ETH Zurich, Department of Computer Science
SS 2017

Dr. Martin Hirt
Chen-Da Liu Zhang

Cryptographic Protocols
Notes 4

Scribe: Sandro Coretti (modified by Chen-Da Liu Zhang)

About the notes: These notes serve as written reference for the topics not covered by the papers that are handed out during the lecture. The material contained therein is thus a *strict* subset of what is relevant for the final exam.

This week, the notes contain the definition of proofs of knowledge and how to show that (most of) the protocols we have seen are proofs of knowledge. Moreover, we discuss a weakening of the zero-knowledge property called *witness hiding* and how to achieve it.

4.1 Definition of Proofs of Knowledge

Proofs of knowledge (POKs) are defined relative to a (efficiently computable) predicate $Q : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}$ (corresponding some **NP**-language L). For some $x \in \{0,1\}^*$, w with $Q(x,w) = 1$ is called a *witness* for x (or, more precisely, for x's membership in L).

To formally define PoKs, one considers a *knowledge extractor*, which is an efficient algorithm K that, by interacting with a prover algorithm P' on some input x, tries to extract a witness w for x. Algorithm K may invoke P' arbitrarily many times and control its random tape.

Definition 4.1. *An interactive protocol (P,V) is a* proof of knowledge for a predicate Q *if there exists a knowledge extractor K such that for any $x \in \{0,1\}^*$, if V accepts an interaction with P' on input x with non-negligible probability, then K (interacting with P') outputs w with $Q(x,w) = 1$.*

4.2 Proving the Proof-of-Knowledge Property

A convenient way of proving that an interactive proof is a proof of knowledge is via the following notion of 2-*extractability*, which we have already encountered (informally) in both the lecture and the exercises.

Definition 4.2. *A three-move round with challenge space $\mathcal{C}$ is* 2-extractable[1] for a predicate Q *if from any two accepting triples (t,c,r) and (t,c',r') with $c \neq c'$ for some input x, one can efficiently compute a w with $Q(x,w) = 1$.*

Theorem 4.1. *An interactive proof (P,V) consisting of s independent 2-extractable three-move rounds in which the challenge is chosen uniformly from some challenge space $\mathcal{C}$ is a proof of knowledge if $1/|\mathcal{C}|^s$ is negligible.*

[1] This is also called *special soundness* in the literature.

Proof. Consider an arbitrary P' and fix $x \in \{0,1\}^*$. Denote by p the probability that V accepts an interaction with P' on input x.

The knowledge extractor K, which interacts with P' and controls its randomness ℓ, works as follows:

1. Choose ℓ uniformly at random.
2. Generate two independent protocol executions between P' with randomness ℓ and V.
3. If V accepts both executions and they have different challenge sequences, identify the first round in which the challenges differ and use 2-extractability to compute a witness w. Otherwise, return to step 1.

First note that since P''s randomness is fixed, the executions generated in step 2 are identical up to the point where V asks a different challenge for the first time. In particular, the first message in that round is the same. Thus, if such a round exists, 2-extractability implies that K indeed recovers w with $Q(x,w) = 1$.

It remains to bound the running time of K. Denote by $f(\ell)$ the probability that V accepts an interaction with P' when the randomness of P' is set to ℓ. Thus, if L denotes the random variable corresponding to the uniform choice of ℓ by K,

$$\mathbf{E}[f(L)] = p.$$

Moreover, the probability that both executions generated in step 2 are accepting is $f(\ell)^2$, and, therefore, the success probability of a single iteration of K is

$$\mathbf{E}[f(L)^2] \geq \mathbf{E}[f(L)]^2 = p^2,$$

where the first step uses Jensen's inequality. (This ignores that with negligible probability $1/|\mathcal{C}|^s$, the two executions are identical.) Hence, K runs in $\mathcal{O}(1/p^2)$ expected time, which is polynomial if p is non-negligible. □

4.3 Weakening Zero-Knowledege: Witness Hiding

Witness Hiding. In certain cases, the zero-knowledge property cannot be achieved. Consider for example the use of interactive proofs as identification protocols (cf. Section 2.5.1), where for efficiency and scalability it is desirable to have constant-round protocols. This can for example be achieved by running several instances of the Fiat-Shamir protocol in parallel or by simply executing a single round of Schnorr's protocol. In either case, the challenge space would, however, be exponentially large, and thus it is not clear whether the protocol would be zero-knowledge. Therefore, one considers a weakening of the zero-knowledge property called *witness hiding (WH)*.

In the following, let (P,V) be an interactive proof for some predicate Q and denote by

$$\mathcal{W}_x := \{w \in \{0,1\}^* \mid Q(x,w) = 1\}$$

the set of witnesses for a particular x.

Informally, an interactive proof of knowledge (P,V) is witness hiding if no cheating verifier V' can convince the honest verifier V *after* interacting with arbitrarily many instances of the prover P. Since for every fixed input x with $\mathcal{W}_x \neq \emptyset$, there is always a cheating verifier V' that can convince V (namely one that has some $w \in \mathcal{W}_x$ hard-wired into its code), WH is defined via an instance distribution: one requires the existence of an efficient *instance generator* G.

Definition 4.3. *An* instance generator *for a PoK is an efficient algorithm G that outputs an instance x with a uniformly random witness $w \in \mathcal{W}_x$.*

Consider now the following random experiment: First, an instance x is sampled along with a witness using G. Then, V' is given x and is allowed to interact with (polynomially) many instances of the prover P on input x using the witness w.[2] In a second phase, V' interacts with (a single instance of) V on the same input x. V' is considered successful if V accepts.

Definition 4.4. *An interactive proof of knowledge (P, V) is* witness hiding *for an instance generator G if every efficient algorithm V' has only negligible success probability in the above experiment.*

Witness Independence. A way of proving that a particular interactive proof is WH is via the notion of *witness independence (WI)* and the so-called *non-collision assumption (NCA)*.

Consider an interaction of P with an arbitrary V' on some input x. Denote by $Z_{x,w}$ the resulting transcript if the witness used by P is w.

Definition 4.5. *A protocol is* witness independent *if for all $x \in \{0,1\}^*$, all $w, w' \in \mathcal{W}_x$, and all V', $Z_{x,w}$ and $Z_{x,w'}$ are identically distributed.*

Often, the predicate Q has the property that for every instance x there are multiple witnesses w, i.e., $|\mathcal{W}_x| > 1$.[3] The non-collision assumption on Q and G says that there exists no efficient algorithm that outputs (x, w, w') with $w, w' \in \mathcal{W}_x$ and $w \neq w'$ such that (x, w) is distributed identically to the output of G.

Theorem 4.2. *If a proof of knowledge (P, V) for a predicate Q is WI and the NCA on Q and G holds, then it is also WH for G.*

Proof. Suppose is not WH for G. Then, there exists a V' with non-negligible probability of succeeding in the WH experiment (see above). V' can be used to violate the NCA on Q and G as follows: Run G to produce (x, w) and let V' interact with instances of P' on x with witness w. If V' afterwards succeeds in convincing V, use the knowledge extractor K for (P, V) on V' to extract a witness w'. Furthermore, by the WI property, the view of V'—and thus that of K—is independent of which witness was used by P'. Thus, $w' \neq w$ with non-negligible probability, since w is a uniformly random element of $\mathcal{W}_x$. □

[2] Note, however, that, unlike the knowledge extractor, V' cannot rewind any of the prover instances.

[3] Unless, of course, there are no witnesses at all.