Problem Set 1

This problem set is due on *Friday, February 24, 2023* at **4:59 PM**. Please note our late submission penalty policy in the course information handout. Please submit your problem set, in PDF format, on Gradescope. *Each problem should be in a separate page.*

You are to work on this problem set in groups. For problem sets 1, 2, and 3, we will randomly assign the groups for the problem set. After problem set 3, you are to work on the following problem sets with groups of your choosing of size three or four. If you need help finding a group, try posting on Piazza. See the course website for our policy on collaboration. Each group member must independently write up and submit their own solutions.

Homework must be typeset in LaTeX*and submitted electronically!* Each problem answer must be provided as a separate page. Mark the top of each page with your group member names, the course number (6.5610), the problem set number and question, and the date. We have provided a template for LaTeX on the course website (see the *Psets* tab at the top of the page).

With the authors' permission, we may distribute our favorite solution to each problem as the "official" solution—this is your chance to become famous! If you do not wish for your homework to be used as an official solution, or if you wish that it only be used anonymously, please note this in your profile on your homework submission.

Problem 1-1. Pseudorandom functions and one-way functions Let $f : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^m$ be a pseudorandom function (PRF) with keyspace $\mathcal{K}$. For each of the following functions g, determine if g is necessarily a PRF. If so, explain in 1-3 sentences why it is a PRF, and if not, provide an attack.

(a) $g : \mathcal{K}^2 \times \{0,1\}^n \to \{0,1\}^{2m}$ where $g((k_1,k_2),x) = (f(k_1,x), f(k_2,x))$.

(b) $g : \mathcal{K} \times \{0,1\}^{2n} \to \{0,1\}^{2m}$ where $g(k,x_1,x_2) = (f(k,x_1), f(k,x_2))$.

(c) $g : \mathcal{K} \times \{0,1\}^{2n} \to \{0,1\}^{m}$ where $g(k,x_1,x_2) = f(k,x_1) \oplus f(k,x_2)$.

(d) $g : \mathcal{K} \times \{0,1\}^{n-1} \to \{0,1\}^{2m}$ where $g(k,x) = (f(k,x||0), f(k,x||1))$.

Let $f : \{0,1\}^n \to \{0,1\}^m$ be a one-way function (OWF). For each of the following functions g determine if g is necessarily a OWF. If so, explain in 1-3 sentences why it is a OWF, and if not, provide an attack.

(e) $g : \{0,1\}^n \to \{0,1\}^{2m}$ where $g(x) = (f(x), 0^m)$.

(f) $g : \{0,1\}^{n/2} \to \{0,1\}^m$ where $g(x) = f(x, 0^{n/2})$, where we assume for simplicity that n is even.

(g) $g : \{0,1\}^{2n} \to \{0,1\}^m$ where $g(x_1,x_2) = f(x_1) \oplus x_2$.

(h) $g : \{0,1\}^n \to \{0,1\}^{m+1}$ where $g(x) = f(x)||x[0]$, where $x[0]$ is the first bit of x

Problem 1-2. From functions to permutations

If $f : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ is a function, then the permutation $D_f : \mathcal{K} \times \{0,1\}^{2n} \to \{0,1\}^{2n}$ associated with f is defined as $D_f(k,(x,y)) = (y, x \oplus f(k,y))$. Suppose that f is a PRF.

(a) Is D_f a PRP (for any such f)? If so, explain your answer, and if not, provide an attack.

(b) Define $D_f^2 : \mathcal{K}^2 \times \{0,1\}^{2n} \to \{0,1\}^{2n}$ as $D_f^2((k_1,k_2),(x,y)) = D_f(k_2,(D_f(k_1,(x,y))$. Is D_f^2 a PRP? If so, explain your answer, and if not, provide an attack.

(c) *Extra Credit.* Argue that D_f^2 has the following weaker security property: For any $m = \mathsf{poly}(n)$, the distributions $\{z_i, D_f^2((k_1,k_2),z_i)\}_{i\in[m]}$ and $\{z_i,u_i\}_{i\in[m]}$ are computationally indistinguishable, where z_i, u_i are uniformly distributed in $\{0,1\}^{2n}$ and where k_1,k_2 are uniformly distributed in $\mathcal{K}$.

Problem 1-3. Pseudorandom permutations

Let $F\colon \mathcal{K} \times \{0,1\}^n$ be a pseudorandom permutation (PRP) with keyspace $\mathcal{K}$. This problem analyzes two constructions, which use the PRP F to construct a PRP on domain $\{0,1\}^{n-1}$.

- The first PRP construction $f_F(k,x)$ is as follows:
  - Given input $x \in \{0,1\}^{n-1}$, compute $F(k, 0\|x)$.
  - If $F(k,0\|x) = (0\|y)$ for some $y \in \{0,1\}^{n-1}$ output y.
  - Otherwise, compute $F(k, F(k,0\|x))$, $F(k, F(k, F(k,0\|x))$, and so on, until reaching an output of the form $(0\|y) \in \{0,1\}^n$. Output y.
- The second PRP construction $f'_F(k,x)$ is the function that outputs the first $n-1$ bits of $F(k,0\|x)$.

(a) For all PRPs F, will f_F produce an output for every input? In other words, is it true that for every PRP F, no input will cause f_F run forever? Explain why or why not.

(b) Argue that $f_F(k,\cdot)$ is a permutation for all $k \in \mathcal{K}$. (Recall that for all PRPs F and for all $k \in \mathcal{K}$, $F(k,\cdot)$ is a permutation.)

(c) Suppose that F is such that for all keys $k \in \mathcal{K}$ and all $x \in \{0,1\}^{n-1}$, the function $f_F(k,x)$ invokes F *only one time*. Show that F is NOT a PRP. In other words, use this property to attack F as a PRP.

(d) Explain why f'_F is not necessarily a PRP for all PRPs F.

Problem 1-4. Programming and substitution ciphers

On piazza, under handouts, you can find a zip file `pset1.zip` that contains the necessary files for this assignment.

(a) The files `c1.bin` and `c2.bin` are one-time pad encryptions using the exclusive-or operator applied character by character. Your task is to decrypt the messages. Fortunately for you, both messages use the same key. All words in the messages appear in the provided English word list `dict.txt`, are lowercase, and punctuation (except for spaces) is omitted. Provide the decrypted messages in your LaTeXdocument. Your code may be in any language and needs only to work on the provided inputs; please attach your code on gradescope.

(b) *Extra credit.*

The file `substitution.bin` is a substitution cipher of an English language document, using ASCII character encoding. Your task is to figure out the substitution pattern used to encrypt the document. Then, using the substitution, provide the encryption of your kerberos email in lowercase (including the `@mit.edu`) as this answer. Write each byte in hexadecimal notation, e.g 03, a2, 4e. Your code may be in any language and needs only to work on the provided inputs; please attach your code on gradescope.