COSC530: Homework 1 - Solutions

Exercise 1. Fix a blockcipher $E: \{0,1\}^\ell \times \{0,1\}^\ell \rightarrow \{0,1\}^\ell$ (for simplicity we assume the key-length $k = \ell$ here; the construction and proof are easy to extend to the more general case). Define $G: \{0,1\}^\ell \times \{0,1\}^\ell \rightarrow \{0,1\}^{2\ell}$ as

$$G(K, x) \overset{\text{def}}{=} E(K_0, x) \parallel E(K_1, x)$$

where $K_0 = E(K, 0^\ell)$ and $K_1 = E(K, 1^\ell)$. (Here the constants $0^\ell$ and $1^\ell$ are arbitrary; any two distinct inputs will do.) The fact that $G$ is a PRF if $E$ is is captured by the following theorem:

**Theorem 1.** Suppose there is a PRF adversary $A$ against $G$ running in time $t$ and making at most $q$ queries. Then there is a PRF adversary $B$ against $E$ such that

$$\text{Adv}_G^\text{prf}(A) \leq 3 \cdot \text{Adv}_E^\text{prf}(B).$$

Furthermore, the running-time of $B$ is at most $t$, and $B$ makes at most $q$ queries.

**Proof.** In order to bound A’s advantage we consider a sequence of games (aka. “hybrid experiments”) $H_1$ to $H_4$ defined as follows for $i \in \{1, 2, 3, 4\}$:

**Game $H_i$:**

$b' \leftarrow_r A^{O_i(x)}$

Return 1 if $b' = 1$, Else return 0

where

**Oracle $O_1(x)$:**

static $K \leftarrow_r \{0,1\}^\ell$

static $K_0 \leftarrow E(K, 0^\ell)$

static $K_1 \leftarrow E(K, 1^\ell)$

Return $E(K_0, x) \parallel E(K_1, x)$

**Oracle $O_2(x)$:**

static $K_0 \leftarrow_r \{0,1\}^\ell$

static $K_1 \leftarrow_r \{0,1\}^\ell$

Return $E(K_0, x) \parallel E(K_1, x)$

**Oracle $O_3(x)$:**

static $K_1 \leftarrow_r \{0,1\}^\ell$

$y_0 \leftarrow_r \{0,1\}^\ell$

Return $y_0 \parallel E(K_1, x)$

**Oracle $O_4(x)$:**

$y_0 \leftarrow_r \{0,1\}^\ell$

$y_1 \leftarrow_r \{0,1\}^\ell$

Return $y_0 \parallel y_1$

Note that we have pushed some code from the “main” code of the games into the oracles for concision, making use of the “static” keyword. Now, the above games can be used to bound A’s
Instructor: Adam O’Neill
adam@cs.georgetown.edu

advantage as follows:

\[
\text{Adv}^\text{prf}_E (A) = \Pr [A^{G_k}(\cdot) \text{ outputs 1}] - \Pr [A^{\$}(\cdot) \text{ outputs 1}]
\]
\[
= \Pr [H_1 \text{ outputs 1}] - \Pr [H_4 \text{ outputs 1}]
\]
\[
= (\Pr [H_1 \text{ outputs 1}] - \Pr [H_2 \text{ outputs 1}]) + (\Pr [H_2 \text{ outputs 1}] - \Pr [H_3 \text{ outputs 1}]) + (\Pr [H_3 \text{ outputs 1}] - \Pr [H_4 \text{ outputs 1}]).
\]

Above, the first line is by definition, the second by construction of \(H_1\) and \(H_4\), and the third is due to the telescoping sum.

Next, we will give PRF adversaries \(B_1, B_2, B_3\) against \(E\) such that for all \(i \in \{1, 2, 3\}\):

\[
\Pr [H_{i+1} \text{ outputs 1}] - \Pr [H_i \text{ outputs 1}] \leq \text{Adv}^\text{prf}_E (B_i)
\]

where \(B_1, B_2, B_3\) each run in time at most \(t\) and make at most \(q\) queries. These adversaries are given below; let us first check that the theorem follows. Indeed, combining the above we have

\[
\text{Adv}^\text{prf}_G (A) \leq \text{Adv}^\text{prf}_E (B_1) + \text{Adv}^\text{prf}_E (B_2) + \text{Adv}^\text{prf}_E (B_3),
\]

and the theorem follows by taking \(B\) to be whichever of \(B_1, B_2, B_3\) has the largest advantage.

We now define \(B_1, B_2, B_3\), as follows:

**Adversary \(B_1^O(\cdot)\):**
- \(K_0 \leftarrow \mathcal{O}(0^\ell)\); \(K_1 \leftarrow \mathcal{O}(1^\ell)\)
- Run \(A\), replying to its queries as follows:
  - On query \(x\) return \(E_{K_0}(x)\|E_{K_1}(x)\)
  - Until \(A\) halts with output \(b'\)
  - Return 1 if \(b' = 1\), Else return 0

**Adversary \(B_2^O(\cdot)\):**
- \(K_1 \leftarrow \{0, 1\}^\ell\)
- Run \(A\), replying to its queries as follows:
  - On query \(x\) return \(\mathcal{O}(x)\|E_{K_1}(x)\)
  - Until \(A\) halts with output \(b'\)
  - Return 1 if \(b' = 1\), Else return 0

**Adversary \(B_3^O(\cdot)\):**
- Run \(A\), replying to its queries as follows:
  - On query \(x\) do:
    - \(y_0 \leftarrow \{0, 1\}^\ell\)
    - Return \(y_0\|\mathcal{O}(x)\)
  - Until \(A\) halts with output \(b'\)
  - Return 1 if \(b' = 1\), Else return 0

It remains to show that these \(B_1, B_2, B_3\) satisfy Equation 1. We just do the first case; the other cases are similar. We have

\[
\text{Adv}^\text{prf}_E (B_1) = \Pr [B_1^{E_K}(\cdot) \text{ outputs 1}] - \Pr [B_1^{\$}(\cdot) \text{ outputs 1}]
\]
\[
= \Pr [H_1 \text{ outputs 1}] - \Pr [H_2 \text{ outputs 1}]
\]

which is what we wanted to show. Above, the first line is by definition, and the second line is by construction of \(H_1\) and \(H_2\). This concludes the proof.

**Exercise 2.** Unfortunately, the assumption that “AES is an ideal cipher” is false, as it is for any blockcipher \(E\). Indeed, for any blockcipher \(E\) there is an adversary \(A\) that trivially distinguishes oracle access to \(E_K\) (let’s ignore the additional inverse oracle given to \(A\) for simplicity) from oracle access to an ideal cipher with advantage \(1 - 2^{-\ell}\):
Adversary $A^{O(\cdot)}$:
Let $K \in \{0, 1\}^k, x \in \{0, 1\}^\ell$ be arbitrary
$y \leftarrow O(K, x)$
If $E(K, x) = y$ then return 1
Else return 0

Nevertheless, a proof in the ideal cipher model rules out attacks on the higher-level protocol which treat the underlying blockcipher $E$ as a blackbox. This is because such an attack must work regardless of how $E$ is implemented, and therefore must work when it is implemented as an ideal cipher, which we know is impossible (by our assumed proof). Note that many attacks are of this “black-box” variety — e.g., the attacks we saw on ECB and CBC-C.

Of course, in the real world the adversary need not treat the underlying blockcipher $E$ as a blackbox when attacking the higher-level protocol. One might then heuristically hope that the adversary additionally cannot find a “non-blackbox attack” when the underlying blockcipher is implemented as AES. Intuitively, such an attack would have to be due to some “bad interaction” between AES and the higher-level protocol, which we might heuristically assume won’t happen because the design of AES is “sufficiently complicated.”

On the other hand, a proof that assumes the underlying blockcipher $E$ is a strong PRP rules out even non-blackbox attacks (assuming the strong PRP assumption really does hold for $E$).