Cryptography – Homework 1

Discussed on 6/11/2018.

Exercise 1.1 Warmup

Consider a CPU with n-bit registers. The n-bit strings $\{0,1\}^n$ are naturally interpreted as the integers $\mathbb{Z}_{2^n} = \{0,1,2,\ldots,2^n-1\}$ using e.g. either "most significant bit first" or "least significant bit first" interpretation (it is not important which one is used). Addition and multiplication on $\mathbb{Z}_{2^n}$ are defined as usual modulo 2^n, e.g. $1+(2^n-1)=0$ and $2\cdot 2^{n-1}=0$ ("overflow"). We define some sets of functions over $\mathbb{Z}_{2^n}$:

- $M_n := \{f\colon \mathbb{Z}_{2^n} \to \mathbb{Z}_{2^n}\}$ the set of all maps from $\mathbb{Z}_{2^n}$ to $\mathbb{Z}_{2^n}$.
- $P_n := \{f \in M_n \mid f_n \text{is a bijection}\}$ the set of all bijections/permutations.
- $A_n = \{f \in M_n \mid \exists a,b \in \mathbb{Z}_{2^n} \forall x \in \mathbb{Z}_{2^n} : f(x) = a\cdot x + b\}$ the set of "affine" functions.

The (mono-alphabetic) substitution cipher generalizes to $\mathbb{Z}_{2^n}$ by taking a subset of P_n as key space with encryption defined by

$$\mathsf{Enc}_f(m_1||m_2||\ldots||m_l) = f(m_1)||f(m_2)||\ldots||f(m_l) \text{ and } \mathsf{Dec}_f(c_1||c_2||\ldots||c_l) := f^{-1}(c_1)||f^{-1}(c_2)||\ldots f^{-1}(c_l)$$

for $f \in P_n$ and $m_i, c_i \in \mathbb{Z}_{2^n}$.

(a) Compute $|M_n|$, $|P_n|$, $|A_n|$, and $|A_n \cap P_n|$.

Hint: Why does it suffice to discuss whether $f(x) = ax+b$ is injective or not?

(b) How would you store and compute $f(\cdot)$ and $f^{-1}(\cdot)$ for $f \in P_n$ resp. $f \in P_n \cap A_n$?

(c) Show that the substitution cipher is perfectly secret if (1) we restrict the message space to $\mathbb{Z}_{2^n}$ and (2) we choose the key uniformly at random from P_n resp. from $A_n \cap P_n$.

(d) Why is the previous result no contradiction to the fact that frequency analysis can be used to break the classical mono-alphabetic substitution cipher?

Exercise 1.2 Generating large odd random numbers

Consider the following algorithm $\mathcal{A}$

- *Input*: Natural number $n > 1$.
- *Output*: A random odd number in $[2^{n-1}, 2^n - 1]$.
- *Algorithm*:
  (a) Set $x_0 = 1$, $x_{n-1} = 1$.
  (b) For i in 1 to $n-2$: Set $x_i = \text{fair\_coin}()$.
  (c) Let x be the integer represented by the bit string $x_0x_1\ldots x_{n-1}$ with the least-significant bit first x_0.

(a) Determine the distribution of the output of $\mathcal{A}(n)$.

(b) What is (a lower bound on) the probability that the output number is a prime?

Hint: Use that within $[2^{n-1}, 2^n-1]$ there are at least $\frac{2^{n-1}}{n}$ primes if n is sufficiently large ($n \geq 20$ suffices).

Exercise 1.3 Perfect Secrecy

(a) Is the following statement true? Every encryption scheme for which the size of the key space equals the size of the message space, and for which the key is chosen uniformly from the key space, is perfectly secret.

(b) Show that every encryption scheme can be transformed into an equivalent encryption scheme defined over a fixed-length key space. More precisely, show that from any encryption scheme $\mathcal{E} = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ we can build an ES $\mathcal{E}' = (\mathsf{Gen}', \mathsf{Enc}', \mathsf{Dec}')$ such that

1. Gen' simply generates a "truly random" (short for "chosen uniformly at random") key $k' \overset{u}{\in} \{0,1\}^T$ for some sufficiently large T. (T should bound the running time of Gen.)
2. The message and ciphertext spaces of $\mathcal{E}$ and $\mathcal{E}'$ are the same.
3. For every message m and every ciphertext c we have $\Pr_{k \overset{r}{:=} \mathsf{Gen}()}[\mathsf{Enc}_k(m) = c] = \Pr_{k' \overset{r}{:=} \mathsf{Gen}'()}\left[\mathsf{Enc}'_{k'}(m) = c\right]$, i.e. the probabilitiy to encrypt m to c is in both ES the same when the key is generated by Gen resp. Gen'.
4. The running times of Enc' and Dec' are bounded by the running times of $\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec}$.

Exercise 1.4 Eavesdropping game

The q-**eavesdropping game** is the version of the eavesdropping game where $\mathcal{A}$ may pass to the oracle two sequences. $m_0 = (m_0^{(1)}, \ldots, m_0^{(q)})$ and $m_1 = (m_1^{(1)}, \ldots, m_1^{(q)})$, and the oracle returns $c = (c^{(1)}, \ldots, c^{(q)})$ where $c^{(i)} \overset{r}{:=} \mathsf{Enc}_k(m_b^{(i)})$ and $m_b^{(i)} \in \mathcal{M}$.

We call an ES *perfectly secure under q encryptions (using the same key)* if there is no $\mathcal{A}$ which can win the q-eavesdropping game with probability strictly greater than $\frac{1}{2}$.

(a) Show that the one-time pad is not perfectly secure under $q = 2$ encryptions.

(b) Propose a q-time pad which is perfectly secure under q encryptions. Does your scheme remember some "state" between two en/decryptions?

Remark: If you want to know more about "classical" ciphers and how to break them:

- `http://www.mysterytwisterc3.org` A very nice collection of flash- animated crypto-puzzles (Ceasar, Substitution, Permutation, Knapsack,...) but also programming challenges and real ciphers.
- *H.F. Gaines, Cryptanalysis, a study of ciphers and their solution. Dover, 1956.*
- *Simon Singh, The Code Book: The Science of Secrecy from Ancient Egypt to Quantum Cryptography. Anchor, 2000.*