show examples of non-terminating programs and how their effectiveness in protecting their secret can be measured.

6.4.2. Equivocation-based Security Measures for Shared-Key Cryptosystems

Ensuring privacy and security of communication is a fundamental concern of cyber-physical systems and handled by encryption. Information-theoretic reasoning allows the modelling of security properties via unconditional security. That is, information-theoretic approaches formalise security properties that do not rely upon unproven computational hardness results, and are not vulnerable to advances in computing hardware, software or theory. For example, such unconditional security guarantees are not weakened by quantum computers, mem-computers, or new mathematical discoveries.

Traditionally the strongest measure of the security of a system is *perfect secrecy* as proposed by Shannon. However, this relies upon having a large key that is used only once. In practice a measure of the security of cryptosystems that does not meet this requirement is more useful. To this end we presented *max-equivocation*, a measure of the maximum achievable security given the keys available. Indeed max-equivocation not only formalizes the best possible security, but also generalizes perfect secrecy.

Max-equivocation holds even when inputs to the systems (i.e. keys and messages) are not uniform. This corresponds to many real world scenarios, and indeed we have shown that existing approaches are non-optimal as they do not consider these perturbations in the inputs. We provide necessary and sufficient conditions for achieving max-equivocation, formalizing exactly when it can be achieved in practice.

We further generalize to consider scenarios where message spaces are not complete, i.e. there are messages that are invalid and could never be produced. This allows reasoning over (and contrasting with) many prior approaches as well as formalizing their strengths and weaknesses under max-equivocation.

The most common attack against such cryptosystems is to consider when the attacker sees a single (encrypted) message and tries to guess the content. This can be measured by the *vulnerability* of the system, i.e. the probability that the attacker will guess correctly the message. We formalize a *relative vulnerability* for when the attacker makes this guess under incorrect assumptions about the messages. We formalize that the attacker can never improve their chances at guessing the message with incorrect assumptions.

Now we consider what information the attacker can gain by observing the cryptosystem. We show that the encryption function alone reveals information about the possible message distributions to the attacker. In the worse case scenario an encryption function may admit only a single message distribution. Thus the construction of the encryption function should consider this and (when possible) admit many solutions.

Further we consider what the attacker can learn by observing the communication of a cryptosystem. We show that the attacker can learn the probability distribution over the ciphertexts (encrypted messages), and combined with the information from the encryption function converge upon a distribution for the messages. Again if the encryption function admits one solution then the attacker learns the exact message distribution. We show that even when a single solution will not be found, the attacker still converges upon a message distribution that can only improve their attacks.

In addition to formalizing how these attacks work, and thus how to protect against them when constructing cryptosystems, we also consider not sharing the encryption function as a mechanism to avoid the attacker exploiting it. We formalize how to still communicate effectively in this scenario, and the advantages and disadvantages of this approach.

We present several algorithms to demonstrate the practicality of the techniques. The algorithms to achieve max-equivocation consider the message distribution and compute an encryption function that achieves close to max-equivocation. Since these algorithms are tailored for the message distributions, they out perform generic algorithms. We also present algorithms that are able to perform well without revealing the entire encryption function, and thus revealing less information to the attacker and hindering cryptoanalysis.

Thus we show that unconditional security is not only more resistant to technology changes, but also can be formalised for many scenarios, and is achievable in practice.

6.4.2.1. Papers:

[29] (C, submitted) Recent work has presented max-equivocation as a measure of the resistance of a cryptosystem to attacks when the attacker is aware of the encoder function and message distribution. Here we consider the vulnerability of a cryptosystem in the one-try attack scenario when the attacker has incomplete information about the encoder function and message distribution. We show that encoder functions alone yield information to the attacker, and combined with inferable information about the ciphertexts, information about the message distribution can be discovered. We show that the whole encoder function need not be fixed or shared a priori for an effective cryptosystem, and this can be exploited to increase the equivocation over an a priori shared encoder. Finally we present two algorithms that operate in these scenarios and achieve good equivocation results, ExPad that demonstrates the key concepts, and ShortPad that has less overhead than ExPad.

[13], [28] (C; J, submitted) Preserving the privacy of private communication is a fundamental concern of computing addressed by encryption. Information-theoretic reasoning models unconditional security where the strength of the results is not moderated by computational hardness or unproven results. Perfect secrecy is often considered the ideal result for a cryptosystem, where knowledge of the ciphertext reveals no information about the message or key, however often this is impossible to achieve in practice. An alternative measure is the equivocation, intuitively the average number of message/key pairs that could have produced a given ciphertext. We show a theoretical bound on equivocation called max-equivocation and show that this generalizes perfect secrecy when achievable, and provides an alternative measure when perfect secrecy is not. We derive bounds for max-equivocation, and show that max-equivocation is achieved when the entropy of the ciphertext is minimized. We consider encryption functions under this new perspective, and show that in general the theoretical best is unachievable, and that some popular approaches such as Latin squares or Quasigroups are also not optimal. We present some algorithms for generating encryption functions that are practical and achieve 90 - 95% of the theoretical best, improving with larger message spaces.

6.4.3. Malware Classification via Deobfuscation and Behavioral Fingerprinting

A fundamental problem to guarantee the security of systems is to be able to discriminate between legitimate processes and processes with malicious behavior. Malicious software, or malware, has to be identified and prevented from executing on the system, and its actions reverted by a disinfection process. To be able to recognize and disinfect malware it is necessary to be able to extract a behavioral fingerprint or signature from a binary file, and to construct a database of such signatures for comparison. The signatures in the database have to be classified accoring to the malware's family and category, allowing the correct disinfection method to be deployed.

Automatic extraction of behavioral signatures in the form of temporal logical graphs or control flow graphs is a recent but very effective technique, and malware developers have already adapted malware compilation chains to include techniques to hinder reverse engineering and thus prevent the extraction of such signatures. These obfuscation techniques include the addition of obfuscated conditional statements leading to dead code, control flow flattening based on complex function like cryptographic hash functions, and source code virtualization on an embedded interpreter.

Consequently, deobfuscation has to be developed along with fingerprinting techniques to be able to effectively extract malware signatures. We are pushing the state of the art in both subjects, advancing generalized and targeted deobfuscation and deploying them on an innovative virtualization and malware fingerprinting tool.

Mixed Boolean Arithmetic (MBA) obfuscation is an obfuscation technique developed by Cloakware Inc. and deployed in obfuscating compilation chains for both legitimate code and malware. We have deployed state-of-the-art SMT solvers to evaluate their effectiveness against MBA-obfuscated conditionals and ascertained their limited effectiveness. So we have developed an algebraic simplification technique targeting the algebraic structure of MBA obfuscation, and proved such technique to be extremely effective, being able to deobfuscate statements in orders of magnitude less time than the time required to obfuscate them in the first place.

While the algebraic simplification technique is very effective against MBA obfuscation, it is completely tailored to MBA obfuscation. We instead explore a completely general method based on dynamic program synthesis. Synthesis algorithms, like the ones based on Reed-Muller expansion techniques, interrogate the target (in this case the obfuscated conditional) multiple times considering it as a black-box oracle, and synthesize the function expressed by the target frm the answers to such interrogation. We determined that synthesis is significantly more efficient than SMT solving in synthesizing the obfuscated function in a very compact form, and thus very promising as a generalized deobfuscation method.

While more targeted deobfuscation techniques are required to coutneract control flow flattening and virtualization, the deobfuscation of conditional statements is an important step for malware fingerprinting. We plan to use our tool to classify a large database of malware, producing an extensive database of malware signatures representing multiple versions and families of malicious code. Malware mining and evolution techniques can be deployed on such database to construct different signatures for unknown variants of similar malware, thus improving the effectiveness of the detection process.

6.4.3.1. Papers:

[30] (C, submitted) The obfuscation of conditional statements is a simple and efficient way to disturb the identification of the control flow graph of a program. Mixed Boolean arithmetics (MBA) techniques provide concrete ways to achieve this obfuscation of conditional statements. In this work, we study the effectiveness of automated deobfuscation of MBA obfuscation, using algebraic, SMT-based and synthesis-based techniques. 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 the deobfuscation complexity can be greatly reduced by algebraic simplification. We also consider synthesis-based deobfuscation and find it to be more effective than SMT-based deobfuscation. We discuss complexity and limits of all methods, and conclude that MBA obfuscation is not effective enough to be considered a reliable method for control flow or white-box obfuscation.

6.5. Energy-Centric Systems

Participants: Axel Legay, Uli Fahrenberg.

This part is concerned with Tasks 1 and 2. Mostly, we focus on quantifying properties of interconnected objects such as Cyber Physical Systems (CPS) (SoS and CPS share a lot of commonalities).

Energy and *resource management* problems are important in areas such as embedded systems or autonomous systems. They are concerned with the question whether a given system admits infinite schedules during which (1) certain tasks can be repeatedly accomplished and (2) the system never runs out of energy (or other specified resources). Formal modeling and analysis of such problems has attracted some attention in recent years.

6.5.1. Papers:

[18] (C; accepted) We define and study basic properties of *-continuous Kleene ω-algebras that involve a *-continuous Kleene algebra with a *-continuous action on a semimodule and an infinite product operation that is also *-continuous. We show that *-continuous Kleene ω-algebras give rise to iteration semiring-semimodule pairs, and that for Büchi automata over *-continuous Kleene ω-algebras, one can compute the associated infinitary power series.

[17] (C; accepted) Energy problems are important in the formal analysis of embedded or autonomous systems. Using recent results on *-continuous Kleene ω-algebras, we show here that energy problems can be solved by algebraic manipulations on the transition matrix of energy automata. To this end, we prove general results about certain classes of finitely additive functions on complete lattices which should be of a more general interest.

[15] (C; accepted) We develop a *-continuous Kleene ω-algebra of real-time energy functions. Together with corresponding automata, these can be used to model systems which can consume and regain energy (or other types of resources) depending on available time. Using recent results on *-continuous Kleene ω-algebras and computability of certain manipulations on real-time energy functions, it follows that reachability and Büchi acceptance in real-time energy automata can be decided in a static way which only involves manipulations of real-time energy functions.

6.6. Languages for composition

Participants: Axel Legay, Thomas Given-Wilson.

This part is concerned with Task 1, especially to describe the composition of complex systems, and to study expressivity of existing formalisms.

Contemporary cyber-physical systems are inherently constructed out of a variety of agents with communication and interaction forming a key role in the behaviour of the system as a whole. Traditional approaches to reasoning over a single computation or treating the system as a single agent prove unsatisfactory for understanding the capabilities, strengths, and weaknesses of such systems.

Since communication is a fundamental to such systems it is necessary to understand the role the communication primitives themselves play. There are many approaches to communication primitives, often chosen for their ability to easily represent desired behaviour. However, the formal properties of many implementations or chosen models have not been presented.

An alternative to formalising each possible model individually is to abstract away and reason over families of models based on their communication primitives. This allows keys results to be achieved in one model, and then generalised to the entire family, or transferred to other families based upon formal relations between these families. Thus making it possible for results to be easily applied to many models or systems without repeating significant effort.

6.6.1. Papers:

[20] (C), [32] (J; submitted) The expressiveness of communication primitives has been explored in a common framework based on the π-calculus by considering four features: synchronism (asynchronous vs synchronous), arity (monadic vs polyadic data), communication medium (shared dataspaces vs channel-based), and pattern-matching (binding to a name vs testing name equality vs intensionality). Here another dimension coordination is considered that accounts for the number of processes required for an interaction to occur. Coordination generalises binary languages such as π-calculus to joining languages that combine inputs such as the Join Calculus and general rendezvous calculus. By means of possibility/impossibility of encodings, this paper shows coordination is unrelated to the other features. That is, joining languages are more expressive than binary languages, and no combination of the other features can encode a joining language into a binary language. Further, joining is not able to encode any of the other features unless they could be encoded otherwise.

[33] (C; submitted) The expressiveness of communication primitives has been explored in a common framework by considering four features: synchronism, arity, communication medium, and pattern-matching. These all assume asymmetric communication between input and output primitives, however some calculi consider more symmetric approaches to communication such as fusion calculus and Concurrent Pattern Calculus. Symmetry can be considered either as allowing a mixture of input and output in an action or co-action, or as the unification of actions. By means of possibility/impossibility of encodings, this paper shows that: the action and co-action approach is related to or more expressive than many previously considered languages; and the unification approach is more expressive than some, but mostly unrelated to other languages.

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. Privacy

Participants: Axel Legay, Fabrizio Biondi, Jean Quilbeuf.

Privacy is a regional project whose objective is to quantify privacy of data. This includes, e.g., quantifying the anonymity of a voting protocol.

7.1.2. Variability

Participants: Axel Legay, Jin Hyun Kim, Louis-Marie Traonouez.

Variability is a regional project whose objective is to lift scheduling techniques to connected-objects. The main application of the project is Systems of Systems.

7.2. National Initiatives

7.2.1. ANR Malthy

Participants: Axel Legay, Rudolf Fahrenberg, Louis-Marie Traonouez.

The objective of this project is to study new models and techniques to reason on quantitative systems. We mainly focus on the composition of timed components in a dynamic setting.

7.2.2. BGLE SyS2Soft

Participants: Axel Legay, Thomas Given-Wilson, Cyrille Jegourel.

This national project studies various languages and techniques for quantitative systems.

7.3. European Initiatives

7.3.1. FP7 & H2020 Projects

7.3.1.1. ACANTO

Title: ACANTO: A Cyber physicAl social NeTwOrk using robot friends

Programm: H2020

Duration: February 2015 - August 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 Spa (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.'

7.3.2. Danse

Program: FP7

Project acronym: DANSE

Project title: Designing for Adaptability and evolutioN in System of systems Engineering

Duration: Octobre 2011 – March 2015

Coordinator: Offis

Abstract: Design and verification of Systems of Systems. We contributed by proposing the first verification engine for Heterogeneous SoS. For doing so, we have combined Plasma with Desyre that is a simulator for SoS described via the standardised FMI/FMU approach.

7.3.3. Meals

Program: Marie Curie

Project acronym: Meals

Project title: Mobility between Europe and Argentina applying Logics to Systems

Duration: Octobre 2012 – Octobre 2015

Coordinator: Germany (Saarbrucken) and Argentina (Corona)

Abstract: Collaborative action on the topic of quantitative systems

7.3.4. Sensation

Program: Fet ProActif

Project acronym: Sensation

Project title: Self Energy-Supporting Autonomous Computation

Duration: Octobre 2012 – Octobre 2015

Coordinator: Aalborg University

Abstract: Development of new results for energy-centric systems. We contributed by proposing new algorithms for rare-event simulation.

7.3.5. EMC2

Program: ARTEMIS

Project acronym: EMC2

Project title: Embedded Multi-Core systems for Mixed Criticality applications in dynamic and changeable real-time environments

Duration: mars 2014 – mars 2017

Coordinator: Infineon

Abstract: Large initiative on embadded systems and SoS. We will contribute with our expertise from DANSE and Sensation projects.

7.3.6. Collaborations with Major European Organizations

- Partner 1: Aalborg University, Computer Science, Denmark
- Statistical Model Checking, and Systems of Systems
- Partner 2: Rice University, Computer Science, USA
- Synthesis of components of Systems of Systems
- Partner 3: Namur University, Computer Science, Belgium
- Variability in software engineering
- Partner 4: Louvain University, Computer Science, Belgium
- Verification of Systems of Systems via Statistical Model Checking, especially train stations in collaboration with Alstom.
- Partner 5: Waterloo University, Computer Science, Canada
- Variability in Systems of Systems

7.4. International Initiatives

7.4.1. Visits of International Scientists

7.4.1.1. Internships

- Karin Quaas, PostDoc at Leipzig University
- Kim Larsen, Professor at Aalborg University
- Rafael Olochea, PhD student at Waterloo University
- Yusuke Yamamoto, Assistant Professor, Japan.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific events selection

- Axel Legay was PC member of MEMOCODE, FORMATS, FORMALIZE, SPLC, ASE, FACS, FORTE.
- Louis-Marie Traonouez was PC member of FORMATS
- Rudolf Fahrenberg was PC member of TACAS

8.1.2. Journal

8.1.2.1. Member of the editorial boards

Axel Legay is a member of the editorial board of the newly created journal for masterminding changes (FOMACS).

8.1.2.2. Reviewer - Reviewing activities

Axel Legay has been reviewer for TCS, FMSD, IandC.

8.1.3. Invited talks

- Axel Legay has been invited at the first EMSIG school, Copenhagen 2015.
- Uli Fahrenberg has been invited at Automatha 2015.

8.1.4. Leadership within the scientific community

Together with the University of Vannes and IRIT Toulouse, the team has initiated the first GT on systems of systems.

8.1.5. Scientific expertise

Axel Legay has been an evaluator for the National Science fund in Belgium. He is also a member of Inria evaluation committee.

8.1.6. Research administration

The members of the team reviewed numerous papers for numerous international conferences

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Software Verification

Axel Legay: Software Verification, 40 hours, Royal Holloway, University of London

Axel Legay: Modélisation et Vérification Formelle par Automates, 12 hours, University of Rennes 1.

Axel Legay: Méthodes d'analyse de risques, 28 hours, ENSIBS2 (Vannes)

Jean Quilbeuf: Modélisation et Vérification Formelle par Automates, TP 12 hours, University of Rennes 1.

8.2.2. Supervision

PhD: Mounir Chadli: Scheduling for Systems of Systems, Started on December 2014, Axel Legay

[1] HdR: Axel Legay, Contribution to Statistical Model Checking, November 2015

8.2.3. Juries

Axel Legay has been a member of the PhD jury of Raphael Michel (Namur).

8.3. Popularization

Together with Vannes University, the team has published three high-level description articles on systems of systems [5], [6], [7].

9. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. LEGAY. *Contributions to Statistical Model Checking*, Inria Rennes, November 2015, Habilitation à diriger des recherches, https://hal.inria.fr/tel-01244469

Articles in International Peer-Reviewed Journals

[2] A. COLOMBO, D. FONTANELLI, A. LEGAY, L. PALOPOLI, S. SEDWARDS. *Efficient customisable dynamic motion planning for assistive robots in complex human environments*, in "Journal of ambient intelligence and smart environments", September 2015 [*DOI :* 10.3233/AIS-150338], https://hal.inria.fr/hal-01239099

[3] P. D'ARGENIO, A. LEGAY, S. SEDWARDS, L.-M. TRAONOUEZ. *Smart Sampling for Lightweight Verification of Markov Decision Processes*, in "International Journal on Software Tools for Technology Transfer (STTT)", August 2015, vol. 17, n° 4, pp. 469-484 [*DOI :* 10.1007/s10009-015-0383-0], https://hal.inria.fr/hal-01088633

[4] A. DAVID, K. G. LARSEN, A. LEGAY, U. NYMAN, L.-M. TRAONOUEZ, A. WASOWSKI. *Real-Time Specifications*, in "Software Tools for Technology Transfer (STTT)", January 2015, vol. 17, n° 1, 29 p. [*DOI :* 10.1007/s10009-013-0286-x], https://hal.archives-ouvertes.fr/hal-01087799

Articles in Non Peer-Reviewed Journals

[5] A. LEGAY, J. QUILBEUF, F. OQUENDO. *Verifying Systems-of-Systems with Statistical Model Checking*, in "ERCIM News", 2015, n° 103, https://hal.inria.fr/hal-01242652

[6] F. OQUENDO, A. LEGAY, K. DRIRA. *GT SoS: Research Network on Trustworthy Software-intensive Systems-of-Systems*, in "ERCIM News", 2015, n° 102, https://hal.inria.fr/hal-01242651

[7] F. OQUENDO, A. LEGAY. *Formal Architecture Description of Trustworthy Systems-of-Systems with SosADL*, in "ERCIM News", 2015, n° 102, https://hal.inria.fr/hal-01242649

International Conferences with Proceedings

[8] V. C. NGO, J.-P. TALPIN, T. GAUTIER, L. BESNARD, P. LE GUERNIC. *Modular translation validation of a full-sized synchronous compiler using off-the-shelf verification tools (abstract)*, in "International Workshop on Software and Compilers for Embedded Systems", St Goar, Germany, ACM, June 2015, https://hal.inria.fr/hal-01148919

[9] V. C. NGO, J.-P. TALPIN, T. GAUTIER, P. LE GUERNIC. *Translation Validation for Clock Transformations in a Synchronous Compiler*, in "FASE - ETAPS 2015", London, United Kingdom, Springer, April 2015, https://hal.inria.fr/hal-01087795

[10] V. C. NGO, J.-P. TALPIN, T. GAUTIER. *Translation Validation for Synchronous Data-flow Specification in the SIGNAL Compiler*, in "International Conference on Formal Techniques for Distributed Objects, Components and Systems", Grenoble, France, Formal Techniques for Distributed Objects, Components, and Systems, Springer, June 2015, vol. 9039, pp. 66-80 [*DOI :* 10.1007/978-3-319-19195-9_5], https://hal.inria.fr/hal-01148901

Conferences without Proceedings

[11] J. M. ATLEE, S. BEIDU, U. FAHRENBERG, A. LEGAY. *Merging Features in Featured Transition Systems*, in "Proceedings of the 12th Workshop on Model-Driven Engineering, Verification and Validation co-located with ACM/IEEE 18th International Conference on Model Driven Engineering Languages and Systems", Ottawa, Canada, CEUR Workshop Proceedings, September 2015, vol. 1514, pp. 38-43, https://hal.inria.fr/hal-01237661

[12] J. M. ATLEE, U. FAHRENBERG, A. LEGAY. *Measuring Behaviour Interactions between Product-Line Features*, in "3rd IEEE/ACM FME Workshop on Formal Methods in Software Engineering", Firenze, Italy, May 2015, https://hal.inria.fr/hal-01237655

[13] F. BIONDI, T. GIVEN-WILSON, A. LEGAY. *Attainable Unconditional Security for Shared-Key Cryptosystems*, in "The 14th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (IEEE TrustCom-15)", Helsinki, Finland, August 2015, https://hal.inria.fr/hal-01192859

[14] F. BIONDI, A. LEGAY, J. QUILBEUF. *Comparative Analysis of Leakage Tools on Scalable Case Studies*, in "22nd International SPIN Workshop on Model Checking of Software", Stellenbosch, South Africa, August 2015 [*DOI :* 10.1007/978-3-319-23404-5_17], https://hal.inria.fr/hal-01241352

[15] D. CACHERA, U. FAHRENBERG, A. LEGAY. *An ω-Algebra for Real-Time Energy Problems*, in "35th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science", Bengaluru, India, December 2015, https://hal.inria.fr/hal-01237667

[16] X. DEVROEY, G. PERROUIN, M. CORDY, P.-Y. SCHOBBENS, P. HEYMANS, A. LEGAY. *State machine flattening, a mapping study and tools assessment*, in "8th IEEE International Conference on Software Testing, Verification and Validation", Graz, Austria, April 2015 [*DOI :* 10.1109/ICSTW.2015.7107408], https://hal.inria.fr/hal-01242787

[17] Z. ESIK, U. FAHRENBERG, A. LEGAY. *-Continuous Kleene ω-Algebras for Energy Problems*, in "Proceedings Tenth International Workshop on Fixed Points in Computer Science", Berlin, Germany, EPTCS, September 2015, vol. 191, pp. 48-59, https://hal.inria.fr/hal-01237653

[18] Z. ESIK, U. FAHRENBERG, A. LEGAY. *-Continuous Kleene ω-Algebras*, in "Developments in Language Theory - 19th International Conference", Liverpool, United Kingdom, Lecture Notes in Computer Science, July 2015, vol. 9168, pp. 240-251, https://hal.inria.fr/hal-01237648

[19] U. FAHRENBERG, A. LEGAY. *Partial Higher-Dimensional Automata*, in "6th Conference on Algebra and Coalgebra in Computer Science", Nijmegen, Netherlands, LIPIcs, June 2015, vol. 35, https://hal.inria.fr/hal-01237643

[20] T. GIVEN-WILSON, A. LEGAY. *On the Expressiveness of Joining*, in "8th Interaction and Concurrency Experience (ICE 2015)", Grenoble, France, June 2015, https://hal.inria.fr/hal-01152456

[21] C. JEGOUREL, A. LEGAY, S. SEDWARDS, L.-M. TRAONOUEZ. *Distributed Verification of Rare Properties using Importance Splitting Observers*, in "Proceedings of the 15th International Workshop on Automated Verification of Critical Systems (AVoCS 2015)", Edinburgh, United Kingdom, September 2015, vol. 72, https://hal.inria.fr/hal-01238982

[22] Y. KAWAMOTO, T. GIVEN-WILSON. *Quantitative Information Flow for Scheduler-Dependent Systems*, in "The 13th International Workshop on Quantitative Aspects of Programming Languages and Systems (QAPL 2015)", London, United Kingdom, Electronic Proceedings in Theoretical Computer Science, April 2015, forthcoming, https://hal.inria.fr/hal-01114778

[23] A. LEGAY, S. SEDWARDS, L.-M. TRAONOUEZ. *Estimating Rewards & Rare Events in Nondeterministic Systems*, in "Proceedings of the 15th International Workshop on Automated Verifica-

tion of Critical Systems (AVoCS 2015)", Edinburgh, United Kingdom, September 2015, vol. 72 [*DOI :* 10.14279/TUJ.ECEASST.72.1023], https://hal.inria.fr/hal-01239051

[24] A. LEGAY, L.-M. TRAONOUEZ. *Statistical Model Checking of Simulink Models with Plasma Lab*, in "Fourth International Workshop on Formal Techniques for Safety-Critical Systems", Paris, France, November 2015, https://hal.archives-ouvertes.fr/hal-01241249

[25] V. C. NGO, A. LEGAY, J. QUILBEUF. *Statistical Model Checking for SystemC Models*, in "High Assurance Systems Engineering Symposium", Orlando, Florida, United States, January 2016, https://hal.inria.fr/hal-01238162

Research Reports

[26] A. ARNOLD, B. MASSIMO, F. ALBERTO, M. MARCO, V. SENNI, A. LEGAY, J. QUILBEUF, E. CHISTOPH. *Statistical Model Checking of Systems of Systems: An Industrial Approach*, Inria, December 2015, n^{o} RR-8828, https://hal.inria.fr/hal-01242864

[27] V. C. NGO, A. LEGAY. *Dependability Analysis of Control Systems using SystemC and Statistical Model Checking*, Inria Rennes - Bretagne Atlantique ; Inria, July 2015, n^{o} RR-8762, https://hal.archives-ouvertes.fr/hal-01180996

Other Publications

[28] F. BIONDI, T. GIVEN-WILSON, A. LEGAY. *Attainable Unconditional Security for Shared-Key Cryptosystems*, November 2015, working paper or preprint, https://hal.inria.fr/hal-01233185

[29] F. BIONDI, T. GIVEN-WILSON, A. LEGAY. *On the Attacker's Knowledge in Shared-Key Cryptosystems*, December 2015, working paper or preprint [*DOI :* 10.1145/1235], https://hal.inria.fr/hal-01241374

[30] F. BIONDI, S. JOSSE, A. LEGAY, T. SIRVENT. *Effectiveness of Synthesis in Concolic Deobfuscation*, December 2015, working paper or preprint, https://hal.inria.fr/hal-01241356

[31] M. CHADLI, J. H. KIM, A. LEGAY, L.-M. TRAONOUEZ, S. NAUJOKAT, B. STEFFEN. *A Model-Based Framework for the Specification and Analysis of Hierarchical Scheduling Systems*, December 2015, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01241681

[32] T. GIVEN-WILSON, A. LEGAY. *On the Expressiveness of Coordination*, December 2015, working paper or preprint, https://hal.inria.fr/hal-01241647

[33] T. GIVEN-WILSON, A. LEGAY. *On the Expressiveness of Symmetry*, December 2015, working paper or preprint, https://hal.inria.fr/hal-01241839

[34] Y. KAWAMOTO, F. BIONDI, A. LEGAY. *Combining Static and Statistical Approaches to Quantitative Information Flow*, December 2015, working paper or preprint, https://hal.inria.fr/hal-01241360

[35] J. H. KIM, A. LEGAY, L.-M. TRAONOUEZ, M. ACHER, S. KANG. *A Formal Modeling and Analysis Framework for Software Product Line of Preemptive Real-Time Systems*, October 2015, Publication acceptée en temps que poster/papier court à la conférence SAC 2016, section software engineering, https://hal.archives-ouvertes.fr/hal-01241673

[36] A. LEGAY, L.-M. TRAONOUEZ. *Statistical Model Checking with Change Detection*, September 2015, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01242138

References in notes

[37] F. BIONDI, A. LEGAY. *Quantitative Anonymity Evaluation of Voting Protocols*, in "12th International Conference on Software Engineering and Formal Methods", Grenoble, France, September 2014, https://hal.inria.fr/hal-01088188

[38] F. BIONDI, A. LEGAY, P. MALACARIA, A. WASOWSKI. *Quantifying Information Leakage of Randomized Protocols*, in "Theoretical Computer Science", 2014, pp. 68 - 87 [*DOI :* 10.1007/978-3-642-35873-9_7], https://hal.inria.fr/hal-01088193

[39] F. BIONDI, A. LEGAY, B. F. NIELSEN, P. MALACARIA, A. WASOWSKI. *Information Leakage of Non-Terminating Processes* , in "IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science", Delhi, India, December 2014 [*DOI :* 10.4230/LIPICS.FSTTCS.2014.517], https://hal.inria.fr/hal-01086879

[40] F. BIONDI, A. LEGAY, L.-M. TRAONOUEZ, A. WASOWSKI. *QUAIL: A Quantitative Security Analyzer for Imperative Code*, in "Computer Aided Verification - 25th International Conference", Saint Petersburg, Russia, Lecture Notes in Computer Science, Springer, July 2013, vol. 8044, pp. 702 - 707 [*DOI :* 10.1007/978-3-642-39799-8_49], https://hal.archives-ouvertes.fr/hal-01087804

[41] F. BIONDI, J. QUILBEUF, A. LEGAY. *Information Leakage by Trace Analysis in QUAIL*, November 2014, working paper or preprint, https://hal.inria.fr/hal-01088208

[42] V. C. NGO, A. LEGAY, J. QUILBEUF. *Dynamic Verification of SystemC with Statistical Model Checking*, Inria Rennes - Bretagne Atlantique, équipe ESTASYS ; Inria, October 2014, n° RR-8644, 25 p. , https://hal.inria.fr/hal-01089742