Sequential and parallel composition of Zero-Knowledge Proofs

Andreas Bomke

April 26, 2001

Outline

- Motivation
- Definition of a Zero-Knowledge Interactive Proof
- Evasive and Pseudorandom Sets
- Sequential Composition
- Parallel Composition
- Outlook

Motivation:
How can a party A *prove* to a party B that x is in a language L so that no more *knowledge* than $x \in L$ is revealed?

Applications:

- prove some property of one's secrets without revealing them
- identification schemes
- as a subprotocol in bigger tasks like
  - voting
  - electronic cash
  - distributed key generation

- Goldreich, Oren, 1992
  Definitions and Properties of Zero-Knowledge Proof Systems

- Goldreich, Krawczyk, 1996
  On the composition of zero-knowledge proof systems

- Kilian, Petrank, Rackoff, 1998
  Lower bounds for zero knowledge on the internet

Definitions

Definition 1. An **interactive proof** $\langle P, V \rangle$ for a language L is a two-party protocol in which a computationally unrestricted **prover**, P, interacts with a probabilistic polynomial-time **verifier**, V, by exchanging messages.
Both parties share a common input x.
At the end V either *accepts* or *rejects* and it holds:

- Completeness
  For any $c > 0$ and sufficiently long $x \in L$,
  $Prob(V \text{ accepts } x) > 1 - |x|^{-c}$.

- Soundness
  For any $c > 0$ and sufficiently long $x \notin L$,
  $Prob(V \text{ accepts } x) < |x|^{-c}$, even if the prover deviates from the prescribed protocol.

number of rounds = number of (alternately) exchanged messages

$\langle \mathbf{P}, \mathbf{V}^* \rangle (\mathbf{x})$ denotes the probability distribution generated by V^* after interacting with P on inputs $x \in L$

$f : \mathbb{N} \to \mathbb{R}$ is called **negligible** if for all $c > 0$ and sufficiently large n, $f(n) < n^{-c}$
f is called **nonnegligible** if there exists a $c > 0$ such that for all sufficiently large n, $f(n) > n^{-c}$

Definition 2. Goldwasser, Micali, Rackoff, 1985
An interactive proof $\langle P, V \rangle$ is called **zero-knowledge** if for *every* probabilistic polynomial-time V^*, there exists a probabilistic expected polynomial-time **simulator** M_{V^*} that on inputs $x \in L$ produces probability distributions $M_{V^*}(x)$ polynomially indistinguishable* from the distributions $\langle P, V^* \rangle (x)$.

*informally this means that there exists no probabilistic polynomial time algorithm which can decide with better than negligible error probability, when given a polynomial number of samples, from which of the distributions they are drawn

On evasive and pseudorandom sets

Definition 3. $S \subseteq \{0,1\}^k$ is called $(\tau(k), \varepsilon(k))$-**pseudorandom** if for any probabilistic circuit C of size $\tau(k)$

$$|p_C(S) - p_C(\{0,1\}^k)| \leq \varepsilon(k),$$

where $p_C(X)$ denotes the probability that C outputs 1 when given elements of X chosen with uniform probability.

Definition 4. A collection of uniform distributions on a sequence of $(\tau(k), \varepsilon(k))$-pseudorandom sets is called a **pseudorandom ensemble** if $\tau(\cdot)$ and $\varepsilon^{-1}(\cdot)$ grow faster than any polynomial.

Definition 5. Let $Q(\cdot)$ be a polynomial and $S_1, S_2, \ldots$ a sequence of nonempty sets with $S_n \subseteq \{0,1\}^{Q(n)}$. Such a sequence is called a **P-evasive ensemble** if for any probabilistic polynomial-time algorithm A and any $x \in \{0,1\}^n$, $Prob_A(A(x) \in S_n)$ is negligible in n.

Theorem 1. *There exists a P-evasive pseudorandom ensemble $S_1, S_2, \ldots$ with $Q(n) = 4n$.*

Definition 6. Let $Q(\cdot)$ be a polynomial, and for every $n \in \mathbb{N}$ let $S^{(n)} = \{S_1^{(n)}, \ldots, S_{2^n}^{(n)}\}$ be a collection of 2^n sets where each $S_i^{(n)} \subseteq \{0,1\}^{Q(n)}$. The sequence $S^{(1)}, S^{(2)}, \ldots$ is called a **P/Poly-evasive ensemble**, if for any $c > 0$, sufficiently large n, and any probabilistic circuit C of size n^c

$$Prob_{C,i}(C(i) \in S_i) < n^{-c}$$

Theorem 2. *There exists a P/Poly-evasive pseudo-random ensemble $S^{(1)}, S^{(2)}, \ldots$ with $Q(n) = 4n$.*

Sequential composition

Theorem 3. *Zero-knowledge* is not closed *under sequential composition.*

A sequential composition of protocols $\langle P_1, V_1\rangle, \dots, \langle P_k, V_k\rangle$ with inputs $x_1, \dots, x_k$ is defined as a protocol $\langle P, V\rangle$ which on input $x_1\%x_2\%\dots\%x_k\%$, where % is a delimiter, proceeds in k stages where in stage i it activates P_i and V_i as subroutines on x_i. V accepts iff all V_i's accept.

Proof. Let $S_1, S_2, \dots$ be a P-evasive pseudorandom ensemble and K a Boolean function with $L_K := \{x : K(x) = 1\} \notin BPP$.

Consider the following protocol $\langle P, V \rangle$ with input x of length $n = |x|$.

P	V
	send $s \in_R \{0,1\}^{4n}$
if $s \in S_n$ then send $K(x)$ else send $s_0 \in_R S_n$	
	accept x

Obviously, $\langle P, V \rangle$ is a zero-knowledge proof for $\{0,1\}^*$, because a simulator only needs to simulate the sending of $s_0 \in_R S_n$, * by randomly choosing s_0 from $\{0,1\}^{4n}$ rather than from S_n. †

*because of the evasivity condition of S_n, there is only a negligible probability that any verifier finds a $s \in S_n$

†this is indistinguishable because S_n is a pseudorandom subset of $\{0,1\}^{4n}$

Now consider a sequential composition of two such protocols with input $x\%x'\%$, where $|x'| = |x| = n$.

Since S_n is fixed for all executions of the single protocol, a cheating verifier V^* can use $s_0 \in_R S_n$, obtained with overwhelming probability in the first execution, to send it instead of the random string s in the second execution.

Thus V^* gets the value of $K(x')$ from the prover.

Since $L_K \notin BPP$ there is no way to simulate in probabilistic polynomial time the interaction of P with V^*. $\square$

A stronger formulation of ZK

Definition 7. An interactive proof $\langle P, V \rangle$ is called **black-box simulation zero-knowledge** if there exists a probabilistic expected polynomial-time oracle machine M such that for any polynomial-time verifier V^* and for $x \in L$, the distributions $\langle P, V^* \rangle(x)$ and $M^{V^*}(x)$ are polynomially indistinguishable.

every black-box simulation zero-knowledge proof is also a zero-knowledge proof, but the converse is not true

Theorem 4. *Black-box simulation zero-knowledge is robust under sequential composition.*

proof idea:
Compose the simulators M_i for $\langle P_i, V_i \rangle$ to a simulator M for the sequentially composed protocol $\langle P, V \rangle$.
Assuming that a distinguisher between the distributions $M^{V'}(x)$ and $\langle P, V' \rangle(x)$ exists, show that for some index j a distinguisher between the distributions $M_j^{V^*}(x)$ and $\langle P_j, V^* \rangle(x)$ exists, thus yielding a contradiction.

Parallel composition

Theorem 5. *Black-box simulation zero-knowledge is not closed under parallel composition.*

A parallel composition of protocols $\langle P_1, V_1 \rangle, \ldots, \langle P_k, V_k \rangle$ on inputs $x_1, \ldots, x_k$, which w.l.o.g. all have exactly m rounds, is defined as a protocol $\langle P, V \rangle$ on input $x_1 \% x_2 \% \ldots \% x_k \%$ whose ith message consists of the ith messages of $\langle P_1, V_1 \rangle, \ldots, \langle P_k, V_k \rangle$. V accepts iff all V_i's accept.

Proof. Let $S^{(1)}, S^{(2)}, \ldots$ be a P/Poly-evasive ensemble and K a Boolean function with $L_K := \{x : K(x) = 1\} \notin BPP$.

Consider the following protocols $\langle P_1, V_1\rangle$ and $\langle P_2, V_2\rangle$ with inputs x of length $n = |x|$.

P_1	V_1	step	P_2	V_2
send $i \in_R \{1, \ldots, 2^n\}$		1	dummy step	
	dummy step	2		send $j \in_R \{1, \ldots, 2^n\}$
dummy step		3	send $r \in_R S_j^{(n)}$	
	send $s \in_R \{0,1\}^{4n}$	4		dummy step
if $s \in S_i^{(n)}$ then send $K(x)$		5	dummy step	

The two protocols are black-box simulation ZK when executed independently due to the following simulators M_1 and M_2.

M_1 sends $i \in_R \{1, \ldots, 2^n\}$ in step 1, like P_1 does. In step 5 M_1 behaves as if the verifier sends $s \notin S_i^{(n)}$, since this is false only with negligible probability because of the evasiveness of $S^{(1)}, S^{(2)}, \ldots$

M_2 sends $r \in_R \{0,1\}^{4n}$ in step 3. This choice is polynomially indistinguishable from the prover's one because of the pseudorandomness of $S_j^{(n)}$.

The parallel composition of $\langle P_1, V_1 \rangle$ and $\langle P_2, V_2 \rangle$ into a single protocol $\langle P, V \rangle$ is not black-box ZK.

Consider the cheating verifier V^*:

step	V^*	comment
1	receive $i \in_R \{1, \ldots, 2^n\}$	
2	send i	$j = i$
3	receive $r \in_R S_i^{(n)}$	
4	send r	$s = r$
5	receive $K(x)$	

Since $L_K \notin BPP$ this interaction cannot be simulated in probabilistic polynomial-time. $\square$

Some results

- Goldreich, Micali, Wigderson, 1986
  if one-way functions exist, then every language in NP has a black-box simulation ZK proof

- Goldreich, Krawczyk, 1996
  if a 3-round protocol for a language L exists, whose parallel composition is black-box simulation ZK, then $L \in BPP$

- Kilian, Petrank, Rackoff, 1998
  if a 4-round protocol for L exists, whose asynchronous parallel composition is black-box simulation ZK, then $L \in BPP$

- Dwork, Naor, Sahai, 1998
  certain variants of the ZK-definition are robust under a parallel composition where constraints on the timing of messages are given

Outlook and open questions

- Is the parallel composition of selected proofs (used in practice) possible?

- How can the notion of ZK be strengthened so that the resulting class is closed under parallel composition?

- Which of the existing definitions are suitable for which applications?

- Practical advantages of using ZK in cryptographic protocols?

- Are the known round-complexities optimal?