DD2448 Foundations of Cryptography (krypto19)
Take-home Exam

Douglas Wikström, dog@kth.se

May 14, 2019

Abstract

Make sure that you read and understand `Files→THE/solution_rules.pdf` at Canvas before you start. This document details the rules for solving and handing in your solutions.

Problem 1 (5T). Your friend is a high school teacher and you are asked to visit his or her class and talk about cryptography. In preparation for your visit, the students are asked to construct their own ciphers. You may assume that the students have no knowledge of cryptography and that none of them is a genius.

Your job is to write a program that can break their ciphers. Please describe what assumptions you make and how your program works in plain English (not using pseudo-code). Your solution is graded based on completeness, clarity, and conciseness.

Problem 2 (15I). Implement the AES cipher. A detailed description is found on Kattis `https://kth.kattis.com/problems/oldkattis.aes`. Feel free to consult different sources on how to make an efficient implementation, but any borrowed ideas should be explained briefly in the solutions submitted on paper. You must also be prepared to explain in detail what you did and why at the oral exam. Make sure that your code is commented and well structured. Up to 15I points may be subtracted if this is not the case.

Problem 3. In one of the exercises we considered the dangers of failing to verify membership in a subgroup. Here we consider another dangerous implementation pitfall that has to do with sampling. Your task is to read about this problem and explain it to a fellow developer at a block chain company who did not solve this problem, but otherwise know as much as you do.

Task 3.1 (0T). Study the following sources and search for additional information to identify the common feature of the fatal flaws in the SwissPost electronic voting system and the Zcash block chain proof parameters:

- `https://people.eng.unimelb.edu.au/vjteague/UniversalVerifiabilitySwissPost.pdf`
- `https://z.cash/blog/zcash-counterfeiting-vulnerability-successfully-remediated`

Task 3.2 (6T). Give an example of a cryptographic construction or protocol where it is possible make this mistake which is as simple as possible.

Task 3.3 (6T). Provide a clear description for how to avoid the mistake, so that your company survives the block chain hype.

Problem 4. Let $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ be a CCA2 secure cryptosystem with message space $\{0,1\}$. For the purpose of this problem we may take *CCA2 secure* to mean that given a public key pk and a set of ciphertexts, any ciphertext computed by an adversary contains a plaintext that is independently distributed from the plaintexts of the given ciphertexts.[1]

Suppose that a key pair $(pk, sk) = \mathsf{Gen}(1^n)$ is generated and consider the following protocol between a prover $\mathcal{P}$ and verifier $\mathcal{V}$, where the verifier holds sk and wishes to verify that $\mathcal{P}$ holds it as well.

1. $\mathcal{V}$ chooses $b_0 \in \{0,1\}$ randomly, computes $c_0 = \mathsf{Enc}_{pk}(b_0)$, and hands c_0 to $\mathcal{P}$.
2. $\mathcal{P}$ computes $b_0 = \mathsf{Dec}_{sk}(c_0)$, chooses $b_1, \ldots, b_k \in \{0,1\}$ randomly conditioned on $\sum_{i=0}^{k} b_i \bmod 2 = 0$ (for some k that it chooses), computes $c_i = \mathsf{Enc}_{pk}(b_i)$ for $i \in [1,k]$, and hands $c_1, \ldots, c_k$ to $\mathcal{V}$.
3. $\mathcal{V}$ computes $b_i = \mathsf{Dec}_{sk}(c_i)$ for $i \in [1,k]$ and accepts **if** $c_i \neq c_0$ **for all** $i \in [1,r]$ and $\sum_{i=0}^{k} b_i \bmod 2 = 0$, and rejects otherwise.

One can prove that if a malicious prover $\mathcal{P}^*$ is not given the secret key sk as input, then it fails with probability negligibly close to $1/2$, but for the purpose of this problem we simply take this to be $1/2$. Now the interesting part!

Task 4.1 (8T). First $\mathcal{V}$ is unhappy about accepting with probability as large as $1/2$ even when $\mathcal{P}^*$ does not hold sk, so it requires $\mathcal{P}^*$ to execute the protocol r times and only accept if each individual execution accepts. Prove under our simplifying assumptions that the probability that $\mathcal{V}$ then accepts when a malicious prover $\mathcal{P}^*$ does not hold sk is at most 2^{-r}.

Task 4.2 (8T). Then $\mathcal{P}$ complains about the many executions and suggests that they run them in parallel instead. Let $\mathcal{P}_j$ and $\mathcal{V}_j$ be copies of $\mathcal{P}$ and $\mathcal{V}$ for $j \in [1,r]$ (given the same input) and add the subscript j to the messages of the jth execution. In a parallel execution, the messages are bundled to avoid increasing the number of rounds:

$$(c_{j,0})_{j\in[1,r]} \text{ and } \textbf{(fixed typo on this line)}$$
$$\big((c_{j,1}, \ldots, c_{j,k})\big)_{j\in[1,r]}$$

Again, $\mathcal{V}$ accepts if each $\mathcal{V}_j$ accepts and rejects otherwise. Prove that the probability that a malicious prover $\mathcal{P}^*$ not holding sk makes $\mathcal{V}$ accept is bounded by $\beta(r)$ for the smallest $\beta(r)$ you can.[2]

Problem 5. (Construction of a PRG.) Let $F_n : \{0,1\}^{2n} \to \{0,1\}^n$ and consider the function $P_k[F_n]$, where k is a positive polynomial in n, defined by

$$P_k[F_n] : \{0,1\}^n \to \{0,1\}^{k\times n}$$
$$P_k[F_n] : s \mapsto r_0|r_1|\cdots|r_{k-1} \ ,$$

where $r_i = F_n(s|\mathsf{bin}(i))$, $|$ denotes concatenation, and $\mathsf{bin}(i)$ denotes the n-bit base-two representation of the integer $i \in [0, k-1]$. We want to prove that that $\{P_k[F_n]\}_{n\in\mathbb{N}_+}$ is a pseudo random generator.

Task 5.1 (5T). Express the weakest possible assumption you can about $\{F_n\}_{n\in\mathbb{N}_+}$ that enables you to prove this claim mathematically.

[1] You may think of each ciphertext as a physical box that hides the plaintext perfectly.

[2] Think about how a malicious prover can behave when it can coordinate its actions freely for all executions.

Task 5.2 (8T). Prove that if $\{P_k[F_n]\}_{n\in\mathbb{N}_+}$ is not a pseudo-random generator, then your assumption is false by describing and analyzing a reduction. Hint: Use a hybrid argument.

Task 5.3 (3T). Argue that for $n = 256$ it is, given your assumption, reasonable in practice to choose $F_{256}(x) = \mathsf{SHA\text{-}256}(x)$ and provide references to literature if needed.

Task 5.4 (1T). State the exact security of your construction for this instantiation, i.e., if an adversary takes time T to break $P_k[\mathsf{SHA\text{-}256}(x)]$, then how much time T' is needed to violate your assumption about $\mathsf{SHA\text{-}256}$ expressed in terms of k? (In other words, we consider k to be variable here.)

Task 5.5 (1T). A friend suggests to modify the construction to $P'_k[F_n] : s \mapsto r'_0|r'_1|\cdots|r'_{2k-1}$, where r'_i is only the first half of r_i. **Can you explain the rationale behind this suggestion!**

Problem 6. (Sampling) Define the statistical distance between distributions over a finite set Ω with probability functions P_X and P_Y by $\|\mathsf{P}_X - \mathsf{P}_Y\| = \frac{1}{2}\sum_{x\in\Omega} |\mathsf{P}_X(x) - \mathsf{P}_Y(x)|$.

Task 6.1 (5T). Let $q > 2$ be a prime, let X be uniformly distributed over $\{0,1\}^{\lceil \log q\rceil + t}$, define Y by $Y = X \bmod q$, and let Z be uniformly distributed over $\mathbb{Z}_q$. Prove the best bound you can of the form $\|\mathsf{P}_Y - \mathsf{P}_Z\| \leq \beta(t)$.[3] Hint: How close to a cylinder is a roll of tape?

Task 6.2 (4T). Let $X_1, \ldots, X_k \in \mathbb{Z}_q^k$ be uniformly and independently distributed. Prove the best bound you can of the form $\Pr\left[\mathsf{span}(X_1, \ldots, X_k) \neq \mathbb{Z}_q^k\right] \leq \ell(q,k)$. (In other words, what is the probability that the vectors are not linearly independent.)

Task 6.3 (4T). We can clearly combine our two results and use a PRG to sample random vectors which are linearly independent with overwhelming probability when $q \gg k$ using little randomness, but we do not need a fully-blown PRG for this![4]

Consider the function $F(x) = (1, x, x^2, \ldots, x^{k-1})$ and suppose that we define $Y_i = F(S_i)$, where the $S_i \in \mathbb{Z}_q$ are uniformly and independently distributed. Prove the best bound you can of the form $\Pr\left[\mathsf{span}(Y_1, \ldots, Y_k) \neq \mathbb{Z}_q^k\right] \leq f(q,k)$.

[3] This situation is quite common in applied cryptography where we need to sample almost uniformly from field, but only have random bit strings to start with.

[4] What we really need is a distribution over some set of vectors which are linearly independent with high probability when the vectors are chosen independently.