



Family Name:

First Name:

Section:

Security and Cryptography

Final Exam - Solutions

January 12th, 2010

Duration: 3 hours

This document consists of 17 pages.

Instructions

Electronic communication devices and documents are *not* allowed.

A pocket calculator is allowed.

Answers must be written on the exercises sheet.

This exam contains 3 *independent* exercises.

Answers can be either in French or English. Readability and style of writing will be part of the grade.

Questions of any kind will certainly *not* be answered. Potential errors in these sheets are part of the exam.

You have to put your full name on the first page and have all pages *stapled*.

1 Vigenère Cipher

We formalize the Vigenère Cipher as follows:

- Let $A = \mathbb{Z}_{26}$ denote the alphabet, A^* denotes the set of all finite sequences (or *strings*) of elements in A . For $s \in A^*$ we denote by $|s|$ its length and s_i its i th term for $i = 0, 1, \dots, |s| - 1$.
- The plaintext space, key space, and ciphertext space are A^* .
- We assume that given a random plaintext $X = (X_0, \dots, X_{n-1})$ of length n , all X_i are independent with distribution p . That is

$$\Pr [X = x \mid |X| = n] = \prod_{i=0}^{n-1} p(x_i)$$

- We assume that given a key $K = (K_0, \dots, K_{k-1})$ of length k , all K_i are independent and follow a uniform distribution. That is

$$\Pr [K = \kappa \mid |K| = k] = \frac{1}{26^k}$$

- The ciphertext is defined by

$$Y_i = X_i + K_{i \bmod k} \bmod 26$$

for $i = 0, 1, \dots, n - 1$.

1. Assuming that the key is of length k , what is the entropy of K in terms of bits?

It is $k \log_2(26) \approx 4.7k$.

2. How large should be k to have an equivalent key length of 80 bits?

We should have $k \geq \frac{80}{\log_2(26)} \approx 17.02$ so $k = 18$ should be enough.

3. Given a string s , we define the index of coincidence $I_c(s)$ as the probability that two elements of s selected at random at different positions are equal. Given $c \in A$, let $n_s(c)$ be the number of index positions i such that $s_i = c$.

Show that

$$I_c(s) = \sum_{c \in A} \frac{n_s(c)(n_s(c) - 1)}{|s|(|s| - 1)}$$

We pick two index positions I and J at random such that they are different. That is, for any i and j such that $i \neq j$ we have $\Pr[I = i, J = j] = \frac{1}{|s|(|s|-1)}$. We have $I_c(s) = \Pr[s_I = s_J] = \sum_{c \in A} \Pr[s_I = s_J = c]$. Now, $\Pr[s_I = s_J = c]$ is $\frac{n_s(c)(n_s(c)-1)}{|s|(|s|-1)}$ so we obtain the formula.

4. Let X be a random plaintext of length $n = |X|$. Express the expected value $I_p = E(I_c(X))$ in terms of n and p .

We have $n_s(c) = \sum_{i=0}^{n-1} 1_{X_i=c}$ so $E(n_s(c)) = np(c)$. Similarly, we have $n_s(c)^2 = \sum_{i,j=0}^{n-1} 1_{X_i=X_j=c}$. If $i = j$, we have $E(1_{X_i=X_j=c}) = p(c)$. If $i \neq j$, we have $E(1_{X_i=X_j=c}) = p(c)^2$. So, $E(n_s(c)^2) = np(c) + n(n-1)p(c)^2$. By linearity of E , we thus obtain

$$I_p = E(I_c(X)) = \sum_{c \in A} p(c)^2$$

We denote I_u the value of I_p when p is the uniform distribution.
Deduce I_u from the previous question.

It is $I_u = \frac{1}{26}$

5. Let $n = qk + r$ be the Euclidean division of n by k . We pick I and J different with uniform distribution and let \mathcal{E} be the event that $I \bmod k = J \bmod k$.

Show that $\Pr[Y_I = Y_J | \neg \mathcal{E}] = I_u$.

We have $E(I_c(Y)) = \Pr[Y_I = Y_J]$ where the probability holds over the distribution of I, J, X , and K . Clearly,
--

$\Pr[Y_I = Y_J \neg \mathcal{E}] = \Pr[X_I + K_{I \bmod k} \equiv X_J + K_{J \bmod k} \pmod{26} \neg \mathcal{E}] = I_u$
--

since $K_{I \bmod k}$ and $K_{J \bmod k}$ are independent and uniformly distributed.
--

Show that $\Pr[Y_I = Y_J | \mathcal{E}] = I_p$.

We have

$\Pr[Y_I = Y_J \mathcal{E}] = \Pr[X_I + K_{I \bmod k} \equiv X_J + K_{J \bmod k} \pmod{26} \mathcal{E}] = \Pr[X_I = X_J \mathcal{E}]$

since $K_{I \bmod k} = K_{J \bmod k}$. We split this probability over all possible values of $I \bmod k$. In each case, we obtain something which is I_p on average since all plaintext elements are independent. Thus, $\Pr[Y_I = Y_J \mathcal{E}] = I_p$.
--

Show that

$$\Pr[\mathcal{E}] = \frac{q(2n - k(q + 1))}{n(n - 1)}$$

For $i = 0, 1, \dots, r - 1$, we have $\Pr[I \bmod k = J \bmod k = i] = \frac{(q+1)q}{n(n-1)}$. For $i = r, r + 1, \dots, k - 1$, we have $\Pr[I \bmod k = J \bmod k = i] = \frac{q(q-1)}{n(n-1)}$. Thus,

$$\Pr[\mathcal{E}] = r \frac{(q + 1)q}{n(n - 1)} + (k - r) \frac{q(q - 1)}{n(n - 1)} = \frac{q(2n - k(q + 1))}{n(n - 1)}$$

Deduce the value $E(I_c(Y))$.

By collecting all previous results we have

$$E(I_c(Y)) = I_p \Pr[\mathcal{E}] + I_u(1 - \Pr[\mathcal{E}]) = (I_p - I_u) \Pr[\mathcal{E}] + I_u$$

Using the expression of $\Pr[\mathcal{E}]$ we finally obtain

$$E(I_c(Y)) = (I_p - I_u)q \frac{2n - k(q + 1)}{n(n - 1)} + I_u$$

Using $n \gg 1$, $q \approx \frac{n}{k}$ and $E(I_c(Y)) \approx I_c(Y)$, deduce a formula to estimate k based on $I_c(Y)$.

We have

$$I_c(Y) \approx (I_p - I_u) \frac{n - k}{nk} + I_u$$

We invert the previous formula. We obtain

$$k \approx \frac{1}{\frac{I_c(Y) - I_u}{I_p - I_u} + \frac{1}{n}}$$

2 Secure Channel

1. Assuming that Alice and Bob share a secret key K and want to set up a secure channel, explain what are the properties of
 - message confidentiality
 - message authenticity
 - message integrity
 - message sequentiality

- message confidentiality: only the legitimate receiver can receive the message in clear
- message authenticity: only the legitimate sender can send a message
- message integrity: the message cannot be modified when being transmitted
- message sequentiality: the order of messages in a protocol cannot be modified (no message swap, no repetition, no deletion)

2. The GSM secure channel works by sending $m \oplus A5(KC, Count)$ where KC is an encryption key and $Count$ is an implicit message counter.

Which of the properties of Q. 1 is guaranteed, which is not? Explain precisely your answer. (If the answer is neither a clear *yes* nor a clear *no*, explain why.)

- message confidentiality: protected (assuming that $A5$ is secure and that the key does not leak for external reasons). As a matter of fact, there are several $A5$ algorithms, including some weak ones and they all use the same KC . So, an active adversary can change the cipher to a weak one and recover KC so confidentiality is not guaranteed (but this is not due to the secure channel).
- message authenticity: not specifically protected by a message authentication code. It is protected in the sense that an adversary cannot push a message which makes sense without knowing KC .
- message integrity: clearly not protected. An adversary can XOR a δ of her choice to the transmitted message. The effect is that a cleartext m will be replaced by $m \oplus \delta$.
- message sequentiality: protected by using the counter.

3. The Bluetooth secure channel works by sending $(m\|CRC(m)) \oplus E0(K_c, CLK)$ where K_c is an encryption key, CLK is the clock value, and CRC is a cyclic redundancy check function (i.e. a linear mapping).

Which of the properties in Q. 1 is guaranteed, which is not? Explain precisely your answer. (If the answer is neither a clear *yes* nor a clear *no*, explain why.)

- message confidentiality: protected (assuming that $E0$ is secure and that the key does not leak for external reasons).
- message authenticity: not specifically protected by a message authentication code. It is protected in the sense that an adversary cannot push a message which makes sense without knowing K_c .
- message integrity: Not protected. The CRC protection is void. An adversary can XOR $(\delta\|CRC(\delta))$ for a δ of her choice to the transmitted message. The effect is that a cleartext m will be replaced by $m \oplus \delta$.
- message sequentiality: semi-protected by using the clock value. The message sequence cannot be modified except by deleting some messages.

4. The WEP secure channel works by sending $IV \parallel ((m \parallel \text{CRC}(m)) \oplus \text{RC4}(K, IV))$ where K is an encryption key, IV is an asynchronous initial vector, and CRC is a cyclic redundancy check function (i.e. a linear mapping).

Which of the properties in Q. 1 is guaranteed, which is not? Explain precisely your answer. (If the answer is neither a clear *yes* nor a clear *no*, explain why.)

- message confidentiality: it would be protected if RC4 were secure in this asynchronous mode, but this is not the case. So, it is not protected.
- message authenticity: not specifically protected by a message authentication code. It is protected in the sense that an adversary cannot push a message which makes sense with an unused IV without knowing K . It is enough to know one plaintext and ciphertext to reuse the IV.
- message integrity: Not protected. The CRC protection is void. An adversary can XOR $(\delta \parallel \text{CRC}(\delta))$ for a δ of her choice to the transmitted message. The effect is that a cleartext m will be replaced by $m \oplus \delta$.
- message sequentiality: not protected. IVs are meant to be used in any order and even reused.

5. The TLS protocol works by sending $\text{Enc}_{K_1}(m \parallel \text{MAC}_{K_2}(m \parallel \text{seq}))$ where K_1 and K_2 are two secret keys and seq is an implicit message counter.

Which of the properties in Q. 1 is guaranteed, which is not? Explain precisely your answer. (If the answer is neither a clear *yes* nor a clear *no*, explain why.)

- message confidentiality: protected by encryption.
- message authenticity: protected by a message authentication code.
- message integrity: protected by a message authentication code.
- message sequentiality: protected by using the counter.

6. The biometric passport works by sending $\text{Enc}_{K_{\text{Senc}}}(m) \parallel \text{MAC}_{K_{\text{Smac}}}(\text{Enc}_{K_{\text{Senc}}}(m))$ where K_{Senc} and K_{Smac} are two secret keys.

Which of the properties in Q. 1 is guaranteed, which is not? Explain precisely your answer. (If the answer is neither a clear *yes* nor a clear *no*, explain why.)

- message confidentiality: protected by encryption.
- message authenticity: protected by a message authentication code.
- message integrity: protected by a message authentication code.
- message sequentiality: not protected.

3 TCHO Encryption

The goal of the exercise is to study the TCHO public-key cryptosystem.

- We consider the usual $+$ and \times operations in \mathbb{Z}_2 .
- The plaintext space is $\{0, 1\}$ (we encrypt a single bit) and the ciphertext space is $\{0, 1\}^\ell$ (the ciphertexts are ℓ -bit long).
- The public key is a polynomial of degree d with coefficients in \mathbb{Z}_2 denoted $P(z) = P_0 + P_1z + \dots + P_dz^d$.
- The secret key is a polynomial of degree d_K with coefficients in \mathbb{Z}_2 denoted $K(z) = K_0 + K_1z + \dots + K_{d_K}z^{d_K}$.
- These two polynomials are such that:
 - $P(z)$ divides $K(z)$ in $\mathbb{Z}_2[z]$;
 - $K(z)$ has a total number w of nonzero coefficients which is low. We assume that w is odd.
- We define four elementary operations.
 - **Repetition:** Given a plaintext x , we define the ℓ -bit vector $C(x) = (x, \dots, x)$ (all components of $C(x)$ are equal to x).
 - **LFSR:** Given a d -bit vector $r = (r_0, r_1, \dots, r_{d-1})$, we define its expansion to an ℓ -bit vector ($\ell > d$) by using the relation

$$r_{i+d} = \sum_{j=0}^{d-1} r_{i+j}P_j$$

for $i = 0, \dots, \ell - 1 - d$ in \mathbb{Z}_2 .

Note that this relation is linear. We let $\mathcal{L}_P(r) = (r_0, r_1, \dots, r_{\ell-1})$.

- **Biased sequence:** Given a random seed r' we define $\mathcal{S}_\gamma(r')$ as a random ℓ -bit string such that the probability that each bit is 0 is given by $\frac{1+\gamma}{2}$ (its probability of being 1 is thus $\frac{1-\gamma}{2}$).
- **Cancellation:** Given $y \in \mathbb{Z}_2^\ell$, we define $K \otimes y \in \mathbb{Z}_2^{\ell-d_K}$ by

$$(K \otimes y)_i = \sum_{j=0}^{d_K} y_{i+j}K_j$$

for $i = 0, \dots, \ell - 1 - d$ in \mathbb{Z}_2 .

- **Encryption:** To encrypt the bit x with randomness r and r' , compute:

$$\text{Enc}_P(x; r, r') = C(x) + \mathcal{L}_P(r) + \mathcal{S}_\gamma(r')$$

with component-wise addition over \mathbb{Z}_2 .

1. Show that given $C(x) + \mathcal{S}_\gamma(r')$, the plaintext x can be recovered if γ is not too small. What is the complexity of the attack in terms of ℓ ?

The $C(x)$ is a repetition of x and \mathcal{S}_γ generates bits which are biased towards 0. We can just look at the majority of $C(x) + \mathcal{S}_\gamma(r')$ which is most likely to be equal to x . The complexity is $\mathcal{O}(\ell)$. There is however a probability of giving an incorrect result which is bounded by

$$p = \sum_{i=\frac{\ell}{2}}^{\ell} \binom{\ell}{i} \left(\frac{1}{2}(1+\gamma)\right)^i \left(\frac{1}{2}(1-\gamma)\right)^{\ell-i}$$

2. Show that given $C(x) + \mathcal{L}_P(r)$, the plaintext x can be recovered. What is the complexity of the attack in terms of d ?

Since $K(1) = w \bmod 2 = 1$ and is a multiple of $P(1)$ we must have an odd number of nonzero terms in $P(z)$. It is easy to check if $C(x) + \mathcal{L}_P(r)$ satisfies the linear relation defined by $P(z)$ or its opposite by just looking at the first d terms. The complexity is $\mathcal{O}(d)$.

3. Show that for any $x \in \mathbb{Z}_2$ we have $K \otimes C(x) = (x, x, \dots, x)$.

From the definition of $K \otimes y$ we can see that $(K \otimes C(x))_i = K(x) \bmod 2$ for all i . Since w is odd, we have $K(x) \bmod 2 = x$ so $K \otimes C(x) = (x, x, \dots, x)$.

4. Show that for any $r \in \mathbb{Z}_2^d$ we have $K \otimes \mathcal{L}_P(r) = 0$.

Let $y = \mathcal{L}_P(r)$. Since $K(z)$ is a multiple of $p(z)$, let us write $K(z) = P(z)Q(z)$. We have $K_s = \sum_{i+j=s} P_i Q_j$ so

$$(K \otimes y)_t = \sum_{i,j} y_{t+i+j} P_i Q_j = \sum_j \sum_{i=0}^d y_{t+i+j} P_i$$

Clearly, we have $\sum_{i=0}^d y_{t+i+j} P_i = 0$ for all j and t . So, $K \otimes \mathcal{L}_P(r) = 0$.

5. Show that for a random r' all bits of $K \otimes \mathcal{S}_\gamma(r')$ have the same distribution and a probability of being 0 of $\frac{1}{2}(1 + \gamma^w)$.

Hint: For any i , $(K \otimes \mathcal{S}_\gamma(r'))_i$ is the XOR of exactly w independent bits of bias γ .

For any i , $(K \otimes \mathcal{S}_\gamma(r'))_i$ is the XOR of exactly w independent bits of bias γ so it has a bias of γ^w . Indeed, if b is a random bit of bias γ , it means that the probability of being 0 is $\frac{1}{2}(1 - \gamma)$ so $-\gamma = E((-1)^b)$. If b_1, \dots, b_w are independent of bias γ we have

$$E((-1)^{b_1 \oplus \dots \oplus b_w}) = E((-1)^{b_1 + \dots + b_w}) = E((-1)^{b_1} \times \dots \times (-1)^{b_w})$$

Due to the independence, this is $E((-1)^{b_1}) \dots E((-1)^{b_w}) = (-\gamma)^w$. Since w is odd, this is $-\gamma^w$.

6. Given $\text{Enc}_P(x; r, r')$ and $K(z)$, give an algorithm to recover x . What is its complexity in terms of the parameters d_K and ℓ ?

We compute $K \otimes \text{Enc}_P(x; r, r')$ in time $\mathcal{O}(d_K \ell)$. Due to the previous questions, this must be equal to $(x, x, \dots, x) + K \otimes \mathcal{S}_\gamma(r')$. Assuming that γ^w is not too small and that $\ell - d_K$ is large enough, we can recover x by computing the majority. The complexity is $\mathcal{O}(d_K \ell)$.

7. To study the security, give an algorithm to recover $K(z)$ given $P(z)$, d_K and w . What is its complexity?

Hint: if $K(z) = 1 + \sum_{j=1}^{w-1} z^{ij}$, it satisfies a condition which can be written

$$1 + \sum_{j=1}^{\frac{w-1}{2}} z^{ij} = \sum_{j=\frac{w-1}{2}+1}^{w-1} z^{ij} \pmod{P(z)}$$

We compute a list of many $1 + \sum_{j=1}^{\frac{w-1}{2}} z^{ij} \pmod{P(z)}$ and another list of many $\sum_{j=\frac{w-1}{2}+1}^{w-1} z^{ij} \pmod{P(z)}$ and look for matching. This works with complexity $\mathcal{O}(2^{\frac{w-1}{2}})$ which is not polynomial.