ETH Zürich, D-INFK
Spring 2018

Prof. Ueli Maurer
Christian Badertscher
Fabio Banfi

Cryptography Foundations
Solution Exercise 12

12.1 Information-Theoretic Authentication Amplification

a) We describe the very simple authenticated channel that transmits one message from Alice to Bob in case the activation sequence is correct.

The authenticated channel AUTH'
(for message space $\mathcal{M}$)

Variables:

- $x \in \mathcal{M} \cup \{\bot\}$: Initialized to $\bot$.
- $\mathsf{ready}, \mathsf{ack}, \mathsf{deliver} \in \{\bot, \top\}$: All initialized to $\bot$.

Interfaces:

- A: On the first input $\xi \in \mathcal{M}$, set $x := \xi$ and $\mathsf{ready} := \top$. On the second (trigger) input send, if $\mathsf{ack} = \top$ set $\mathsf{deliver} := \top$. Ignore all subsequent inputs at this interface.
- B: On the first (trigger) input query, if $\mathsf{ready} = \top$ set $\mathsf{ack} := \top$. On the second (trigger) input read, if $\mathsf{deliver} = \top$ output x, otherwise output nothing. Ignore all subsequent inputs at this interface.
- E: On any (trigger) input read, output $(x, \mathsf{sent}, \mathsf{deliver}, \mathsf{ack})$.

b) The converter π_1, π_2 and the simulator σ are described in Figure 1. As mentioned in the lecture, for the sake of simplicity, we describe the simulator σ in a way that it also performs apparently non-trivial tasks (for example sampling a key). Since this example is an information-theoretically secure construction, we can be more generous in what we call "trivial" operations.

Let $\mathbf{R}$ and $\mathbf{S}$ be as defined in the exercise sheet. We want to show that any distinguisher for $\pi_1^A \pi^B \mathbf{R}$ and $\sigma^E \mathbf{S}$ has advantage at most δ.

We do this following the proof of the Hash-based scheme discussed in the lecture. We define an MBO for both systems $\pi_1^A \pi^B \mathbf{R}$ and $\sigma^E \mathbf{S}$: Let m be the message input at interface A and m' the message potentially injected at interface E (note that we assume that $q_E = 1$, i.e., Eve injects at most one message). The MBO takes the value 1 if and only if $H_K(m) = H_K(m') \wedge m \neq m'$ (the condition is defined on the systems themselves, so in case of $\sigma^E \mathbf{S}$, the random variable K is sampled by the simulator). The two resulting systems are game-equivalent. (Note that when the game is won, Eve could make Bob output the message m'.) The probability of winning the game, i.e., provoking a collision, is at most δ by definition of $H_K(\cdot)$. This is an upper bound on the distinguishing advantage of any distinguisher by Lemma 4.16.

The proof generalizes in a straightforward way to the the case where Eve can inject more, say up to q_E, messages. By a union-bound argument (similar to the one used for Exercise 2.3), the achieved error bound will be worse by a factor of q_E.

Alice's protocol π_1

- On the first input $m \in \mathcal{M}$ at the outside, output m at the inside interface connected to the insecure channel $\overset{\mathcal{M}}{\twoheadrightarrow}$.
- On the second (trigger) input send, input read the inside interface connected to the unary authenticated channel $\overset{\{\diamond\}}{\longleftarrow\bullet}$.
  - If $\diamond$ is received, choose $K \in \mathcal{K}$ uniformly at random, compute $z := H_K(m)$ and output the pair (K, z) at the inside interface connected to the authenticated channel $\overset{\mathcal{K}\times\mathcal{Z}}{\bullet\!\!\longrightarrow}$.
  - Otherwise, take no action.
- Ignore all subsequent inputs.

Bob's protocol π_2

- On the first (trigger) input query at the outside interface, input read to the inside interface connected to the insecure channel $\overset{\mathcal{M}}{\twoheadrightarrow}$.
  - If the set of obtained messages $\mathcal{B} \neq \emptyset$, then output $\diamond$ at the inside interface connected to the unary authenticated channel $\overset{\{\diamond\}}{\longleftarrow\bullet}$. Store the set $\mathcal{B}$.
  - Otherwise, take no action.
- On the second (trigger) input read at the outside, input read to the inside interface connected to the authenticated channel $\overset{\mathcal{K}\times\mathcal{Z}}{\bullet\!\!\longrightarrow}$.
  - If a pair (K, z) is obtained and $\mathcal{B} \neq \emptyset$ then do the following: for each message $m \in \mathcal{B}$, check whether $z = H_K(m)$ holds and output the first message m that satisfies the equation.
- Ignore all subsequent inputs.

The simulator σ

Variables:

- $x \in \mathcal{M} \cup \{\perp\}$: Initialized to $\perp$.
- $\mathcal{B} \subseteq \mathcal{M}$: Initialized to $\emptyset$.
- sent, ack, deliver $\in \{\perp, \top\}$: All initialized to $\perp$.

Interfaces (emulated):

(Recall that on each (trigger) input read at the outside interface, the simulator obtains the values $(x, \mathsf{sent}, \mathsf{deliver}, \mathsf{ack})$ from AUTH'.)

- E of $\overset{\mathcal{M}}{\twoheadrightarrow}$:
  - On input read, obtain x. Once $x \neq \perp$ include it in $\mathcal{B}$. Output $\mathcal{B}$ in any case.
  - On input (inj, m'), add m' to $\mathcal{B}$.
- E of $\overset{\{\diamond\}}{\longleftarrow\bullet}$: On input read, output $\diamond$ if $\mathsf{ack} = \top$ (and $\perp$ otherwise).
- E of $\overset{\mathcal{K}\times\mathcal{Z}}{\bullet\!\!\longrightarrow}$: On input read, once (i.e., the first time that) $\mathsf{deliver} = \top$, sample K uniformly at random, set $z := H_K(x)$, and output (K, z). On later inputs read output the previously sampled pair (K, z).

Figure 1: Converters π_1, π_2 and the simulator σ.

12.2 CBC-MAC and Prefix-Free Encodings

a) We design a distinguisher $\mathbf{D}$ that for $r \geq 6$ and sufficiently large n has advantage larger than the bound of Theorem 6.1. The distinguisher exploits the fact that the given encoding is not prefix-free. First, it queries the n-bit 0-string and obtains the value z_1. Then, it queries the value $0\dots0|10\dots0|z_1$ to obtain the value z_2 (note that θ_r allows those queries since they result in exactly 6 blocks, as shown below). The distinguisher outputs 1 if and only if $z_1 \neq z_2$.

Let F be the function that corresponds to $\mathbf{R}_{n,n}$. If $\mathbf{D}$ is connected to $\theta_r \mathsf{CBC}' \mathbf{R}_{n,n}$, the n-bit 0-string is padded to $0\dots0|10\dots0$ and

$$z_1 = F(F(0\dots0) \oplus 10\dots0).$$

Then, $0\dots0|10\dots0|z_1$ is padded to $0\dots0|10\dots0|z_1|10\dots0$ and we have

$$\begin{aligned} z_2 &= F(F(\underbrace{F(F(0\dots0) \oplus 10\dots0)}_{=z_1} \oplus z_1) \oplus 10\dots0) \\ &= F(F(0\dots0) \oplus 10\dots0) \\ &= z_1. \end{aligned}$$

Therefore, $\mathbf{D}$ never outputs the bit 1 when connected to $\theta_r \mathsf{CBC}' \mathbf{R}_{n,n}$.

Note that the second query has length $3n$ while the first one has length n. In particular, the second query is different from the first one. Hence, if $\mathbf{D}$ is connected to $\theta_r \mathbf{V}_n$, z_2 will be a uniformly random n-bit string independent of z_1. Therefore, we have $z_1 \neq z_2$ with probability $1 - 2^{-n}$ in this case. This implies that the advantage of $\mathbf{D}$ is $1 - 2^{-n}$, which is larger than $\frac{1}{2}r^2 2^{-n}$ for sufficiently large n.

b) We design a prefix-free encoding as follows. Let $x \in \{0,1\}^*$ and let $\ell := |x|$. Append sufficiently many "0"-bits to $0^\ell|1|x|1$ to obtain $\tilde{x}$ such that n divides $|\tilde{x}|$.

To see that this encoding is prefix-free, assume there are $x_1 \neq x_2$ such that the encoding of x_1 is a prefix of the encoding of x_2. This implies that both encodings start with the same number of "0"-bits, i.e., $|x_1| = |x_2| = \ell$. By assumption, there exists a bit string x_3 such that $0^\ell|1|x_1|10\dots0|x_3 = 0^\ell|1|x_2|10\dots0$, which contradicts $x_1 \neq x_2$.

12.3 Uniform Random Functions with Variable Input-Length

We first describe a converter β that constructs a key *and* a URF from a URF, and then use Corollary 6.4 from the lecture notes to construct a VIL-URF.

Let k be the key size of a δ-AUH H and let $t := \lceil (k+m)/n \rceil$. The converter β performs the following setup: It outputs some *fixed* distinct queries $\hat{x}_1, \dots, \hat{x}_t \in \{0,1\}^m$ at the inside interface[1] and combines the returned values into a single string $y = y_1|\dots|y_t$. Let y' be the first k bits of y and y'' be the next m bits of y. The converter β then outputs y' at the first sub-interface of its outside interface. On input $x \in \{0,1\}^m$ at the second sub-interface of its outside interface, β outputs $x \oplus y''$ at the inside interface and then outputs the returned value at the second sub-interface of its outside interface.

We claim that

$$[-,r]\beta[s_r]\mathbf{R}_{m,n} \;\in\; ([\mathbf{U}_k, [r]\mathbf{R}_{m,n}])^{rt2^{-m}} \tag{1}$$

for $s_r = r + t$, where $[-,r]$ restricts access to the connected system to at most r queries to the second sub-interface and does not restrict access to the first sub-interface, i.e., $[-,r][\mathbf{U}_k, \mathbf{R}_{m,n}] \equiv [\mathbf{U}_k, [r]\mathbf{R}_{m,n}]$. To bound the distinguishing advantage $\overline{\langle [-,r]\beta[s_r]\mathbf{R}_{m,n} \,|\, [\mathbf{U}_k, [r]\mathbf{R}_{m,n}] \rangle}$, observe

[1] We have to assume that $t \leq 2^m$ to assure the existence of t distinct elements in $\{0,1\}^m$.

that the bit string y used by β is uniformly random since it is obtained from outputs of $\mathbf{R}_{m,n}$ for distinct inputs. Thus, the value $y' \in \{0,1\}^k$ output at the first sub-interface of $[-,r]\beta[s_r]\mathbf{R}_{m,n}$ is uniformly random and therefore identically distributed to the output at the first sub-interface of $[\mathbf{U}_k, [r]\mathbf{R}_{m,n}]$. Further note that repeated inputs to the second sub-interface of $[-,r]\beta[s_r]\mathbf{R}_{m,n}$ are answered consistently and distinct inputs result in fresh uniformly random outputs. However, these outputs and the outputs at the first sub-interface might not be independent if β produces the outputs at the second sub-interface by querying $\mathbf{R}_{m,n}$ on one of the values $\hat{x}_1, \ldots, \hat{x}_t$. We thus define an MBO $A_1, A_2, \ldots$ by

$$A_i = \begin{cases} 1, & \exists j \in \{1,\ldots,i\}\ \exists j' \in \{1,\ldots,t\} \quad x_j \oplus y'' = \hat{x}_{j'} \\ 0, & \text{else,} \end{cases}$$

where x_i is the ith input to the second sub-interface. Let $[-,r]\hat{\beta}[s_r]\mathbf{R}_{m,n}$ denote the game obtained by enhancing $[-,r]\beta[s_r]\mathbf{R}_{m,n}$ with this MBO. By the argument above, we have

$$[-,r]\hat{\beta}[s_r]\mathbf{R}_{m,n} \ \models\!\!\equiv\ [\mathbf{U}_k, [r]\mathbf{R}_{m,n}].$$

Thus, Theorem 4.23 implies $\overline{\langle [-,r]\beta[s_r]\mathbf{R}_{m,n} \,|\, [\mathbf{U}_k, [r]\mathbf{R}_{m,n}] \rangle} \ \leq\ \Gamma(\mathbf{b}[-,r]\hat{\beta}[s_r]\mathbf{R}_{m,n})$. To analyze this winning probability, let $\mathbf{W}$ be a winner and let $X_1, \ldots, X_r$ be the random variables corresponding to the (non-adaptive) queries to the second sub-interface asked by $\mathbf{W}$. Moreover, let Y'' denote the random variable corresponding to the value y'' produced by β during setup. We then have

$$\begin{aligned} \overline{\mathbf{b}[-,r]\hat{\beta}[s_r]\mathbf{R}_{m,n}}\ (\mathbf{W}) &= \mathsf{Pr}(\exists i \in \{1,\ldots,r\}\ \exists j \in \{1,\ldots,t\} \quad X_i \oplus Y'' = \hat{x}_j) \\ &\leq \sum_{i=1}^{r}\sum_{j=1}^{t} \mathsf{Pr}(Y'' = X_i \oplus \hat{x}_j) \\ &\leq rt2^{-m}, \end{aligned}$$

where the first inequality follows from the union bound and the second inequality follows from the fact that Y'' is distributed uniformly over $\{0,1\}^m$. Since this holds for all winners $\mathbf{W}$, this yields

$$\overline{\langle [-,r]\beta[s_r]\mathbf{R}_{m,n} \,|\, [\mathbf{U}_k, [r]\mathbf{R}_{m,n}] \rangle} \ \leq\ \Gamma(\mathbf{b}[-,r]\hat{\beta}[s_r]\mathbf{R}_{m,n}) \leq rt2^{-m},$$

which concludes the proof of our claim.

Now let α be the converter from Corollary 6.4 and let $\epsilon_{r,l} := \frac{1}{2}r^2\delta(l)$. We then have by Corollary 6.4

$$\tau_{r,l}\alpha[\mathbf{U}_k, [r]\mathbf{R}_{m,n}] \ \in\ (\tau_{r,l}\mathbf{V}_n)^{\epsilon_{r,l}}. \tag{2}$$

Using the composition property of constructions stated in Lemma 5.1 as well as (1) and (2), we obtain

$$\gamma[s_r]\mathbf{R}_{m,n} \ \in\ (\tau_{r,l}\mathbf{V}_n)^h,$$

with $\gamma = \tau_{r,l}\alpha[-,r]\beta$ and $h = rt2^{-m}\ \rho^{\tau_{r,l}\alpha} + \epsilon_{r,l}$. Since $\tau_{r,l}$ already restricts access to r queries and α makes exactly one query to $\mathbf{R}_{m,n}$ per query to its second sub-interface, we have $\gamma \equiv \tau_{r,l}\alpha\beta$. We further have $h = rt2^{-m} + \epsilon_{r,l} = rt2^{-m} + \frac{1}{2}r^2\delta(l)$ because $rt2^{-m}$ is a constant function that does not depend on the distinguisher and therefore attaching $\tau_{r,l}\alpha$ to the distinguisher does not influence the function. Thus, we can set $\alpha' := \alpha\beta$, $\epsilon'_{r,l} := rt2^{-m} + \frac{1}{2}r^2\delta(l)$, and conclude

$$\tau_{r,l}\alpha'[s_r]\mathbf{R}_{m,n} \ \in\ (\tau_{r,l}\mathbf{V}_n)^{\epsilon'_{r,l}}.$$