Privacy Enhancing Technologies **Fall 2024**

Problem Set 3 — 23/11/2024 update

Due: Fri, Nov 29 at 11:59pm CET (submit via Gradescope)

Instructions: You **must** typeset your solution in LaTeX using the provided template:

`https://spylab.ai/teaching/pets-f24/hws/template.tex`

Submission Instructions: You must submit your problem set via Gradescope. Please use the course code provided in Moodle to sign up. Note that Gradescope requires that the solution to each problem starts on a **new page**.

Bugs: We make mistakes! If it looks like there might be a mistake in the statement of a problem, please ask a clarifying question on Moodle.

Problem 1: Conceptual Questions [8 points]. For each of the following statements, say whether it is TRUE or FALSE. **Justify your answer in one sentence.**

(a) Which of the following are true in a world where $\mathsf{P} = \mathsf{NP}$.

i. The ZK proof system for secret-shared data from lecture 7 is sound (if both servers are honest).

ii. The ZK proof system for secret-shared data from lecture 7 is zero-knowledge (if the client is honest, and at least one server is honest).

iii. The Randomized Response mechanism with $\gamma = 1/4$ is ϵ-DP with $\epsilon = \ln(3)$.

iv. Let M be an ϵ-DP mechanism, and Enc be a semantically secure encryption scheme with cryptographic security (e.g., AES). Then releasing $\mathsf{Enc}(k, M(D))$ is ϵ-DP (where k is a random key).

(b) You and your friends want to determine which one of you has the lowest salary. You design and run a protocol, at the end of which all your friends learn that their "Big 4 salaries"$^{\text{TM}}$ are higher than yours. This blatant invasion of your privacy could have been avoided if you had used a proper maliciously-secure MPC protocol for the function $f(x_1, \dots, x_n) = \arg\min_i x_i$.

(c) For a database $D = \{d_1, d_2, \dots, d_n\} \in \{0,1\}^n$, the curator replies to each counting query S with $\text{Ans}(S) = \left(\frac{1}{n}\sum_{i\in S} d_i\right) + e$, where $e \in \{-2, +2\}$ is sampled uniformly at random for each query. An adversary can reconstruct the entire database with n subset queries.

Problem 2: Proofs on distributed data [4 points]. In class, we saw zero-knowledge proofs on secret-shared data, in which the verifiers holds additive shares $[x]_A, [x]_B$ of an input $x \in \mathbb{F}_p^n$ (i.e., $x = [x]_A + [x]_B$). The prover convinces the verifiers that $C(x) = 0$ for some circuit C.

Construct a zero-knowledge proof on *distributed* data, where each verifier holds a piece of the input (i.e., $x_A, x_B \in \mathbb{F}_p^{n/2}$) and the prover convinces the verifiers that $C(x_A || x_B) = 0$.

Your proof system should require the prover to send $O(|C|)$ field elements to the verifiers and require the verifiers to exchange $O(1)$ field elements. Zero-knowledge should hold as long as the verifiers don't collude.

You can assume the existence of a sound and zero-knowledge fully-linear PCP with a proof size of $O(|C|)$ field elements, where the verifier makes $O(1)$ queries to the proof.

Problem 3: Data de-anonymization from auxiliary information [12 points]. A credit-card company has a database that contains the transaction history of its customers. The company wants to release this data to the public to enable data-driven research. To prevent privacy leaks, they "anonymize" the data by removing any identifiers (names, addresses, etc.) and release a database $D \in \mathcal{X}^{n \times m}$ where each row contains m transactions made by one individual.

The adversary wants to de-anonymize one individual in the database, and learn something about their spending habits. Assume that for some row $r \in \mathcal{X}^m$, the adversary gets access to auxiliary information $\mathsf{aux}(r) \in (\mathcal{X} \cup \{\perp\})^m$. This auxiliary information consists of information about k transactions selected uniformly at random from the row r. Formally, we have that $r_i = \mathsf{aux}_i$ for $i \in I$, and $\mathsf{aux}_i = \perp$ for $i \notin I$, where the set of indices I is chosen uniformly at random among all subsets of size k of $[m]$.

(a) Let $r, r' \in \mathcal{X}^m$ be two (fixed) rows of transactions that differ in an α-fraction of the indices. Give either a tight upper bound or an algebraic upper bound[1] $f(\alpha, k)$ on the probability that r' is consistent with $\mathsf{aux}(r)$, i.e.,

$$\Pr[r'_i = \mathsf{aux}(r)_i \text{ for all } i \in I] \leq f(\alpha, k)\,,$$

where the probability is taken over the random choice of the indices set I.

(b) Assume that $k > \frac{\log(\alpha/n)}{\log(1-\alpha)}$ for some $\alpha \in (0,1)$. Provide an algorithm that, given $\mathsf{aux}(r)$ (with a uniformly random index set I) for a uniformly random $r \in D$, outputs a guess $\hat{r} \in \mathcal{X}^m$ such that $\hat{r}$ and r agree in at least a $(1-\alpha)$-fraction of the indices, with probability at least $1-\alpha$. This probability is taken over the random choice of r and I, and the randomness (if any) of your algorithm.

(c) We say that a database D is (α, β)*-sparse* if at most for a β-fraction of the rows r of D, there exists another row $r' \in D$ that agrees with r in at least a $(1-\alpha)$-fraction of the indices.

Let D be (α, β)-sparse, and let r be a uniformly random row in D. Let $k > \frac{\log(\alpha/n)}{\log(1-\alpha)}$ as in part (b). Provide an algorithm that, given $\mathsf{aux}(r)$ (with a uniformly random index set I), outputs a guess $\hat{r} \in \mathcal{X}^m$ such that $\hat{r} = r$, with probability at least $1 - \alpha - \beta$. This probability is taken over the random choice of r and I, and the randomness (if any) of your algorithm.

Problem 4: On the tightness of Dinur-Nissim [12 points].

(a) Consider the Randomized Response mechanism from the lecture with $\gamma = 1/4$ applied to a database $D \in \{0,1\}^n$. That is, we release a fixed database $D' \in \{0,1\}^n$ where $D'_i = D_i$ with probability $3/4$ and $1 - D_i$ with probability $1/4$. Since the adversary gets access to the whole database of noisy responses D', they can ask as many (noisy) subset queries as they want.

   i. Show that, with high probability $(1-o(1))$ over the randomness in the Randomized Response mechanism, there is a subset query S for which the adversary incurs error $e(S) = \left|\frac{1}{n}\sum_{i\in S} D_i - \frac{1}{n}\sum_{i\in S} D'_i\right| \in \Omega(1)$.

   ii. Show that the adversary can recover (in expectation) $3/4$ of the entries of D.

   iii. Is this consistent with Theorem 1 of Dinur and Nissim?

(b) Let M be an ϵ-DP mechanism that operates over a database $D \in \mathcal{X}^n$, where $\mathcal{X}$ is a finite discrete set with $|\mathcal{X}| \geq 2$. Show that there exists a distribution over databases D such that, in expectation (over the sampling of D, the randomness of M, and the randomness of the adversary), an adversary can reconstruct at most a fraction $e^\epsilon/|\mathcal{X}|$ of the entries in D from the output $M(D)$.

(c) Suppose we use the Gaussian mechanism to answer any set of k subset queries, with noise of standard deviation $o\left(\frac{\sqrt{k}}{n \ln n}\right)$ (so slightly smaller than what we saw in class). Show that if the adversary can make $k = n$ queries, then this mechanism is blatantly non-private. In other words, show that the adversary can reconstruct all but $o(n)$ entries of the database with high probability.

[1] I.e., involving α and k and basic arithmetic operations, additions, multiplications, exponentials, etc. Try to make the bound tight enough that you can apply it in the next part.

Problem 5: Feedback [0 points]. Please answer the following questions to help us design future problem sets. You are not required to answer these questions, and if you would prefer to answer anonymously, please use this form. However, we do encourage you to provide us feedback on how to improve the course experience.

(a) Roughly how long did you spend on this problem set?

(b) What was your favorite problem on this problem set?

(c) What was your least favorite problem on this problem set?

(d) Any other feedback for this problem set?