ISSN 2186-7437

NII Shonan Meeting Report

No. 2015-16

Logical and Verification Methods in Security and Privacy

Marco Gaboardi, University of Dundee
Vivek Nigam, Federal University of Paraiba
Tachio Terauchi, Japan Advanced Institute of Science and Technology

October 26–29, 2015

National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-Ku, Tokyo, Japan

Logical and Verification Methods in Security and Privacy

Marco Gaboardi
University of Dundee

Vivek Nigam
Federal University of Paraiba

Tachio Terauchi
Japan Advanced Institute of Science and Technology

Report number: 2015-16
Date of the seminar: October 26 - 29, 2015

Description of the Meeting

Logic and verification methods are standard tools useful for reasoning in a formal way about programs. Tools and techniques based on logic have been used to prove software correct with respect to different requirements: functional correctness, resource consumption, integrity of data, etc. An area where the usefulness of these methods was observed early on is the one of security and privacy. Formal logic methods have been used for formally specifying security and privacy policies and models, and for verifying that programs respect them.

Recently however traditional notions of security and privacy have been extended to take into account new refined aspects of programs. In particular, we assisted at the extension of traditional security properties to notions that are quantitative or probabilistic in nature, and to notions that require complex policies models based on different notions of capabilities or authorizations. Quantitative information flows, differential privacy and distance-bounding protocols are three important examples of such security properties. Quantitative information flows is a natural quantitative extension of the traditional information flow theory where the information that is leaked by a program is measured in terms of notions like min-entropy and g-leakage. Differential privacy is a strong statistical notion of data privacy requiring that the result of a data analysis is probabilistically almost the same in the case the individual participate in the data or not. Distance-bounding protocols are cryptographic protocols where it is critical to control also the physical distance between the parties involved. These three properties are emblematic examples of the kind of security and privacy requirements that software is required today to ensure.

Verifying and reasoning in a formal way about this kind of security and privacy properties require extensions of traditional techniques that are able at the same time of describing more refined formal models - accounting for the quantities and/or capabilities specific to these properties - and more refined proof techniques designed specifically for these formal models. In order to deal with this supplementary complexity of the security and privacy models, several techniques have enriched traditional logical methods with quantitative informations - like probabilities, timestamps, etc - or authentication mechanisms - like capabilities, labels, etc. These techniques have provided new theoretical foundations and practical successful tools.

Part of this success is certainly to ascribe to the advancements in the technology of SMT solvers and interactive proof assistants. These advancements have permitted to reach results that only few

years ago seemed impossible to reach. SMT solvers and proof assistants can be nowadays easily integrated in tools for security and privacy analysis, and serve as basic component for the resolution of numeric and/or symbolic constraints relative to quantities and/or capabilities. However, the development of more and more specialized techniques require also a more and more customized use of these tools. For this reason it is important to promote the interaction between the members of the security and privacy community and the community working on verification tools.

The overall goal of the meeting was to foster the discussion between researchers in academia and industry that are working in different areas of security and privacy, logic and verification. The common ground between the different participants was the use of logic and verification methods for formally reasoning about the different aspects of security and privacy. On the application side, the discussion was around the different tools that are needed to reason about traditional and quantitative notion of security and privacy. On the theoretical side instead, the discussion was around common foundations for the different aspects of security and privacy. A further goal of the meeting was exploring the applicability of the most recent techniques developed in the setting of security and privacy to problem in different research areas.

Participants were from different research areas of security and privacy such as:

- Quantitative information flow
- Differential Privacy
- Distance-bounding protocols
- Information flow controls
- Access control policies
- Anonymization
- Obfuscation
- Capabilities models
- Malware detection

Participants had also specific expertise in logic and verification methods as:

- Relational Logics
- Linear type systems
- Coalgebraic methods
- Higher order verification methods
- Dependent type systems
- Abstract Interpretation
- Proof assistants
- SMT solvers

The meeting was a great occasion for exchange among researchers representing the different research areas and communities. The expected outcomes of the meeting are new collaborations between participants traditionally working in separate research areas.

Program

Monday October 26

9:00 Tutorial - Andre Scedrov Multiset rewriting with dense time and the analysis of cyber-physical security protocols.

10:10 Break

10:50 Research Problems - Carolyn Talcott Trust and Security Challenges for Networked Distributed Cyber-Physical Agent Systems; Yusuke Kawamoto Combining Static and Statistical Approaches to Quantitative Information Flow; Limin Jia Challenges in Engineering a Provably Secure Hypervisor Framework.

12:00 Lunch and Group Photo.

14:00 Tutorial - Dusko Pavlovic Cyber-physical security in actor networks.

15:00 Advertisement Talks - Vivek Nigam Timed Intruder Models; Sergio Maffeis Automated Testing of Browser Security Policies.

15:30 Break

16:30 Talk - Carroll Morgan Greatest pre-uncertainties for hyperGCL: a backwards semantics for abstract HMMs.

17:00 Open Panel Discussion.

Tuesday October 27

9:00 Tutorial - Bart Jacobs Attribute-based authentication in practice.

10:10 Break

10:50 Research Problems - Silvia Ghilezan Types in access control and privacy; Stephen Chong Knowledge and Effect: A Logic for Reasoning about Confidentiality and Integrity Guarantees; Dave Sands Language-based Data Minimization.

12:00 Lunch

14:00 Tutorial - Sbastien Gambs Inference attacks in location data.

15:00 Advertisement Talks - Ichiro Hasuo Kleisli Simulations, for Quantitative Verification (in General) and Probabilistic Anonymity (in Particular); Takeuti Izumi Logical system for negligible probability.

15:30 Break

16:30 Talk - Martin Hofmann GuideForce: type-based enforcement of secure coding guidelines.

17:00 Open Panel Discussion

Wednesday October 28

9:00 Tutorial - Gergei Bana Computationally Sound Security Analysis with First Order Logic An Introduction to the Computationally Complete Symbolic Attacker Based on Indistinguishability.

10:10 Break

10:50 Advertisement Talks - Mitsuhiro Okada French-Japanese cybersecurity framework (with special focus on formal methods); David Baelde Partial Order Reduction for Security Protocols: Improving Automated Trace Equivalence Checking in the Symbolic Model; Lucca Hirschi Automatic Verification of Privacy Protection for Unbounded Sessions.

11:30 Tool Overview - Deian Stefan Building Least Privileged Web Applications with Node.js.

12:00 Lunch

14:00 Tutorial - Deepak Garg CostIt: Using dependent types and co-monads for incremental complexity analysis.

15:00 Advertisement Talks - Ugo Dal Lago On Equivalences, Metrics, and Polynomial Time; Sin-ya Katsumata From DCC to local DCC.

15:30 Break

16:30 Tool Demo - Alejandro Russo Two can keep a secret if one of them uses Haskell.

17:00 Open Panel Discussion.

Thursday October 29 9:00 Tutorial - Justin Hsu An introduction to language-based techniques for verifying differential privacy.

10:10 Break

10:50 Research Problems - Roberto Giacobazzi Towards systematic code obfuscation (theory and practice).

11:30 Conclusion

12:00 Lunch

Abstracts

Stephen Chong - Harvard University - Cryptographic Enforcement of Language-Based Erasure Information erasure is a formal security requirement that stipulates when sensitive data must be removed from computer systems. In a system that correctly enforces erasure requirements, an attacker who observes the system after sensitive data is required to have been erased cannot deduce anything about the data. Practical obstacles to enforcing information erasure include: (1) correctly determining which data requires erasure; and (2) reliably deleting potentially large volumes of data, despite untrustworthy storage services.

I present a novel formalization of language-based information erasure that supports cryptographic enforcement of erasure requirements: sensitive data is encrypted before storage, and upon erasure, only a relatively small set of decryption keys needs to be deleted. This cryptographic technique has been used by a number of systems that implement data deletion to allow the use of untrustworthy storage services. However, these systems provide no support to correctly determine which data requires erasure, nor have the formal semantic properties of these systems been explained or proven to hold. We address these shortcomings. Specifically, we study a programming language extended with primitives for public-key cryptography, and demonstrate how information-flow control mechanisms can automatically track data that requires erasure and provably enforce erasure requirements even when programs employ cryptographic techniques for erasure.

This is joint work with Aslan Askarov, Scott Moore, and Christos Dimoulas, and was presented at the 28th IEEE Computer Security Foundations Symposium in July 2015.

Bart Jacobs - Radboud University - Attribute-based authentication in practice Traditionally authentication means proving who you are. But in practice, a transaction often requires proving certain properties about yourself (called attributes), like what your (email)address is, whether you're over 21, or whether you're a US-citizen. Such attributes allow flexible, privacy-friendly, proportional authentication. This talk describes the (cryptographic) basis of the IRMA system that implements these ideas, and also discusses the societal aspects, especially related to privacy. See www.irmacard.org

Yusuke Kawamoto - AIST - Combining Static and Statistical Approaches to Quantitative Information Flow Quantitative information flow analysis has been studied to model and quantify the amount of secret information leaked by observable output of a system. 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 work 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. This is joint work with Fabrizio Biondi and Axel Legay.

Deepak Garg - CostIt: Refinement types for incremental complexity analysis The problem of reasoning about the relative use of resources by two similar programs arises in many diverse applications ranging from proving the timing leaks in crypto implementations, proving the benefits of compiler optimizations and establishing the benefits of incremental computation. CostIt is a type-theory aimed at such analysis. In its current form, CostIt targets analysis of the asymptotic cost of incremental computation. The primary insight so far is that lightweight refinement types, fine-grained dependence analysis (from literature on information flow control) and co-monads together provide sufficient expressiveness to reason about relational resource consumption, at least for incremental complexity analysis. At the same time, the resulting type-theory is simple and easy to apply. Step-indexed logical relations can be used to prove soundness of the type-theory relative to a model with computational resources.

Andre Scedrov – Multiset Rewriting with Dense Times in the Analysis of Cyber-Physical Security Protocols Many security protocols rely on the assumptions on the physical properties in which its protocol sessions will be carried out. For instance, Distance Bounding Protocols take into account the round trip time of messages and the transmission velocity to infer an upper bound of the distance between two agents. We classify such security protocols as cyber-physical. Time plays a key role in design and analysis of many of these protocols. We investigate the foundational diff erences and the impacts on the analysis when using models with discrete time and models with dense time. We show that there are attacks that can be found by models using dense time, but not when using discrete time. We illustrate this with a novel attack that can be carried out on most distance bounding protocols. In this attack, the prover exploits the execution delay of instructions during one clock cycle to convince the verifier that the prover is in a location diff erent from its actual position. We propose a model for specifying cyber-physical security protocols which

extends Multiset Rewriting with dense times. We introduce Circle-Configurations and show that they can be used to symbolically solve the reachability problem for our model. Finally, we show that for the important class of balanced transition systems the reachability problem is PSPACE-complete. This is joint work with Max Kanovich, Tajana Ban Kirigin, Vivek Nigam, and Carolyn Talcott.

Carolyn Talcott – Computing devices and Cyber-Physical Systems are everywhere There are exciting opportunities for new infrastructure, medical, transportation, entertainment and other amazing apps. This also means scary responsibilites to make systems that are usable, safe, responsive, trustworthy. We present an overview of limitations and desiderate for NCPAS. We then discuss challenges in balancing need for autonomy, opportunistic collaboration, building trust, and providing adequate security.

Ichiro Hasuo – Kleisli Simulations, for Quantitative Verification (in General) and Probabilistic Anonymity (in Particular) The theory of forward and backward simulations by Lynch and Vaandrager has seen its categorical generalization, in which simulations are identified with (op)lax coalgebra homomorphisms in a Kleisli category. At the heart of the generalization is that the notion of branching—divergence, nondeterminism, probability, etc.—is parametrized as a monad. This readily yields as an instance a notion of quantitative simulation, formulated with matrices and hence amenable to search by linear programming. The resulting quantitative verification method has been applied to probabilistic notions of anonymity, too.

David Baelde – Partial Order Reduction for Security Protocols: Improving Automated Trace Equivalence Checking in the Symbolic Model I present recent work with Stphanie Delaune and Lucca Hirschi.

Security protocols are concurrent processes that communicate using cryptography with the aim of achieving various security properties. Recent work on their formal verification has brought procedures and tools for deciding trace equivalence properties (e.g., anonymity, unlinkability, vote secrecy) for a bounded number of sessions. However, these procedures are based on a naive symbolic exploration of all traces of the considered processes which, unsurprisingly, greatly limits the scalability and practical impact of the verification tools.

We overcome this difficulty by developing partial order reduction techniques for the verification of security protocols. We provide reduced transition systems that optimally eliminate redundant traces, and which are adequate for model-checking trace equivalence properties of protocols by means of symbolic execution. We have implemented our reductions in the tool Apte, and demonstrated that it achieves the expected speedup on various protocols.

Sébastien Gambs – Inference attacks on location data The advent of personal devices equipped with positioning, computational and communication capabilities, such as smartphones, has led to the large scale collection of the mobility data of individuals and the emergence of Location-Based Services (LBSs), which are personalized according to the position of the users. Examples of innovative LBSs include the search for neighboring services around the user, carpooling application dynamically matching a driver with a potential passenger or real-time traffic monitoring based on information sensed through users' smartphones, just to name a few. However among all the personal data, learning the location of an individual is one of the greatest threat against his

privacy. In particular, the mobility data of an individual can be used to learn the points of interests characterizing his mobility such as his home and place of work, to predict his past, current and future locations or even to discover his social network. In this talk, I will illustrate these risks by demonstrating how it is possible through inference attacks working on mobility traces to deduce other type of personal data. I will also discuss the protection mechanisms that can be used to mitigate these risks and to build privacy-preserving LBSs.

Alejandro Russo – Two Can Keep a Secret, If One of Them Uses Haskell For several decades, researchers from different communities have independently focused on protecting confidentiality of data. Two distinct technologies have emerged for such purposes: Mandatory Access Control (MAC) and Information-Flow Control (IFC)the former belonging to operating systems (OS) research, while the latter to the programming languages community. These approaches restrict how data gets propagated within a system in order to avoid information leaks. In this scenario, Haskell plays a unique privileged role: it is able to protect confidentiality via libraries. This pearl presents a monadic API which statically protects confidentiality even in the presence of advanced features like exceptions, concurrency, and mutable data structures. Additionally, we present a mechanism to safely extend the library with new primitives, where library designers only need to indicate the read and write effects of new operations.

Vivek Nigam – Timed Intruder Models The Dolev-Yao intruder has become the standard intruder model for protocol security verification. Many properties and tools have been built adopting this model. However, for some types of protocols, such as Cyber-Physical security protocols, the Dolev-Yao intruder is not suitable as it does not take into account, for instance, transmission delays.In this talk, we investigate timed intruder models describing our initial efforts in formalizing this intruder and using existing techniques to check in an automated fashion properties of Cyber-Physical security protocols.

Silvia Ghilezan – Types in access control and privacy Types have gained an important role in the analysis of formal systems. A type system splits elements (terms) of the language, into sets, called types, and proves absence of certain undesired behaviours. In programming languages, types represent a well-established technique to ensure program correctness. We present and overview of type systems for linked documents, linked data and communication-centered calculi. We then discuss the role of types in security, access control and privacy issues in these settings.

Limin Jia – Challenges in Engineering a Provably Secure Hypervisor Framework In the past decade, a lot of progress has been made in constructing high-assurance software. For instance, fully verified kernels have been shown to be possible. However, challenges remain in constructing provably secure software infrastructure from scratch. We examine these challenges through our experiences of constructing a provably secure hypervisor framework.

Shinya Katsumata – From DCC to local DCC Abadi et al.'s dependency core calculus (DCC) is a calculus that can host various information flow analysis. The main feature of DCC is the security level monad T_l t, which denotes the type of values of a type t at a security level l. The noninterference result was proven with a concrete denotational semantics of DCC.

In this study, we propose a categorical structure that generalises the denotational semantics of DCC. We introduce a new calculus called local DCC, which syntactically embodies the proposed categorical structure. Local DCC contains DCC as a subcalculus, and decomposes the security level monad into an adjunction. We then give a sound categorical interpretation of local DCC.

Gergei Bana - Computationally Sound Security Analysis with First Order Logic An Introduction to the Computationally Complete Symbolic Attacker Based on Indistinguishability The computationally complete symbolic attacker is a technique that was created as a way to find all possible attacks a probabilistic polynomial time adversary can carry out on a protocol, using symbolic methods. In this talk, we first briefly review the idea of verifying complexity-theoretic security guarantees with symbolic techniques, and mention attempts of various research groups to achieve this goal. We then present the elements of our computationally complete symbolic attacker based on indistinguishability, and show how convenient it is to formalize in this framework standard complexity-theoretic hardness assumptions and cryptographic security notions such as the DDH assumption, CPA, CCA security, or unforgeability. Finally we indicate what proofs we have carried out so far for anonymity, real-or-random secrecy, agreement and authentication, and present some attacks we detected with this technique.

Ugo Dal Lago - On equivalences, metrics and polynomial time Interactive behaviors are ubiquitous in modern cryptography, but are also present in lambda-calculi, in the form of higher-order constructions. Traditionally, however, typed lambda-calculi simply do not fit well into cryptography, being both deterministic and too powerful as for the complexity of functions they can express. We study interaction in a lambda-calculus for probabilistic polynomial time computable functions. In particular, we show how notions of context equivalence and context metric can both be characterized by way of traces when defined on linear contexts. We then give evidence on how this can be turned into a proof methodology for computational indistinguishability, a key notion in modern cryptography. We also hint at what happens if a more general notion of a context is used.

Lucca Hirschi - Automatic Verification of Privacy Protection for Unbounded Sessions I will present a new sound approach to check unlinkability and anonymity for an unbounded number of sessions. This is possible for a large class of 2-agents protocols. I identify two sufficient conditions implying unlinkability and anonymity that are much easier to check. Very often, the tool ProVerif is able to check them automatically. This allows for the first proofs for several protocols (e.g., Epassport, Hash-Lock) and finding new flaws in others (e.g., Epassport protocols, LAK).

Deian Stefan - Building Least Privileged Web Applications with Node.js Modern web applications handle user-specific, often sensitive, information. Unfortunately, protecting user data is notoriously difficult todayweb frameworks do not provide a way for declaring and enforcing application-specific security policies. In response, developers often specify and enforce security policy in an ad hoc fashion (e.g., by strewing security checks throughout the codebase). Recent headlines alone serve to highlight that this is not workingweb applications are plagued by privacy leaks. In this talk, I will describe ESpectro, a new framework for building least-privileged Node.js applications. ESpectro provides developers with libraries for compartmentalizing applications and declaring high-level security policies. ESpectro then enforces these policies on the unchanged application code, assuming it to have been compromised, by employing application-level virtualization.

To demonstrate the generality of application-level virtualization, we describe how other web security architectures, including, OKWS, Passe, and Hails can be implemented as libraries atop ESpectro.

Sergio Maffeis - Automated Testing of Browser Security Policies The security of the client side of a web application relies on browser features such as cookies, the same-origin policy and HTTPS. As the client side grows increasingly powerful and sophisticated, browser vendors have stepped up their offering of security mechanisms which can be leveraged to protect it. These are often introduced experimentally and informally and, as adoption increases, gradually become standardised (e.g., CSP, CORS and HSTS). Considering the diverse land- scape of browser vendors, releases, and customised versions for mobile and embedded devices, there is a compelling need for a systematic assessment of browser security.

This talks gives a brief introduction and demo of BrowserAudit, our tool for testing that a deployed browser enforces the guarantees implied by the main standardised and experimental security mechanisms. It includes more than 400 fully automated tests that exercise a broad range of security features, helping web users, application developers and security researchers to make an informed security assessment of a deployed browser. We have validated BrowserAudit by discovering both fresh and known security bugs in major browsers.

Justin Hsu - An introduction to language-based techniques for verifying differential privacy The differential privacy property has been intensively studied throughout computer science since its introduction. In particular, differential privacy is an intriguing property for formal verification, since it is both a (1) relational property and a (2) probabilistic property. I will discuss the composition principles that make differential privacy quite tractable for formal verification, and survey how these principles underlie the three main language-based approaches: (1) runtime verification, (2) linear type systems, and (3) Hoare logics.

Dusko Pavlovic – Cyber physical computation and security in Actor Networks I will present the Actor Network (ANt) model of cyber physical system, and the underlying motivation and conceptual analysis of cyber physical computation, as traced back to von Neumann and his models of universal computer with sensors and actuators. The logical specification languag developed for ANts will also be introduced, and several examples will be derived in it, demonstrating the reasoning about noniterference and about authenticity in multi-channel protocols in general.

Carroll Morgan – Greatest pre-uncertainties for hyperGCL: a backwards semantics for abstract HMMs Quantitative Information Flow (QIF) strives to measure the amount of information leaked by a computational process (eg a channel, or a computer program). Originally (Shannon) the "leak" was considered to be a good thing, the amount of information a channel successfully conveys; nowadays, in cyper-security, leaks are considered to be bad things, information that escapes out control and comes to the attention of an attacked.

Either way, measuring it is important.

The correspondence above, i.e. channels with programs, is instructive – but it can be taken too far. In particular, the assumption that what's good for media (Shannon Entropy) is also good for programs has been shown to be false. As a result, new kinds of entropy have been proposed, and new computational/denotational models designed within which they can operate.

This talk describes one of those, and introduces hyper-distributions, ie distributiond of distributions, that can conventiently be used to construct a monadic semantics based on Hidden Markov Models that is suitable for QIF analysis of real programs.

Participant List

1. Aslan Askarov, Aarhus University
2. David Baelde, LSV, ENS Cachan
3. Gergei Bana, INRIA Paris-Rocquencourt
4. Stephen Chong Harvard University
5. Marco Gaboardi, University of Dundee
6. Sbastien Gambs, Universit de Rennes 1 Inria
7. Deepak Garg, Max Planck Institute for Software Systems
8. Silvia Ghilezan, University of Novi Sad
9. Roberto Giacobazzi, University of Verona and IMDEA Software
10. Ichiro Hasuo, The University of Tokyo
11. Martin Hofmann, LMU Munich
12. Lucca Hirschi, LSV, ENS Cachan
13. Justin Hsu, University of Pennsylvania
14. Bart Jacobs, Radboud University Nijmegen
15. Limin Jia, Carnegie Mellon University
16. Shin-ya Katsumata, RIMS/ Kyoto University
17. Yusuke Kawamoto, INRIA and LIX Ecole Polytechnique
18. Ugo Dal Lago, University of Bologna and INRIA
19. Sergio Maffeis, Imperial College London
20. Carrol Morgan, UNSW, Australia
21. Vivek Nigam, Federal University of Paraiba
22. Mizuhito Ogawa, JAIST
23. Mitsuhiro Okada, Keio University
24. Dusko Pavlovic, University of Hawaii at Manoa

25. Alejandro Russo, Chalmers
26. David Sands, Chalmers University of Technology
27. Andre Scedrov, University of Pennsylvania
28. Deian Stefan, Stanford University
29. Kohei Suenaga, Kyoto University
30. Izumi Takeuti, AIST
31. Carolyn Talcott, SRI International
32. Tachio Terauchi, Japan Advanced Institute of Science and Technology