COMP 3632 Assignment 3

MENG Zihan
20412027
zmengaa@connect.ust.hk

1.

1) Differential privacy: suitable
   Since you are reluctant to reveal your personal information, by adding noise, your personal data can be protected.
   k-anonymity: suitable
   This method will also add some noise to your personal data.
   Secure multiparty computation: not suitable
   This is irrelevant because we are not comparing two pieces of data.
   Private information retrieval: not suitable
   This is irrelevant since we want to protect the data instead of the query.

2) Secure multiparty computation: suitable
   This method provides a solution to the scenario where two parties with different data can jointly compute a function on the union of their data while sharing no data at all.
   k-anonymity: not suitable
   Adding noise may give incorrect result.
   Differential privacy: not suitable
   Same as k-anonymity, adding noise may give wrong answer.
   Private information retrieval: not suitable
   This is irrelevant because we are not retrieving private data from a database, but comparing two pieces of data.

3) Private information retrieval: suitable
   Since you do not want DNS servers knows your private information, we use private information retrieval to ensure query will be protected.
   k-anonymity: not suitable
   This is irrelevant since it cannot serve the purpose of protecting the query.
   Differential privacy: not suitable
   This is also irrelevant.
   Secure multiparty computation: not suitable
   This is irrelevant since we are not comparing two data.

4) k-anonymity: suitable
   By k-anonymity, each set of identifiers must appear at least k times. This ensures that it won't simply releasing everyone's raw data.
   Differential privacy: suitable
   This method will add noise to the original raw data, which can protect the raw data without losing integrity.
   Secure multiparty computation: not suitable
   This is irrelevant since we are not comparing data.
   Private information retrieval: not suitable
   This is irrelevant since we are not protecting the query.

2.

1) Sunday (full): D
   Monday (differential): pD
   Tuesday (incremental): pD
   Wednesday (differential): $3pD$
   Thursday (incremental): pD
   Friday (differential): $5pD$
   Saturday (incremental): pD
   Sunday (full): D

2) Since a file can be changed twice, for all seven days in the week, the same pD files could be modified. Hence, the above answers for Wednesday and Friday may be different, which means the file changed on Wednesday can differ from pD to $3pD$ with the corresponding changes to Friday.

3) Data on Tuesday cannot be corrupted. Since the next backup is based on this backup, if the backup on Tuesday was corrupted, the next backup could no longer be produced.

3.

1) The cost of winning the reward:
$$\frac{10^{20}}{10^{16}} \times 60 \times 10 \times 0.002 = 12000$$
So to make miners will not suffer a loss, the minimum price of a bitcoin is:
$$12000 \div 6 = 2000$$
Which is $2000

2) Each block has at most 1048576 bytes;
Each transaction has size at least 166 bytes;
Then there exists at most $1048576 \div 166 \div 60 \div 10 \approx 10.53$ transactions per second

3) the transaction fee should be $12000 \div 2 \div (1048576 \div 166) \approx 0.95$

4) Assume we purchased x bitcoin mining devices.
Cost: $8000x + 0.002 \times 3600 \times 24 \times 365x = 71072x$
Reward: $\frac{10^{16}x}{10^{20}+10^{16}x} \times 365 \times 24 \times 6 \times 6 \times 2500 = \frac{788400000x}{10000+x}$
To make at least $1000000 profits per year:
$$\frac{788400000x}{10000 + x} - 71072x > 1000000$$
And we have $x \geq 152$
So the minimum number of bitcoin mining machine we need to buy is 152.