How to Break and Repair a Universally Composable Signature Functionality

Michael Backes[1] and Dennis Hofheinz[2]

[1] IBM Zurich Research Lab
mbc@zurich.ibm.com
[2] IAKS, Arbeitsgruppe Systemsicherheit, Prof. Dr. Th. Beth,
Fakultät für Informatik, Universität Karlsruhe
hofheinz@ira.uka.de

Abstract. Canetti and Rabin recently proposed a universally composable ideal functionality $\mathcal{F}_{\text{SIG}}$ for digital signatures. We show that this functionality cannot be securely realized by *any* signature scheme, thereby disproving their result that any signature scheme that is existentially unforgeable under adaptive chosen-message attack is a secure realization.
Next, an improved signature functionality is presented. We show that our improved functionality can be securely realized by precisely those signature schemes that are secure against existential forgery under adaptive chosen-message attacks.

Keywords: Digital signature schemes, universal composability, adaptive chosen-message attack.

1 Introduction

In this contribution, we investigate the idealization $\mathcal{F}_{\text{SIG}}$ of digital signatures, as defined in [6] for the framework of universal composability [2]. This framework enjoys a composition theorem which states that specifically, larger protocols may be formulated and investigated by means of an idealization of, e.g., digital signatures, while later a concrete digital signature scheme may be plugged in for the idealization while preserving the security of the large protocol. Certainly, this does not work for any signature scheme, but the scheme must—in a specified sense—*securely realize* the considered idealization, which makes the notion of secure realization (often also called *emulation*) the central notion of the framework.

We show that the idealization $\mathcal{F}_{\text{SIG}}$ cannot be securely realized by *any* real signature scheme. This in particular invalidates the results of [6, Claim 2] and [2, Claim 14 of full version].[1,2]

[1] Our proof applies to the $\mathcal{F}_{\text{SIG}}$-formulation from [6] as well as to the slightly older formulation in [2].

[2] After we had completed and published this manuscript as an IBM research report [1], the paper [6] was updated. In the updated version [7], the functionality $\mathcal{F}_{\text{SIG}}$ was re-

K. Zhang and Y. Zheng (Eds.): ISC 2004, LNCS 3225, pp. 61–72, 2004.

Next, we propose an improvement of $\mathcal{F}_{\text{SIG}}$ and we show that it can be securely realized by suitable real signature schemes, i.e., by precisely those ones that are secure against existential forgery under adaptive chosen-message attack as defined in [11].

The proof of unrealizability reveals a general problem with detached idealizations of digital signatures: In case of a corrupted signer, signatures for arbitrary messages may be generated *locally* by anyone who has knowledge of the signing key, hence it cannot be guaranteed that the ideal functionality, i.e., the idealization of digital signatures is notified upon every single signature generation. (Consider a larger protocol that honestly generates digital signatures using the publicly distributed signing key of a corrupted signer.) Thus, considering signatures as invalid which are not explicitly "registered" at the ideal functionality causes problems and indeed leads to our attack on $\mathcal{F}_{\text{SIG}}$ discussed below. On the other hand, all signatures not obtained via explicit signing queries to the ideal functionality should intuitively be rejected when they are verified. Our modification of $\mathcal{F}_{\text{SIG}}$ does not have this intuitive rejection property.

1.1 Overview of This Paper

We first briefly review the universal composability framework in Section 2 to prepare the ground for our subsequent results.

In Section 3, we review the ideal signature functionality proposed by Canetti and Rabin and show that it is not securely realizable at all.

In Section 4, we propose an improved functionality for digital signatures, and we show that it can be securely realized precisely by those signature schemes that are existentially unforgeable under adaptive-chosen message attack.

The paper ends with a conclusion (Section 5).

2 Preliminaries

To start, we shortly outline the framework for multi-party protocols as defined in [2]. First of all, *parties*, denoted by P_1 through P_n, are modeled as *interactive Turing machines (ITMs)* and are supposed to run some fixed protocol π. There also is an *adversary*, denoted $\mathcal{A}$ and modeled as an ITM as well, which carries out attacks on protocol π. $\mathcal{A}$ may corrupt parties in which case it learns their current and all past states as well as the contents of all their tapes; furthermore, it controls their future actions. $\mathcal{A}$ may further intercept or, when assuming unauthenticated message transfer (which is called the "bare" model in [2]), also fake messages sent between parties. If $\mathcal{A}$ corrupts parties only *before* the actual protocol run of π takes place, $\mathcal{A}$ is called *non-adaptive*, otherwise $\mathcal{A}$ is said to be *adaptive*. The respective local inputs for protocol π are supplied by an *environment machine*, which is also modeled as an ITM and denoted $\mathcal{Z}$, that

placed by a functionality $\mathcal{F}_{\text{CERT}}$. Furthermore, a modification of $\mathcal{F}_{\text{SIG}}$—independent of the one described here—was put forward in [3]. The generic attack discussed in this paper does not apply to the modified $\mathcal{F}_{\text{SIG}}$ functionality in [3].