Sound and Complete Computational Interpretation of Symbolic Hashes in the Standard Model

Flavio D. Garcia, Peter Van Rossum

Sound and Complete Computational Interpretation of Symbolic Hashes in the Standard Model

Flavio D. Garcia and Peter van Rossum

Institute for Computing and Information Sciences,
Radboud University Nijmegen, The Netherlands.
{flaviog,petervr}@cs.ru.nl

Abstract. This paper provides one more step towards bridging the gap between the formal and computational approaches to the verification of cryptographic protocols. We extend the well-known Abadi-Rogaway logic with probabilistic hashes and give a precise semantic interpretation to it using Canetti's oracle hashes. These are probabilistic polynomial-time hashes that hide all partial information. Finally, we show that this interpretation is computationally sound and complete.

1 Introduction

In the last few years much research has been done to relate the symbolic and computational views on cryptography [AR02,AJ01,MW04,Her05]. Such a relation takes the form of a function mapping algebraic messages m to (distributions over) bit strings $[\![m]\!]$. This map then should relate messages that are observationally equivalent in the algebraic world (meaning that a Dolev-Yao attacker [DY83] can see no difference between them) to indistinguishable distributions over bit strings (meaning that a computationally bounded adversary can only with negligible probability distinguish the distributions). Such a map allows one to use algebraic methods, possibly even automated, to reason about security properties of protocols and have those reasonings be valid also in the computational world.

Micciancio and Warinschi [MW04] briefly but explicitly question if this logical approach can be extended to, among other things, collision resistant hashes. Backes, Pfitzmann, and Waidner [BPW06] show that in their simulatability framework [PW00] a sound interpretation of hashes cannot exist, but that it is possible to give a sound interpretation of formal hashes in the simulatability framework using random oracles.

The problem with hashes is that in the algebraic world $h(m)$ and $h(m')$ are indistinguishable for a Dolev-Yao attacker if the attacker does not know m and m'. In the computational world, however, the standard security definition — it must be computationally infeasible to compute any pre-image of a hash value or a hash collision [RS04] — does not guarantee that the hash function hides all partial information about the message; hence there is no guarantee that $[\![h(m)]\!]$ and $[\![h(m')]\!]$ are computationally indistinguishable.

We give an interpretation of formal hashes using the notion of perfectly one-way functions (a.k.a. oracle hashing) from Canetti and others [Can97a,CMR98]. These are computable probabilistic hash functions that hide all partial information of their input. Under suitable security conditions, we have shown soundness [GR06b] and completeness [GR06a] of this interpretation in the standard model.

2 The algebraic setting

Messages are constructed from constants c, keys k, nonces n and randomness labels r using algebraic encryption, hashing, and pairing operations:

$$\mathrm{Msg} \ni m := c \mid k \mid n \mid \{\!|m|\!\}_k^r \mid \mathrm{h}^r(m) \mid \langle m, m\rangle \mid \square^r \mid \boxtimes^r .$$

The *closure* $\overline{U}$ of a set of messages U is the set of all messages that can be constructed from U using tupling, detupling, and decryption. It represents the information an adversary could deduce knowing U.

As in [AR02] we define the function *pattern* that replaces unknown encryptions and hashes by boxes.

$$\mathit{pattern}(m) = \mathit{pattern}(m, \overline{\{m\}}),$$

where

$$\begin{aligned}
\mathit{pattern}(\langle m_1, m_2\rangle, U) &= \langle \mathit{pattern}(m_1, U), \mathit{pattern}(m_2, U)\rangle \\
\mathit{pattern}(\{\!|m|\!\}_k^r, U) &= \begin{cases} \{\!|\mathit{pattern}(m, U)|\!\}_k^r, & \text{if } k \in U; \\ \square^{\mathcal{R}(\{\!|m|\!\}_k^r)}, & \text{otherwise.} \end{cases} \\
\mathit{pattern}(\mathrm{h}^r(m), U) &= \begin{cases} \mathrm{h}^r(\mathit{pattern}(m, U)), & \text{if } m \in U; \\ \boxtimes^{\mathcal{R}(\mathrm{h}^r(m))}, & \text{otherwise.} \end{cases} \\
\mathit{pattern}(m, U) &= m \quad \text{in any other case.}
\end{aligned}$$

Here $\mathcal{R}$ is an injective function assigning randomness labels to messages.

Two messages m and m' are said to be *observationally equivalent*, notation $m \cong m'$, if there is a type preserving permutation σ of keys, nonces, randomness labels, and boxes such that $\mathit{pattern}(m) = \mathit{pattern}(m')\sigma$.

3 Oracle hashing

The underlying secrecy assumptions behind formal or Dolev-Yao hashes are very strong. It is assumed that given a hash value $h(x)$, it is not possible for an adversary to learn any information about the pre-image x. In the literature this idealization is often modelled with the random oracle [BR93]. Such a primitive is not computable and therefore it is also an idealization. Practical hash functions like SHA or MD5 are very useful cryptographic primitives even though these functions might leak partial information about their input. Moreover, under the traditional security notions (one-wayness), a function that reveals half of its input

is considered secure. In addition, any deterministic hash function h leaks partial information about x, namely $h(x)$. Through this paper we consider a new primitive introduced by Canetti [Can97a] called *oracle hashing*, that mimics what semantic security is for encryption schemes. This hash function is probabilistic and therefore it needs a verification function, just as in a signature scheme. A *hash scheme* consists of two algorithms $\mathcal{H}$ and $\mathcal{V}$. The probabilistic algorithm $\mathcal{H}\colon \text{Param} \times \text{Str} \to \text{Str}$ takes as arguments a unary sequence and a message and outputs a hash value; the verification algorithm $\mathcal{V}\colon \text{Str} \times \text{Str} \to \{0,1\}$, given two messages x and c, correctly decides whether c is a hash of x or not. As an example we reproduce here a hash scheme proposed in the original paper. Let p be a large (i.e., scaling with η) safe prime. Take $\mathcal{H}(x) = \langle r^2, r^{2 \cdot h(x)} \mod p \rangle$, where r is a randomly chosen element in $\mathbb{Z}_p^*$ and h is any collision resistant hash function. The verification algorithm $\mathcal{V}(x, \langle a, b \rangle)$ just checks whether $b = a^{h(x)} \mod p$.

Canetti gives essentially two security notions for such a hash scheme. The first one, *oracle indistinguishability*, guarantees that an adversary can gain no information at all about a bit string, given its hash value (or rather, with sufficiently small probability). The second one is an appropriate form of *collision resistance*. It guarantees that an adversary cannot (or rather, again, with sufficiently small probability) compute two distinct messages that successfully pass the verification test with the same hash value.

Definition 3.1 (Oracle indistinguishability). A hash scheme $\langle \mathcal{H}, \mathcal{V} \rangle$ is said to be *oracle indistinguishable* if for every family of probabilistic polynomial-time predicates $\{D_\eta\colon \text{Str} \to \{0,1\}\}_{\eta \in \mathbb{N}}$ and every positive polynomial p there is a polynomial-size family $\{L_\eta\}_{\eta \in \mathbb{N}}$ of subsets of Str such that for all large enough η and all $x, y \in \text{Str} \setminus L_\eta$:

$$\mathbb{P}[D_\eta(\mathcal{H}(1^\eta, x)) = 1] - \mathbb{P}[D_\eta(\mathcal{H}(1^\eta, y)) = 1] < \frac{1}{p(\eta)}.$$

Here the probabilities are taken over the choices made by $\mathcal{H}$ and the choices made by D_η. This definition is the non-uniform [Gol01] version of oracle indistinguishability proposed by Canetti [Can97a] as it is finally used throughout the proof (See the full version [Can97b], Appendix B).

Definition 3.2 (Collision resistance). A hash scheme $\langle \mathcal{H}, \mathcal{V} \rangle$ is said to be *collision resistant* if for every probabilistic polynomial-time adversary A, the probability

$$\mathbb{P}[\langle c, x, y \rangle \xleftarrow{\$} A(1^\eta); x \neq y \wedge \mathcal{V}(x, c) = \mathcal{V}(y, c) = 1]$$

is a negligible function of η.

4 Soundness & Completeness

Symbolic messages are interpreted in the standard way [AR02], using $\mathcal{H}$ to interpret hashes.

Theorem 4.1 (Soundness) *Assume that the encryption scheme is type-0 secure and that the hash scheme is oracle indistinguishable and collision resistant. Let m and m' be acyclic messages. Then $m \cong m' \implies [\![m]\!] \equiv [\![m']\!]$.* □

This result also holds for the standard ind-cca security notion by making *pattern* length- or structure-aware, as in, for instance, [Her05,MP05].

Theorem 4.2 (Completeness) *Assume also that the encryption scheme is confusion free. Let m_1 and m_2 be acyclic messages. Then $[\![m_1]\!] \equiv [\![m_2]\!] \implies m_1 \cong m_2$.* □

Acknowledgements. We are thankful to David Galindo for providing the reference to [Can97a] and insightful comments.

References

[AJ01] Martín Abadi and Jan Jürjens. Formal eavesdropping and its computational interpretation. In Naoki Kobayashi and Benjamin C. Pierce, editors, *Proceedings of the Fourth International Symposium on Theoretical Aspects of Computer Software (TACS'01)*, volume 2215 of *Lecture Notes in Computer Science*, pages 82–94. Springer, 2001.

[AR02] Martín Abadi and Phillip Rogaway. Reconciling two views of cryptography (the computational soundness of formal encryption). *Journal of Cryptology*, 15(2):103–127, 2002.

[BPW06] Michael Backes, Birgit Pfitzmann, and Michael Waidner. Limits of the reactive simulatability/UC of Dolev-Yao models with hashes. Cryptology ePrint Archive, Report 2006/014 (`http://eprint.iacr.org/2006/068`), 2006.

[BR93] Mihir Bellare and Phillip Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In *Proceedings of the 1st ACM CCS*, pages 62–73. ACM, 1993.

[Can97a] Ran Canetti. Towards realizing random oracles: Hash functions that hide all partial information. In Burt Kaliski, editor, *Advances in Cryptology, 17th Annual International Cryptology Conference (CRYPTO'97)*, volume 1294 of *Lecture Notes in Computer Science*, pages 455–469. Springer, 1997.

[Can97b] Ran Canetti. Towards realizing random oracles: Hash functions that hide all partial information. Cryptology ePrint Archive, Report 1997/007 (`http://eprint.iacr.org/1997/007`), 1997.

[CMR98] Ran Canetti, Danielle Micciancio, and Omer Reingold. Perfectly one-way probabilistic hash functions (preliminary version). In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing (STOC'98)*, pages 131–140. ACM, 1998.

[DY83] Danny Dolev and Andrew C. Yao. On the security of public key protocols. *IEEE Transactions on Information Theory*, 29(2):198–208, 1983.

[Gol01] Oded Goldreich. *Foundations of Cryptography*, volume 1. Cambridge University Press, 2001.

[GR06a] Flavio D. Garcia and Peter van Rossum. Completeness of formal hashes in the standard model. Cryptology ePrint Archive, Report 2006/146 (`http://eprint.iacr.org/2006/146`), 2006.

[GR06b] Flavio D. Garcia and Peter van Rossum. Sound computational interpretation of formal hashes. Cryptology ePrint Archive, Report 2006/014 (`http://eprint.iacr.org/2006/014`), 2006.

[Her05] Jonathan Herzog. A computational interpretation of Dolev-Yao adversaries. *Theoretical Computer Science*, 340(1):57–81, 2005.

[MP05] Daniele Micciancio and Saurabh Panjwani. Adaptive security of symbolic encryption. In Joe Kilian, editor, *Theory of Cryptography: Second Theory of Cryptography Conference (TCC'05)*, volume 3378 of *Lecture Notes in Computer Science*, pages 169–187. Springer, February 2005.

[MW04] Daniele Micciancio and Bogdan Warinschi. Completeness theorems of the Abadi-Rogaway logic of encrypted expressions. *Journal of Computer Security*, 12(1):99–129, 2004.

[PW00] Birgit Pfitzmann and Michael Waidner. Composition and integrity preservation of secure reactive systems. In *Proceedings of the 7th ACM CCS*, pages 245–254, 2000.

[RS04] Phillip Rogaway and Thomas Shrimpton. Cryptographic hash-function basics: Definitions, implications, and separations for preimage resistance, second-preimage resistance, and collision resistance. In Bimal Roy and Willi Meier, editors, *Fast Software Encryption: 11th International Workshop (FSE'04)*, volume 3017 of *Lecture Notes in Computer Science*, pages 371–388. Springer, 2004.