Fundamentals of Cryptography: Midterm

Wednesday Nov 5, 3-5PM

Problem 1 (1pt) Complete the definition of polynomial growth. For a functions $f : \mathbb{N} \to \mathbb{R}^+$. We say $f(n) = \text{poly}(n)$ if ____ fill the blank ____.

Problem 2 (1pt) Complete the definition of negligible functions. A function $f : \mathbb{N} \to \mathbb{R}^+$ is *negligible*, if ____ fill the blank ____.

Problem 3 (2pt) Complete the definition of PRF. An efficiently-computable function $F : \{0,1\}^\lambda \times \{0,1\}^\lambda \to \{0,1\}^\lambda$ is a PRF (for simplicity, assume key/input/output are of the same length) if for any polynomial-time distinguisher $\mathcal{D}$, the probability $\mathcal{D}$ outputs 1 in the real world is close to (off by at most $\text{negl}(\lambda)$) the probability $\mathcal{D}$ outputs 1 in the ideal world.

Real World:	Ideal World:
The challenger samples _____.	The challenger samples _____.
The distinguisher is allowed/given ____ describe the interaction ____.	The distinguisher is allowed/given ____ describe the interaction ____.
The distinguisher outputs a bit.	The distinguisher outputs a bit.

Problem 4 (2pt) The assumption that PRGs exist is known to be equivalent to the assumption that choose all correct answers

(a) OWFs exist; (b) weak OWFs exist; (c) OWPs exist;
(d) CRHFs exist; (e) PRFs and PRPs exist; (f) $P \neq NP$.

Problem 5 (2pt) Sort the following security definitions, from weakest to strongest.

(a) Authenticated Encryption; (b) CPA-security;
(c) CCA1-security; (d) CCA2-security;
(e) indistinguishable encryptions in the presence of an eavesdropper.

Problem 6 (2pt) Let $f_0, f_1 : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be two PRFs, let $g : \{0,1\}^n \to \{0,1\}^{2n}$ be a length-doubling PRG. Then f' must be a PRF if f' is defined as choose all correct answers.

(a) $f'(k,x) = f_0(k_0,x) \| f_1(k_1,x)$, where $k_0 \| k_1 = g(k)$;
(b) $f'(k,x) = f_b(k_b,x)$, where b is the parity of x, and $k_0 \| k_1 = g(k)$;
(c) $f'(k,x) = f_0(k,x) \oplus f_1(k,x)$; (d) $f'(k,x) = f_0(k, f_1(k,x))$.

Choose any 4 of the following problems (problem 7,8,9,10,11) to solve.

Problem 7 (5pt) Non-cryptographic application uses linear PRGs. A linear PRG can be generically defined as follows:

- It is parameterized by a public matrix $M \in \{0,1\}^{\ell \times \ell}$, where $\ell = \text{poly}(\lambda)$.
- The seed is a non-zero vector $s \in \{0,1\}^{\ell}$.
- Define $s_0 = s$, $s_{i+1} = Ms_i$, and the i-th bit of the output is the first bit in s_i.

Denote this PRG by G_M, the pseduocode of G_M is

```
Given seed s_0 = s
For i = 1, 2, 3, ... :
  Let s_i ← M s_{i-1}
  Output the first bit of s_i
```

Part A. Prove that G_M is not a (secure) PRG. Presents an efficient distinguisher without using M (but may know an upper bound of ℓ).

CSS stream cipher is built on top of two linear PRGs G_A, G_B. The i-th output byte of CSS stream cipher is the addition (mod 256) of the i-th output bytes from G_A and G_B. Its pseudocode is

```
Given seed s_A, s_B for G_A, G_B respectively
For i = 1, 2, 3, ... :
  Run G_A for eight cycles to obtains x_i
  Run G_B for eight cycles to obtains y_i
  Output x_i + y_i (mod 256)
```

Part B. Prove or disprove that CCS stream cipher is a (secure) PRG.

Problem 8 (5pt) Given functions $h_1, \ldots, h_t : \{0,1\}^n \to \{0,1\}^n$, the t-round Feistel network is defined as:

> $\text{Feistel}_{h_1,\ldots,h_t}(x_0, x_1)$ takes (x_0, x_1) as the input.
>
> For each $1 \le i \le t$:
>
> Set $x_{i+1} \leftarrow x_{i-1} \oplus h_i(x_i)$.
>
> Output (x_t, x_{t+1}) as the output.

Let $F : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a secure PRF, we know that

$$F'((k_1,k_2,k_3),(x_0,x_1)) = \text{Feistel}_{F(k_1,\cdot),F(k_2,\cdot),F(k_3,\cdot)}(x_0,x_1)$$
$$F''((k_1,k_2,k_3,k_4),(x_0,x_1)) = \text{Feistel}_{F(k_1,\cdot),F(k_2,\cdot),F(k_3,\cdot),F(k_4,\cdot)}(x_0,x_1)$$

are PRP and strong PRP respectively. In this problem, we wonder what if the round keys in Feistel network are not independent.

Is $F'''((k_1,k_2),(x_0,x_1)) = \text{Feistel}_{F(k_1,\cdot),F(k_2,\cdot),F(k_1,\cdot),F(k_2,\cdot)}(x_0,x_1)$ a strong PRP? Briefly prove your claim.

Problem 9 (5pt) Let $\Pi_{\text{CPA}} = (\mathsf{Gen}_{\text{CPA}}, \mathsf{Enc}_{\text{CPA}}, \mathsf{Dec}_{\text{CPA}})$ be a CPA-secure encryption scheme. Let $(\mathsf{Gen}_{\text{MAC}}, \mathsf{MAC}, \mathsf{Verify})$ be a strongly secure MAC scheme, whose MAC algorithm is deterministic and whose verification algorithm is canonical. Let $F : \{0,1\}^\lambda \times \{0,1\}^\lambda \to \{0,1\}^\lambda$ be a secure PRF. Prove that $\Pi_{\text{AE}} = (\mathsf{Gen}_{\text{AE}}, \mathsf{Enc}_{\text{AE}}, \mathsf{Dec}_{\text{AE}})$ is a secure authenticated encryption scheme, or give a counter-example

- $\mathsf{Gen}_{\text{AE}}(1^\lambda)$ samples three keys: k_{CPA} for CPA-secure scheme, k_{PRF} for PRF, k_{MAC} for MAC.
- $\mathsf{Enc}_{\text{AE}}(k, m)$ samples random $v \leftarrow \{0,1\}^\lambda$, and compute ciphertext c as

$$t = \mathsf{MAC}(k_{\text{MAC}}, m\|v),$$
$$c = \mathsf{Enc}_{\text{CPA}}(k_{\text{CPA}}, m\|v\|t; F(k_{PRF}, v)),$$

  where $\mathsf{Enc}_{\text{CPA}}$ uses $F(k_{PRF}, v)$ as its randomness tape. (W.l.o.g. we may assume $\mathsf{Enc}_{\text{CPA}}$ consumes λ-bit of randomness.)
- $\mathsf{Dec}_{\text{AE}}(k, m)$ decrypts the ciphertext using k_{CPA}. Say the decryption is $m\|v\|t$. If $c = \mathsf{Enc}_{\text{AE}}(k, m; v)$, output m; otherwise, output $\perp$.

Problem 10 (5pt) Let $h : \{0,1\}^* \to \{0,1\}^\lambda$ be a (keyless) collision-resistant hash function (CRHF). We will extend the definition of H, and define the hash value of files and folders.

For simplicity, assume that the file system has a tree structure; a file must be a leaf in the tree; there are no additional attributes besides names; assume that file names and folder names are λ-bit strings; all files/folders in the same folder have different names.

For a file F, its hash value is naturally defined as the hash value of its data.

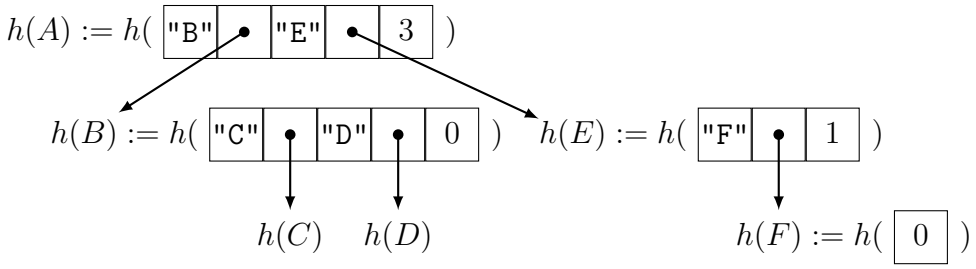
$$h(F) := h(\text{bit string data of } F)$$

For a folder F, assume the it contains files/folders $F_1, F_2, \ldots, F_n$ (sorted by name in alphabetical order) as its immediate children in the file tree, then its hash value is defined as

$$h(F) := h\left(\text{name of } F_1 \,\middle\|\, h(F_1) \,\middle\|\, \ldots \,\middle\|\, \text{name of } F_n \,\middle\|\, h(F_n) \,\middle\|\, \begin{array}{l}\text{total number of folders}\\ \text{in } F \text{ and its subfolders}\end{array}\right)$$

Here is an example.

```
folder A
| folder B
| | file C
| | file D
| folder E
| | folder F
```



Apparently, it is computationally difficult to find two files with different contents but same hash value. Is it computationally hard to find two folders with different contents but same hash value? Prove your statement.

Problem 11 (5pt) Let $P : \{0,1\}^\lambda \times \{0,1\}^\lambda \to \{0,1\}^\lambda$ (together with $P^{-1} : \{0,1\}^\lambda \times \{0,1\}^\lambda \to \{0,1\}^\lambda$) be an efficiently-computable keyed permutation. Is

$$h(x,y) = P(x \oplus y, y) \oplus x \oplus y$$

collision-resistant? Prove one of the following.

- h is a collision-resistant hash function (CRHF) if P is a strong PRP.
- h is not collision-resistant if P is modeled as an ideal cipher.