CS401: Data Privacy

Instructor: Dr. Stacey Truex
Due: Thursday October 21, 2021 5:00pm

Programming Project #3

The goal of this programming project is to expand your familiarity with differential privacy through hands on experience. Projects should be completed individually.

Your project should expand on your Project 2 where you compared the runtime of 2 different SMPC approaches (Paillier and Shamir's Secret Sharing) with the non-private setting. You will now add a differentially private mechanism to your comparison. Additionally, you will expand your analysis to include the accuracy of each method. Your code will therefore support the following four (4) approaches to computing an average of n different randomly generated integer values:

1. No privacy protection.
2. The additively homomorphic Paillier encryption scheme.
3. Shamir's secret sharing. Assume each value is held by a different party (n total parties) and that at half of the parties are trusted (i.e. $t = \lfloor n/2 \rfloor$).
4. Differential privacy (assume a privacy budget of $\varepsilon = 1.0$) and public knowledge of n.

You should consider at least 5 different values of n. You may use the same settings for n from Project 2 or choose to change them if you believe a different set of values would improve your analysis. Your analysis should include, for each approach and value of n:

1. A run-time analysis.
2. An accuracy analysis.

Reminder: Be careful to run your experiments in identical compute settings.

Thinks to consider:

- Impact of set-up process for both Paillier and Shamir's secret sharing.
- Values of n which communicate the trade-offs of SMPC **and** DP.
- Effectively communicating results via well formatted graphs.
- Is one run in each setting sufficient?

Deliverables: Your submission should contain the following items:

1. PDF file containing:
   (a) Graphs presenting the results of your experiments.
   (b) Short ($\leq$ 1 page) discussion on the implications of your results.
2. Compressed folder containing:
   (a) Code which can reproduce your results.
   (b) A README file detailing how to run your code and reproduce your results.