MECD 2023/2024

Security and Privacy

Assignment 2: Secure Multiparty Computation

1. Introduction

Objective: Explore the concepts of Secure Multiparty Computation (SMC), as well as the tools and libraries that implement those concepts.

Groups: maximum 2 students
Specify your group and fill the following form:
https://docs.google.com/spreadsheets/d/1z1Fk4NhQnEUlINT4YifpjwUColbD1hxqcBb02zA8cBM/edit?usp=sharing

Final deadline: December 17, 2023

Delivery: Write a final report that should include the reasoning behind the choices made in each of the steps defined below. Include all the relevant results and discussion in your report. Together with the report you should submit the sources of your solutions in a zip file.

Assignment defenses: slots will be made available.

2. Description

Step #1 – Problem definition and selection of dataset

In this step you will have to **define a problem that can be solved with PSI protocols**. Obtain and prepare the dataset related to that problem and execute the protocols as needed to find the intersections between the data.

Your dataset should have two independent parts, each one corresponding to the data of one party. You should clean and uniformize the data of your dataset in some format such as csv. You must make sure that these parts of the dataset include common data.

After you have your dataset established, generate 5 versions of the dataset in which the main varying factor is the size. In practice, you can use 5 subsets of your dataset of different sizes.

(Groups of 2 only) Implement a solution for your problem that uses a trusted centralized party and evaluate the performance in terms of time and exchanged data for this solution.

In your report, explain the problem, list the input sets and intersection. Justify why you think that the selected data type is a good fit for PSI.

Step #2 – Evaluation with Protocols for Private Set Intersection

In the description below we will use part1.csv and part2.csv to refer to each of the parts of your dataset. The Lab exercises of class PL04 and PL05 provide a good introduction to what you should do, but are not the complete answer to this exercise, so explore the tools looking for answers, and document your observations in the report.

Evaluate and compare each of the 3 (groups: also, your solution) protocols with respect to the execution time and exchanged data. Use several versions of your dataset, in which you vary the size of both parties' data.

At the end of each execution, the tool will give you the following outputs: required time, data sent and data received. Create an excel (or use any other tool of your liking) to register these values for each of your runs and create the following graphs:

- create a graph with three lines, one for each of the protocols, where the x axis is the set size and the y axis is the required time;
- create a graph with three lines, one for each of the protocols, where the x axis is the set size and the y axis is the total data exchanged (data sent + data received);

Add the excel table and the graphs to your report and discuss the results.