MATC16 Cryptography and Coding Theory
Gábor Pete
University of Toronto Scarborough
gpete at utsc dot utoronto dot ca

Solutions to HW Assignment 4

Problem 1. Peggy claims she knows an RSA plaintext. That is, n, e, c are public, and she claims to know an m such that $m^e \equiv c \pmod{n}$. She wants to prove this to Victor using a zero-knowledge protocol. They perform the following steps:

1. Peggy chooses a random integer r_1 with $\gcd(r_1, n)$=1, and computes $r_2 \equiv m \cdot r_1^{-1} \pmod{n}$.
2. Peggy computes $x_i \equiv r_i^e \pmod{n}$ for $i = 1, 2$, and sends x_1, x_2 to Victor.
3. Victor checks if $x_1 x_2 \equiv c \pmod{n}$.

Give the remaining steps of the protocol. Victor wants to be at least 99% sure that Peggy is not lying. **(2 pts)**

Solution. Victor asks for one of the r_i's, $i = 1$ or 2, randomly. Then he checks if this satisfies $r_i^e \equiv x_i \pmod{n}$. They repeat this 6 more times, with Peggy choosing a new random r_1 each time. (Note that $2^{-7} < 1\%$.)

(Explanation: if Peggy does not know m, then she could still produce r_1 and $x_1 \equiv r_1^e \pmod{n}$ then $x_2 \equiv c \cdot x_1^{-1} \pmod{n}$, but would not have a suitable r_2. Or she could choose r_2 and compute x_2 then x_1 from it, but would not have a suitable r_1. Whatever she does, if Victor asks r_1 or r_2 randomly, she will have only 50% chance of surviving his test.)

Problem 2. List the points on the elliptic curve $\{(x, y) : y^2 \equiv x^3 - 2 \pmod{7}\}$. **(2 pts)**

Solution. Let $x = 0, 1, 2, \dots, 6$, and see which yield quadratic residues (mod 7), hence values of y. The quadratic residues are $1 \equiv (\pm 1)^2$ and $4 \equiv (\pm 2)^2$ and $2 \equiv (\pm 3)^2 \pmod{7}$. We obtain the seven points $(3, 2), (3, 5), (5, 2), (5, 5), (6, 2), (6, 5), \infty$.

Problem 3. Factor $n = 35$ by the elliptic curve method, using the curve $y^2 = x^3 + 26$ and calculating $P \boxplus P \boxplus P$ for $P = (10, 9)$. **(2 pts)**

Solution. Using the addition formulas in the book, you first have to compute the slope $m = dy/dx = 3x^2/(2y) = 300/18 = 100/6 \equiv 100 \cdot 6 \equiv 5 \pmod{35}$, which worked without problems, then plug this into the other formulas to get $P \boxplus P = (5, 16)$. Then you have to calculate the coordinates of $(P \boxplus P) \boxplus P$, starting with the slope $m = (16-9)/(5-10) = -7/5$. But gcd(5,35)= $5 \neq 1$, so this point does not exist, but we don't care, because have just found the nontrivial factor 5 of 35.

Problem 4. On Thursday we will prove that, for any random variable X and any function f, we have $H(f(X)) \le H(X)$. (In words, we cannot increase the entropy by doing something deterministic to X.)

(a) Letting X take on the values ± 1, and letting $f(x) = x^2$, show that it is possible that $H(f(X)) < H(X)$. **(1 pt)**

(b) Show that $H(f(X)) = H(X)$ if and only if f is one-to-one on the set of values that are taken by X with positive probability. **(2 pts)**

Solution. For (a), if $\mathbf{P}[X = 1] = p = 1 - \mathbf{P}[X = -1]$ with $p \notin \{0, 1\}$, then $H(X) = -p \log_2 p - (1-p)\log_2(1-p) > 0$, while $f(X) = 1$ with probability one, hence $H(f(X)) = -1 \log_2 1 = 0$, and we are done.

For (b), if we go back to the proof of the inequality in Exercise 6 (a) on page 343-344, we see that we need to show $H(X \mid f(X)) = 0$ if and only if f is 1-to-1. By definition,

$$H(X \mid f(X)) = \sum_y \mathbf{P}[f(X) = y]\, H(X \mid f(X) = y),$$

where y in the summation runs over all the possible values of $f(X)$. If f is 1-to-1, then, for any y, the condition $f(X) = y$ determines the value of X, i.e., the conditioned random variable $(X \mid f(X) = y)$ takes a single value with probability one, hence its entropy is $H(X \mid f(X) = y) = 0$, and the total sum is 0. On the other hand, if f is not 1-to-1, then there is a y such that $\mathbf{P}[f(X) = y] > 0$ and the conditioned random variable $(X \mid f(X) = y)$ has actual randomness, i.e., its entropy has a non-zero term $-p\log_2 p > 0$ for some $p \notin \{0, 1\}$. Thus the total sum will also be positive.

Problem 5. Consider the Hadamard matrix H that is used in defining the Hadamard code, Example 6 of page 397. Namely, H is the 32×32 matrix whose entry h_{ij} in the ith row and jth column, for $0 \le i, j \le 31$, is given by

$$h_{ij} = (-1)^{a_0 b_0 + a_1 b_1 + \cdots + a_4 b_4},$$

where $i = a_4 \ldots a_0$ and $j = b_4 \ldots b_0$ in binary. For instance, for $i = 31$ and $j = 3$, we have $i = 11111$ and $j = 00011$, hence $h_{31,3} = (-1)^2 = 1$.

Prove that the dot product of any two different rows of H is 0. **(2 pts)**

Solution. Let the index of the two rows be $i = a_4 \ldots a_0$ and $i' = a'_4 \ldots a'_0$. The dot product is then

$$\begin{aligned}\sum_{j=0}^{31} (-1)^{(a_0+a'_0)b_0(j) + \cdots + (a_4+a'_4)b_4(j)} &= \sum_{b_0=0}^{1}\sum_{b_1=0}^{1}\sum_{b_2=0}^{1}\sum_{b_3=0}^{1}\sum_{b_4=0}^{1} (-1)^{(a_0+a'_0)b_0 + \cdots + (a_4+a'_4)b_4} \\ &= \left(\sum_{b_0=0}^{1} (-1)^{(a_0+a'_0)b_0}\right) \cdots \left(\sum_{b_4=0}^{1} (-1)^{(a_4+a'_4)b_4}\right).\end{aligned}$$

If $i \neq i'$, then there is some $k \in \{0, 1, \ldots, 4\}$ with $a_k \neq a'_k$, hence $a_k + a'_k \not\equiv 0 \pmod 2$, hence, in the above product of five factors, the kth factor is $+1 - 1 = 0$, hence the entire product is 0, as we wanted.

Problem 6. The following is a parity check matrix for a binary $[n, k]$ code C:

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

What is n and k? Find a generating matrix for C. List the codewords in C. What is the minimal distance in C? What is the code rate of C? **(4 pts)**

Solution. This is a 4×6 matrix, with a 4×4 identity matrix at the end. Cut that off, transpose the beginning, get a 2×4 matrix, then append a 2×2 identity matrix at the beginning, say, to get

$$G = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

This is a 2×6 generating matrix in systematic form. Clearly, $n = 6$ and $k = 2$. We get all the codewords as the linear combinations of the rows of this G. Since we are over the field $\mathbb{Z}_2$, the linear combinations are just the sums, so we get four codewords: (1 0 1 0 1 1), (0 1 1 1 0 1), (0 0 0 0 0 0), (1 1 0 1 1 0). The minimal distance in a linear code equals the minimal Hamming weight (the number of nonzero coordinates) over all non-zero vectors, which is 4 here. Finally, the code rate in a linear $[n, k]$ code is always k/n, which is $1/3$ here.

(Max possible score: 15 pts)