CS 401: Data Privacy

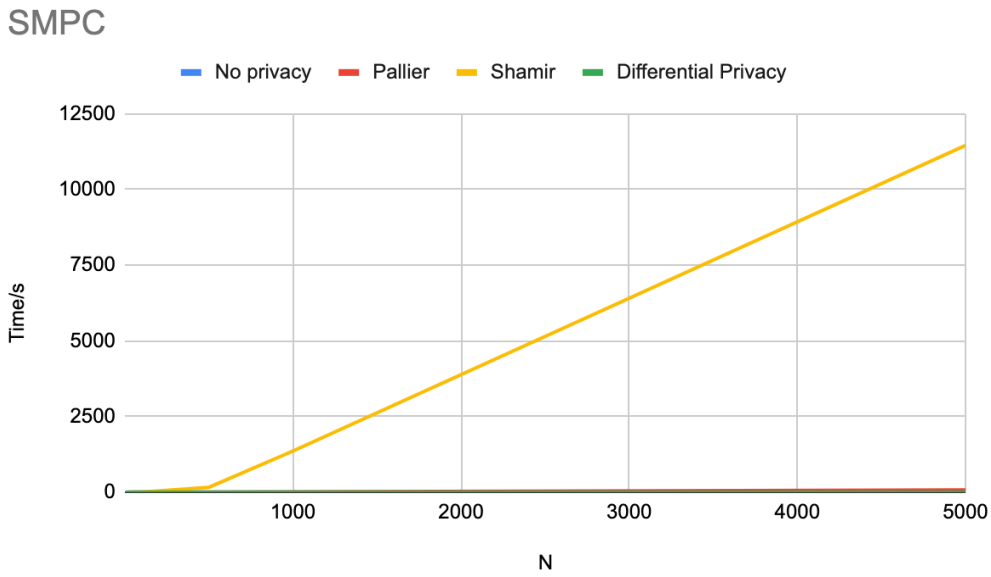
Project 2: Secure Multiparty Computations

Author: Amna Khalid

To explore the implications of various N values on computation time of different privacy schemes, the code for the four - no privacy, Pallier, Shamir's Secret Scheme, Differential Privacy - is run with five different N values. Since randomization is used in this project, the schemes are run multiple times before appending results. The code is computing the average of the N numbers using privacy schemes or not; the results follow:

Time/s Taken by each SMPC Approach				
N	**No privacy**	**Pallier**	**Shamir**	**Differential Privacy**
50	2.15E-06	0.9031407833	0.222990036	4.03E-05
100	1.79E-05	1.74E+00	1.557534933	3.48E-05
500	4.05E-06	7.168787956	160.8889573	9.49E-05
1000	1.00E-05	14.30216312	1356.862339	3.48E-05
5000	3.70E-05	72.13065767	11456.88765	7.80E-05

The table above lists the time taken by each scheme for the corresponding N value. To get a better sense of the values, a graph is generated:



The graph shows how time varies for each different privacy notion. From the graph, the following deductions can be made:

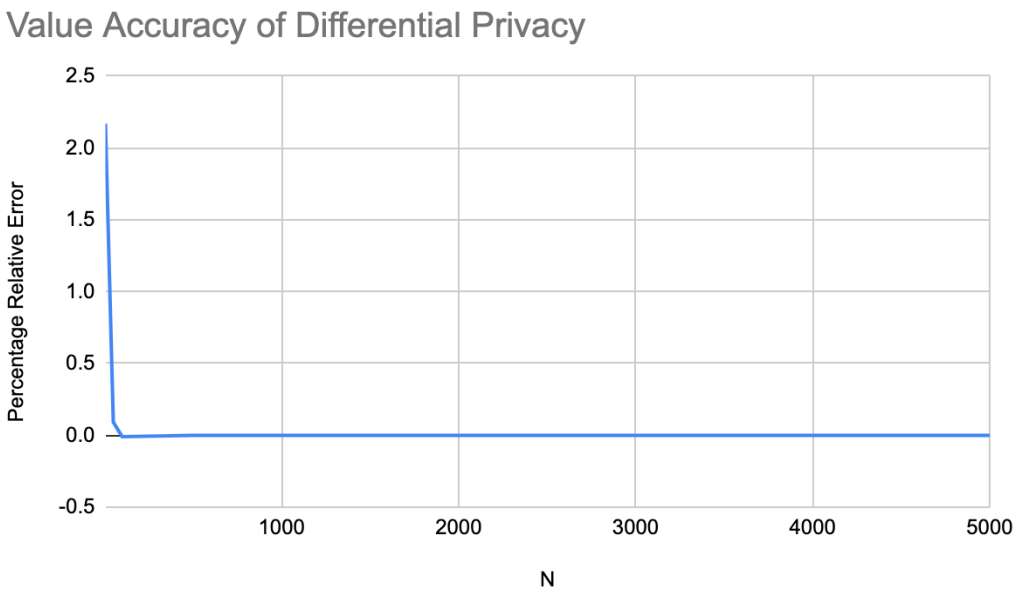
1. Shamir's Secret Scheme is increasing exponentially in comparison to the other privacy schemes showing that Shamir's is running in exponential time.
2. Pallier's takes more time in comparison to no privacy but only by a very slight increase in time can be seen when reaching N =5000. Pallier's in comparison to Shamir's is running in *approximately* constant time.
3. No privacy shows to run have in the least amount of time with the largest N value i.e constant time.

4. Differential privacy corresponds with respect to the time taken time with no privacy i.e it also runs in constant time.

To compute the accuracy of the averages calculated i.e. to observe how far the calculated value is from the true value, percentage relative error is used as a tool of measurement. For an accurate measurement, the percentage relative error will be very low, close to 0% whereas in an inaccurate measurement, the percentage relative error will be higher.

Percentage Relative Error for each SMPC approach				
N	**No privacy**	**Pallier**	**Shamir**	**Differential Privacy**
50	0	0	0	0.09171579147
100	0	0	0	-0.00993077478
500	0	0	0	-1.22E-05
1000	0	0	0	-0.0001457693576
5000	0	0	0	-6.83E-07

All approaches with the exception of Differential Privacy are 100% accurate with their results i.e the percentage relative error is 0%. Differential privacy and its percentage relative error are graphed below:



It is observed as the value of N increases, the accuracy of the value is almost 100% with percentage relative error being almost 0%.

Considering the time and accuracy of each SMPC approach, differential privacy can be considered the best one due to its runtime being minimal and accuracy being >99% with multiple iterations.