This page is intentionally left blank.
Please do not open this question sheet until you are told to do so.

4810-1184 Algorithms for Information Security and Privacy 2018 – Final Examination

Problem 1

In this problem, we will try to find a relationship between the t-closeness anonymity scheme and the private PAC learning. Suppose that all possible sensitive information, which is annual salary in this problem, is $\{i$ millions: $3 \leq i \leq 11\}$.

Question 1.1: Calculate the Earth Moving Distance for a group of two persons with salary 3 million and 6 million.

Suppose that your answer to Question 1.1 is t.

Question 1.2: Consider a group of two persons with the Earth Moving Distance no smaller than t. Discuss why the differences in salary between the two persons in the groups must be no larger than 3 million.

Consider the following type of queries from the machine learning algorithm:

$$f_i(T) \coloneqq \frac{1}{||T||}\left(\frac{(Salary\ of\ i) - 3}{8}\right)$$

Question 1.3: Discuss why $f_i(T)$ can be written in the form of $\sum_{j \in T} f_i(j)$ where $0 \leq f_i(j) \leq 1$.

Suppose that we group all persons in our table to clusters with two persons. Each cluster must have the Earth Moving Distance no smaller than t. When the query from a machine learning algorithm is f_i, we will look at a cluster $\{i, j\}$ which i belongs to. Then, we will give the machine learning algorithm the following information:

$$f_i'(T) \coloneqq \frac{1}{2||T||}\left(\frac{(salary\ of\ i) - 3}{8} + \frac{(salary\ of\ j) - 3}{8}\right)$$

Question 1.4: Calculate the maximum value for $|f_i'(T) - f_i(T)|$

Question 1.5: Discuss how we can use an SQ learning algorithm with all queries in the form of $f_i(T)$ when the query answers is giving as $f_i'(T)$.

Question 1.6: In the t-closeness anonymity scheme, we require that the Earth Moving Distance at each cluster must be smaller than some value. Discuss why it is also important to require that the Earth Moving Distance must be larger than some value.

4810-1184 Algorithms for Information Security and Privacy 2018 – Final Examination

Problem 2

We will work on divisors in this problem. However, instead of working on divisors for elliptic curve as in the class, we will work on divisors on prime field $\mathbb{F}_{11}$ here. Consider a function f from $\mathbb{F}_{11}$ to real numbers. Suppose that $f(x) = (x \oplus 4)^2$.

Question 2.1: Find all zero points of the function f.

Question 2.2: Calculate the orders of the zero points in Question 2.1.

Question 2.3: Discuss why there is no pole (x such that $f(x) \to \infty$) in the function f.

Question 2.4: What is the divisor of the function f?

Question 2.5: Discuss why there is no function f such that $div(f) = 2[3]$.

Question 2.6: From your answer in Question 2.5, discuss why it is not possible to replace elliptic curve $E(\mathbb{F}_p)$ with the set $\mathbb{F}_p$ to save the computation time.

Problem 3

Elliptic curve cryptography uses 5 times less memory than RSA. In this problem, we will try to further reduce its memory consumption. From the following question, we will consider elliptic curve with Weirstrass equation $y^2 = x^3 + Ax + B$.

Question 3.1: Discuss why for any $x' \in \mathbb{F}_p$, there is at most 2 points on elliptic curve of which the x-coordinate is x'.

Question 3.2: Discuss what it is enough to use 1 bit to remember y-coordinate of all the points on elliptic curve.

From next question, recall that $(\mathbb{F}_p - \{0\}, \otimes)$ is an abelian group.

Question 3.3: Use a property of an abelian group to explain why, for any $k \in \mathbb{F}_p - \{0\}$, $k^p = k$.

Question 3.4: Discuss why, when we have X on the right side of the Weirstrass equation, y is either $X^{p/2}$ or $-X^{\frac{p}{2}}$.

Question 3.5: Discuss the computation time of calculating y-coordinate from x-coordinate.

Question 3.6: Discuss why "compressing" y-coordinate to 1 bit is not very helpful from the computational point of view.