On the Composition of Zero-Knowledge Proof Systems

Mikhail Anokhin

Information Security Institute, Lomonosov University, Moscow

Based on the work by O. Goldreich and H. Krawczyk with the same title (Proc. of the ICALP'90, p. 268–282, and SIAM J. on Computing, 1996, v. 25(1), p. 169–192.)

Preliminaries

Definition (negligible function)

A function $\epsilon\colon N \to [0, +\infty)$, where N is an infinite subset of $\mathbb{N}$, is called negligible if for any polynomial p there exists a nonnegative integer m such that $\epsilon(n) \leq 1/p(n)$ whenever $n \in N$ and $n \geq m$. Any negligible function is denoted by negl.

Definition (computational indistinguishability)

Let I be an infinite subset of $\mathbb{N}$ or $\{0,1\}^*$. Suppose $(X_i)_{i\in I}$ and $(Y_i)_{i\in I}$ are probability ensembles consisting of random variables that take values in $\{0,1\}^*$. Then these ensembles are said to be computationally indistinguishable if for every probabilistic polynomial-time algorithm D (called a distinguisher) and every $i \in I$,

$$|\Pr[D(1^i, X_i) = 1] - \Pr[D(1^i, Y_i) = 1]| = \mathrm{negl}(i) \text{ if } I \subseteq \mathbb{N},$$
$$|\Pr[D(i, X_i) = 1] - \Pr[D(i, Y_i) = 1]| \leq \mathrm{negl}(|i|) \text{ if } I \subseteq \{0,1\}^*.$$

Interactive Proof Systems

Let $L \subseteq \{0,1\}^*$. An interactive proof system (protocol) for the language L is a protocol between two parties, called the prover (P) and the verifier (V). The prover is probabilistic computationally unlimited, whereas the verifier is probabilistic polynomial-time. Both parties share a common input $x \in \{0,1\}^*$. The aim of the prover is to convince the verifier that $x \in L$. At the end of the interaction, the verifier either accepts the proof for x ($\langle P, V \rangle(x) = \text{accept}$) or rejects it ($\langle P, V \rangle(x) = \text{reject}$). The proof system must satisfy the following two conditions:

- Completeness: For any $x \in L$, $\Pr[\langle P, V \rangle(x) = \text{accept}] \geq 1 - \text{negl}(|x|)$.
- Soundness: For any $x \in \{0,1\}^* \setminus L$ and any (not necessarily honest) prover P', $\Pr[\langle P', V \rangle(x) = \text{accept}] \leq \text{negl}(|x|)$.

The strongest forms of these conditions are as follows:

- Perfect completeness: For every $x \in L$, always $\langle P, V \rangle(x) = \text{accept}$.
- Perfect soundness: For every $x \in \{0,1\}^* \setminus L$ and every (not necessarily honest) prover P', always $\langle P', V \rangle(x) = \text{reject}$.

Sequential Composition of Interactive Proof Systems

Let $\Pi_i = (P_i, V_i)$ be an interactive proof system for a language L_i, $i = 1, \ldots, k$.

Definition (sequential composition)

On common input $(x_1, \ldots, x_k)$, where $x_i \in \{0,1\}^*$, the execution of the sequential composition $\Pi_1 \circ_s \ldots \circ_s \Pi_k$ of $\Pi_1, \ldots, \Pi_k$ consists of k stages. At stage i, Π_i on common input x_i is executed as a subroutine. The verifier of $\Pi_1 \circ_s \ldots \circ_s \Pi_k$ accepts if and only if V_i has accepted for all $i \in \{1, \ldots, k\}$.

$\Pi_1 \circ_s \ldots \circ_s \Pi_k$ is an interactive proof system for $L_1 \times \cdots \times L_k$. If Π_i is perfectly complete (resp., perfectly sound) for all $i \in \{1, \ldots, k\}$, then $\Pi_1 \circ_s \ldots \circ_s \Pi_k$ is also perfectly complete (resp., perfectly sound).

The honest prover in $\Pi_1 \circ_s \ldots \circ_s \Pi_k$ is denoted by $P_1 \circ_s \ldots \circ_s P_k$.

Parallel Composition of Interactive Proof Systems

Assume that either all provers or all verifiers initiate $\Pi_1, \ldots, \Pi_k$ and that these proof systems have the same number of rounds (denoted by J). If this is not the case, the proof systems should be padded with dummy steps.

Definition (parallel composition)

The parallel composition $\Pi_1 \circ_p \ldots \circ_p \Pi_k$ of $\Pi_1, \ldots, \Pi_k$ is the J-round protocol such that the jth message in this protocol executed on common input $(x_1, \ldots, x_k)$ is $(M_1, \ldots, M_k)$, where M_i is the jth message of Π_i executed on common input x_i $(j = 1, \ldots, J)$. The verifier of $\Pi_1 \circ_p \ldots \circ_p \Pi_k$ accepts if and only if V_i has accepted for all $i \in \{1, \ldots, k\}$.

$\Pi_1 \circ_p \ldots \circ_p \Pi_k$ is an interactive proof system for $L_1 \times \cdots \times L_k$. If Π_i is perfectly complete (resp., perfectly sound) for all $i \in \{1, \ldots, k\}$, then $\Pi_1 \circ_p \ldots \circ_p \Pi_k$ is also perfectly complete (resp., perfectly sound).

The honest prover in $\Pi_1 \circ_p \ldots \circ_p \Pi_k$ is denoted by $P_1 \circ_p \ldots \circ_p P_k$.

Zero-Knowledge Property

Let (P, V) be an interactive proof system for an infinite language L. For any $x \in \{0,1\}^*$, the random variable $\text{view}_V^P(x)$ is defined as $(r; w_1, \ldots, w_k)$, where $w_1, \ldots, w_k$ is the sequence of messages received by V from P during an execution of the proof system on common input x provided that V uses $r \in \{0,1\}^*$ as a source of random bits.

Definition (zero-knowledge proof system)

The proof system (P, V) for the language L is called (computational) zero-knowledge if for every probabilistic polynomial-time interactive algorithm V' (dishonest verifier) there exists a probabilistic polynomial-time algorithm S (called a simulator) such that the ensembles $(\text{view}_{V'}^P(x))_{x \in L}$ and $(S(x))_{x \in L}$ are computationally indistinguishable.

P-Evasive and Pseudorandom Families of Sets

Let $E = (E_n)_{n\in\mathbb{N}}$ be a family of nonempty sets such that $E_n \subseteq \{0,1\}^{m(n)}$ for any $n \in \mathbb{N}$, where m is a polynomial.

Definition (P-evasive family of sets)

The family E is called P-evasive if for every probabilistic polynomial-time algorithm A and every $x \in \{0,1\}^*$, $\Pr[A(x) \in E_{|x|}] \leq \mathrm{negl}(|x|)$.

For a finite nonempty set M, denote by U_M a random variable uniformly distributed on M.

Definition (pseudorandom family of sets)

The family E is called pseudorandom if the ensembles $(U_{E_n})_{n\in\mathbb{N}}$ and $(U_{\{0,1\}^{m(n)}})_{n\in\mathbb{N}}$ are computationally indistinguishable.

Existence of P-Evasive Pseudorandom Families of Sets

Theorem (Goldreich and Krawczyk)

There exists a family of sets $E = (E_n)_{n\in\mathbb{N}}$ satisfying the following conditions:

- *$E_n \subseteq \{0,1\}^{4n}$ and $|E_n| = 2^n$ for any $n \in \mathbb{N}$.*
- *E is P-evasive and pseudorandom.*
- *There exists a deterministic algorithm G such that $G(1^n) = E_n$ for any $n \in \mathbb{N}$.*

Choose a family of sets $E = (E_n)_{n\in\mathbb{N}}$ satisfying the conditions of this theorem. Also, let $K\colon \{0,1\}^* \to \{0,1\}$ be a computable predicate such that $L_K = \{x \in \{0,1\}^* \mid K(x) = 1\} \notin \mathrm{BPP}$. Since $\mathrm{BPP} \subseteq \mathrm{PSPACE} \subset \mathrm{ESPACE} \subseteq \mathrm{R}$ (the second inclusion follows from the space hierarchy theorem), such a predicate exists.

The Protocol Π

Suppose $x \in \{0,1\}^*$ is a common input for P and V and n is the length of x.

	Π	
	P	V
1		$\longleftarrow v \in_{\mathcal{U}} \{0,1\}^{4n}$
2	$w = \begin{cases} K(x) & \text{if } v \in E_n \\ e \in_{\mathcal{U}} E_n & \text{otherwise} \end{cases} \longrightarrow$	
		accepts

Theorem

The protocol Π is a perfectly complete, perfectly sound zero-knowledge interactive proof system for the language $\{0,1\}^*$, but $\Pi \circ_s \Pi$ is not zero-knowledge.

Sketch of Proof of the Theorem (1)

The perfect completeness and perfect soundness of Π are trivial.

For random variables $z = z_x \in \{0,1\}^*$ and $z' = z'_x \in \{0,1\}^*$, where x ranges over $\{0,1\}^*$, $z \approx z'$ denotes that $(z_x)_{x\in\{0,1\}^*}$ and $(z'_x)_{x\in\{0,1\}^*}$ are computationally indistinguishable.

Let V' be a probabilistic polynomial-time interactive algorithm (dishonest verifier). Then $\mathrm{view}^P_{V'}(x) = (r; w)$, where r is the random bit string used by V' and w is as in Π, except that $v = V'(x; r)$ rather than $v \in_{\mathcal{U}} \{0,1\}^{4n}$. Since $\Pr[v \in E_n] \leq \mathrm{negl}(n)$, we have $(r; w) \approx (r; e)$, where $e \in_{\mathcal{U}} E_n$. Moreover, if $u \in_{\mathcal{U}} \{0,1\}^{4n}$, then $(r; e) \approx (r; u)$ because E is pseudorandom. Thus, $S(x) = (r; u)$, where r and u are as above, defines a simulator for V'.

Sketch of Proof of the Theorem (2)

Let V'' be the dishonest verifier for $\Pi \circ_s \Pi$ that interacts with $P \circ_s P$ on common input (x, x) ($x \in \{0,1\}^n$, $n \in \mathbb{N}$) as follows:

exec. of Π		$P \circ_s P$	V''
1st	1		$\longleftarrow v_1 \in_{\mathcal{U}} \{0,1\}^{4n}$
	2	$w = \begin{cases} K(x) & \text{if } v_1 \in E_n \\ e \in_{\mathcal{U}} E_n & \text{otherwise} \end{cases} \longrightarrow$	
2nd	1		$\longleftarrow v_2 = \begin{cases} v_1 & \text{if } w \in \{0,1\} \\ w & \text{otherwise} \end{cases}$
	2	$v_2 \in E_n \implies K(x) \longrightarrow$	
			accepts

Then $\text{view}_{V''}^{P \circ_s P}(x, x) = (v_1; w, K(x))$. Therefore any simulator for V'' can be used for computing $K(x)$, given $x \in \{0,1\}^n$, with error probability at most $\text{negl}(n)$ ($n \in \mathbb{N}$). Since $L_K \notin \text{BPP}$, no such simulator exists. □

Auxiliary-Input Model

Let V have a private auxiliary input $y \in \{0,1\}^*$ representing a priori information to the verifier. Note that V is supposed to be polynomial-time in $|x|$ rather than in $|x| + |y|$, where x is the common input. Therefore we can assume that $y \in \{0,1\}^{p(|x|)}$, where p is an arbitrary (but fixed) polynomial. The (perfect) completeness and (perfect) soundness conditions in this model are obtained by replacing V by $V(y)$, where y ranges over $\{0,1\}^*$:

- Completeness: For any $x \in L$ and any $y \in \{0,1\}^*$, $\Pr[\langle P, V(y)\rangle(x) = \text{accept}] \geq 1 - \text{negl}(|x|)$.
- Soundness: For any $x \in \{0,1\}^* \setminus L$, any $y \in \{0,1\}^*$, and any (not necessarily honest) prover P', $\Pr[\langle P', V(y)\rangle(x) = \text{accept}] \leq \text{negl}(|x|)$.
- Perfect completeness: For every $x \in L$ and every $y \in \{0,1\}^*$, always $\langle P, V(y)\rangle(x) = \text{accept}$.
- Perfect soundness: For every $x \in \{0,1\}^* \setminus L$, every $y \in \{0,1\}^*$, and every (not necessarily honest) prover P', always $\langle P', V(y)\rangle(x) = \text{reject}$.

Auxiliary-Input Zero-Knowledge Property

Definition (auxiliary-input zero-knowledge proof system)

The proof system (P, V) for an infinite language L in the auxiliary-input model is called auxiliary-input (computational) zero-knowledge if for every polynomial p and every probabilistic polynomial-time interactive algorithm V' (dishonest verifier) there exists a probabilistic polynomial-time algorithm S (called a simulator) such that the ensembles $(\text{view}^P_{V'(y)}(x))_{x\in L, y\in\{0,1\}^{p(|x|)}}$ and $(S(x,y))_{x\in L, y\in\{0,1\}^{p(|x|)}}$ are computationally indistinguishable.

Theorem (Goldreich and Oren)

Let $\Pi_1, \ldots, \Pi_k$ be auxiliary-input zero-knowledge proof systems for infinite languages $L_1, \ldots, L_k$, respectively. Then $\Pi_1 \circ_s \ldots \circ_s \Pi_k$ is an auxiliary-input zero-knowledge proof system for $L_1 \times \cdots \times L_k$.

P/poly-Evasive and Non-Uniformly Strong Pseudorandom Families of Sets

Let $E = (E_{n,s})_{n\in\mathbb{N},s\in\{0,1\}^n}$ be a family of nonempty sets such that $E_{n,s} \subseteq \{0,1\}^{m(n)}$ for any $n \in \mathbb{N}$ and any $s \in \{0,1\}^n$, where m is a polynomial.

Definition (P/poly-evasive family of sets)

The family E is called P/poly-evasive if for every family $(C_n\colon \{0,1\}^n \to \{0,1\}^{m(n)})_{n\in\mathbb{N}}$ of polynomial-size probabilistic circuits, $\Pr[C_n(s) \in E_{n,s}] = \mathrm{negl}(n)$, where $s \in_{\mathcal{U}} \{0,1\}^n$.

Definition (non-uniformly strong pseudorandom family of sets)

The family E is called non-uniformly strong pseudorandom if for any family $(D_n\colon \{0,1\}^{m(n)} \to \{0,1\})_{n\in\mathbb{N}}$ of polynomial-size probabilistic circuits, $\max_{s\in\{0,1\}^n}|\Pr[D_n(v) = 1] - \Pr[D_n(w) = 1]| = \mathrm{negl}(n)$, where $v \in_{\mathcal{U}} E_{n,s}$ and $w \in_{\mathcal{U}} \{0,1\}^{m(n)}$.

Existence of P/poly-Evasive Non-Uniformly Strong Pseudorandom Families of Sets

Theorem (Goldreich and Krawczyk)

There exists a family of sets $E = (E_{n,s})_{n \in \mathbb{N}, s \in \{0,1\}^n}$ *satisfying the following conditions:*

- $E_{n,s} \subseteq \{0,1\}^{4n}$ *and* $|E_{n,s}| = 2^n$ *for any* $n \in \mathbb{N}$ *and any* $s \in \{0,1\}^*$.
- E *is P/poly-evasive and non-uniformly strong pseudorandom.*
- *There exists a deterministic algorithm* G *such that* $G(1^n, s) = E_{n,s}$ *for any* $n \in \mathbb{N}$ *and any* $s \in \{0,1\}^*$.

Choose a family of sets $E = (E_{n,s})_{n \in \mathbb{N}, s \in \{0,1\}^n}$ satisfying the conditions of this theorem. As above, $K \colon \{0,1\}^* \to \{0,1\}$ denotes a computable predicate such that $L_K = \{x \in \{0,1\}^* \mid K(x) = 1\} \notin \mathrm{BPP}$.

The Protocols Π_1 and Π_2

Suppose $x \in \{0,1\}^*$ is a common input for all the parties and n is the length of x. Moreover, let $\tau_n(t) = t$ if $t \in \{0,1\}^n$ and $\tau_n(t) = 0^n$ otherwise.

Π_1			Π_2	
P_1	V_1		P_2	V_2
$s \in_{\mathcal{U}} \{0,1\}^n \longrightarrow$		1	dummy step	
	dummy step	2		$\longleftarrow t \in_{\mathcal{U}} \{0,1\}^n$
dummy step		3	$w \in_{\mathcal{U}} E_{n,\tau_n(t)} \longrightarrow$	
	$\longleftarrow v \in_{\mathcal{U}} \{0,1\}^{4n}$	4		dummy step
$\begin{cases} K(x) & \text{if } v \in E_{n,s} \\ \perp & \text{otherwise} \end{cases} \longrightarrow$		5	dummy step	
accepts			accepts	

Theorem

The protocols Π_1 and Π_2 are perfectly complete, perfectly sound zero-knowledge (even against non-uniform verifiers and distinguishers and hence in the auxiliary-input model) interactive proof systems for the language $\{0,1\}^*$, but $\Pi_1 \circ_p \Pi_2$ is not zero-knowledge.

Sketch of Proof of the Theorem (1)

The perfect completeness and perfect soundness of Π_1 and Π_2 are trivial.

For any non-uniform (possibly dishonest) verifier V_1' for Π_1, a simulator S_1 is defined by $S_1(x) = (r_1; s, \perp)$, where r_1 is the random bit string used by V_1' and $s \in_{\mathcal{U}} \{0,1\}^n$. This is because E is P/poly-evasive.

For any non-uniform (possibly dishonest) verifier V_2' for Π_2, a simulator S_2 is defined by $S_2(x) = (r_2; u)$, where r_2 is the random bit string used by V_2' and $u \in_{\mathcal{U}} \{0,1\}^{4n}$. This is because E is non-uniformly strong pseudorandom.

Sketch of Proof of the Theorem (2)

Let V' be the dishonest verifier for $\Pi_1 \circ_p \Pi_2$ that interacts with $P_1 \circ_p P_2$ on common input (x, x) ($x \in \{0,1\}^n$, $n \in \mathbb{N}$) as follows:

	$P_1 \circ_p P_2$	V'
1	$s \in_{\mathcal{U}} \{0,1\}^n \longrightarrow$	
2		$\longleftarrow s$
3	$w \in_{\mathcal{U}} E_{n,s} \longrightarrow$	
4		$\longleftarrow w$
5	$w \in E_{n,s} \implies K(x) \longrightarrow$	
		accepts

Then $\text{view}_{V'}^{P_1 \circ_p P_2}(x, x) = (s, w, K(x))$. Therefore any simulator for V' can be used for computing $K(x)$, given $x \in \{0,1\}^n$, with error probability at most $\text{negl}(n)$ ($n \in \mathbb{N}$). Since $L_K \notin \text{BPP}$, no such simulator exists. □