Notes for Lecture 28

Posted May 6, 2009

Summary

Today we define the notion of computational zero knowledge and show that the simulator we described in the last lecture establishes the computational zero knowledge property of the 3-coloring protocol.

1 The Protocol and the Simulator

Recall that we use a commitment scheme (C, O) for messages in $\{1, 2, 3\}$, and that the common input to the prover and the verifier is a graph $G = ([n], E)$, where $[n] := \{1, 2, \ldots, n\}$. The prover, in addition, is given a valid 3-coloring $\alpha : [n] \to \{1, 2, 3\}$ of G.

The protocol is defined as follows:

- The prover picks a random permutation $\pi : \{1, 2, 3\} \to \{1, 2, 3\}$ of the set of colors, and defines the 3-coloring $\beta(v) := \pi(\alpha(v))$. The prover picks n keys $K_1, \ldots, K_n$ for (C, O), constructs the commitments $c_v := C(K_v, \beta(v))$ and sends $(c_1, \ldots, c_n)$ to the verifier;
- The verifier picks an edge $(u, v) \in E$ uniformly at random, and sends (u, v) to the prover;
- The prover sends back the keys K_u, K_v;
- If $O(K_u, c_u)$ and $O(K_v, c_v)$ are the same color, or if at least one of them is equal to $FAIL$, then the verifier rejects, otherwise it accepts

For every verifier algorithm V^*, we defined a simulator algorithm S^* which repeats the following procedure until the output is different from $FAIL$:

Algorithm S^*_{1round}

- Input: graph $G = ([n], E)$

- Pick random coloring $\gamma : [n] \to \{1, 2, 3\}$.
- Pick n random keys $K_1, \ldots, K_n$
- Define the commitments $c_i := C(K_i, \gamma(i))$
- Let (u, v) be the 2nd-round output of V^* given G as input and $c_1, \ldots, c_n$ as first-round message
- If $\gamma(u) = \gamma(v)$, then output FAIL
- Else output $((c_1, \ldots, c_n), (u, v), (K_u, K_v))$

We want to show that this simulator construction establishes the *computational zero knowledge* property of the protocol, assuming that (C, O) is secure. We give the definition of computational zero knowledge below.

Definition 1 (Computational Zero Knowledge) *We say that a protocol (P, V) for 3-coloring is (t, ϵ) computational zero knowledge with simulator overhead $so(\cdot)$ if for every verifier algorithm V^* of complexity $\leq t$ there is a simulator S^* of complexity $\leq so(t)$ on average such that for every algorithm D of complexity $\leq t$, every graph G and every valid 3-coloring α we have*

$$|\,\mathbb{P}[D(P(G, \alpha) \leftrightarrow V^*(G)) = 1] - \mathbb{P}[D(S^*(G)) = 1]| \leq \epsilon$$

Theorem 2 *Suppose that (C, O) is $(2t + O(nr), \epsilon/(4 \cdot |E| \cdot n))$-secure and that C is computable in time $\leq r$.*

Then the protocol defined above is (t, ϵ) computational zero knowledge with simulator overhead at most $1.6 \cdot t + O(nr)$.

2 Proving that the Simulation is Indistinguishable

In this section we prove Theorem 2.

Suppose that the Theorem is false. Then there is a graph G, a 3-coloring α, a verifier algorithm V^* of complexity $\leq t$, and a distinguishing algorithm D also of complexity $\leq t$ such that

$$|\,\mathbb{P}[D(P(G, \alpha) \leftrightarrow V^*(G)) = 1 - \mathbb{P}[D(S^*(G)) = 1]| \geq \epsilon$$

Let $2R_{u,v}$ be the event that the edge (u, v) is selected in the second round; then

$$
\begin{aligned}
\epsilon &\leq \left| \mathbb{P}[D(P(G,\alpha) \leftrightarrow V^*(G)) = 1] - \mathbb{P}[D(S^*(G)) = 1] \right| \\
&= \left| \sum_{(u,v)\in E} \mathbb{P}[D(P(G,\alpha) \leftrightarrow V^*(G)) = 1 \wedge 2R_{u,v}] \right. \\
&\quad \left. - \sum_{(u,v)\in E} \mathbb{P}[D(S^*(G)) = 1 \wedge 2R_{u,v}] \right| \\
&\leq \sum_{(u,v)\in E} | \mathbb{P}[D(P(G,\alpha) \leftrightarrow V^*(G)) = 1 \wedge 2R_{u,v}] \\
&\quad - \mathbb{P}[D(S^*(G)) = 1 \wedge 2R_{u,v}] |
\end{aligned}
$$

So there must exist an edge $(u^*, v^*) \in E$ such that

$$
| \mathbb{P}[D(P \leftrightarrow V^*) = 1 \wedge 2R_{u^*,v^*}] - \mathbb{P}[D(S^*) = 1 \wedge 2R_{u^*,v^*}] | \geq \frac{\epsilon}{|E|} \tag{1}
$$

(We have omitted references to G, α, which are fixed for the rest of this section.)

Now we show that there is an algorithm A of complexity $2t + O(nr)$ that is able to distinguish between the following two distributions over commitments to $3n$ colors:

- **Distribution (1)** commitments to the $3n$ colors $1, 2, 3, 1, 2, 3, \ldots, 1, 2, 3$;
- **Distribution (2)** commitments to $3n$ random colors

Algorithm A:

- Input: 3n commitments $d_{a,i}$ where $a \in \{1, 2, 3\}$ and $i \in \{1, \ldots, n\}$;
- Pick a random permutation $\pi : \{1, 2, 3\} \to \{1, 2, 3\}$
- Pick random keys K_{u^*}, K_{v^*}
- Construct the sequence of commitments $c_1, \ldots, c_n$ by setting:
  - $c_{u^*} := C(K_{u^*}, \pi(\alpha(u^*))$
  - $c_{v^*} := C(K_{v^*}, \pi(\alpha(v^*))$
  - for every $w \in [n] - \{u^*, v^*\}$, $c_w := d_{\pi(\alpha(w)),w}$
- If the 2nd round output of V^* given G and $c_1, \ldots, c_n$ is different from (u^*, v^*) output 0

- Else output $D((c_1, \ldots, c_n), (u^*, v^*), (K_{u^*}, K_{v^*}))$

First, we claim that

$$\mathbb{P}[A(\text{Distribution 1}) = 1] = \mathbb{P}[D(P \leftrightarrow V^*) = 1 \wedge 2R_{u^*,v^*}] \tag{2}$$

This follows by observing that A on input Distribution (1) behaves exactly like the prover given the coloring α, and that A accepts if and only if the event $2R_{u^*,v^*}$ happens and D accepts the resulting transcript.

Next, we claim that

$$|\,\mathbb{P}[A(\text{Distribution 2}) = 1] - \mathbb{P}[D(S^*) = 1 \wedge 2R_{u^*,v^*}]| \leq \frac{\epsilon}{2|E|} \tag{3}$$

To prove this second claim, we introduce, for a coloring γ, the quantity $DA(\gamma)$, defined as the probability that the following probabilistic process outputs 1:

- Pick random keys $K_1, \ldots, K_n$
- Define commitments $c_u := C(K_u, \gamma(u))$
- Let (u, v) be the 2nd round output of V^* given the input graph G and first round message $c_1, \ldots, c_n$
- Output 1 iff $(u, v) = (u^*, v^*)$, $\gamma(u^*) \neq \gamma(v^*)$, and

$$D((c_1, \ldots, c_n), (u^*, v^*), (K_{u^*}, K_{v^*})) = 1$$

Then we have

$$\mathbb{P}[A(\text{Distribution 2}) = 1] = \sum_{\gamma : \gamma(u^*) \neq \gamma(v^*)} \frac{3}{2} \cdot \frac{1}{3^n} \cdot DA(\gamma) \tag{4}$$

Because A, on input Distribution 2, first prepares commitments to a coloring chosen uniformly at random among all $1/(6 \cdot 3^{n-2})$ colorings such that $\gamma(u^*) \neq \gamma(v^*)$ and then outputs 1 if and only if, given such commitments as first message, V^* replies with (u^*, v^*) and the resulting transcript is accepted by D.

We also have

$$\mathbb{P}[D(S^*) = 1 \wedge 2R_{u^*,v^*}] = \frac{1}{\mathbb{P}[S^*_{1Round} \neq FAIL]} \cdot \sum_{\gamma : \gamma(u^*) \neq \gamma(v^*)} \frac{1}{3^n} \cdot DA(\gamma) \tag{5}$$

To see why Equation (5) is true, consider that the probability that S^* outputs a particular transcript is exactly $1/\mathbb{P}[S^*_{1Round} \neq FAIL]$ times the probability that S^*_{1Round} outputs that transcript. Also, the probability that S^*_{1Round} outputs a transcript which involves (u^*, v^*) at the second round and which is accepted by $D()$ conditioned on γ being the coloring selected at the beginning is $DA(\gamma)$ if γ is a coloring such that $\gamma(u^*) \neq \gamma(v^*)$, and it is zero otherwise. Finally, S^*_{1Round} selects the initial coloring uniformly at random among all possible 3^n coloring.

From our security assumption on (C, O) and from Lemma 6 in Lecture 27 we have

$$\left| \mathbb{P}[S^*_{1Round} \neq FAIL] - \frac{2}{3} \right| \leq \frac{\epsilon}{4|E|} \tag{6}$$

and so the claim we made in Equation (3) follows from Equation (4), Equation (5), Equation (6) and the fact that if p, q are quantities such that $\frac{3}{2}p \leq 1$, $\frac{1}{q} \cdot p \leq 1$, and $\left|q - \frac{2}{3}\right| \leq \delta \leq \frac{1}{6}$ (so that $q \geq 1/2$), then

$$\left| \frac{3}{2}p - \frac{1}{q}p \right| = \frac{3}{2} \cdot p \cdot \frac{1}{q} \cdot \left| q - \frac{2}{3} \right| \leq 2\delta$$

(We use the above inequality with $q = \mathbb{P}[S^*_{1Round} \neq FAIL]$, $\delta = \epsilon/4|E|$, and $p = \sum_{\gamma:\gamma(u^*)\neq\gamma(v^*)} \frac{1}{3^n} DA(\gamma)$.)

Having proved that Equation (3) holds, we get

$$|\,\mathbb{P}[A(\text{Distribution 1}) = 1] - \mathbb{P}[A(\text{Distribution 2}) = 1]| \geq \frac{\epsilon}{2|E|}$$

where A is an algorithm of complexity at most $2t + O(nr)$. Now by a proof similar to that of Theorem 3 in Lecture 27, we have that (C, O) is not $(2t + O(nr), \epsilon/(2|E|n))$ secure.