CMPT 403: Written Assignment 3

Vaibhav Saini – 301386847

Question 1

(a)

The total amount of advertised bandwidth of relays with the “Guard” flag: 249.635765144 Gbit/s

The total amount of advertised bandwidth of relays with the “Exit” flag: 19.30846652 Gbit/s

The relays with “Guard” flag are more than relays with “Exit” flag. Tor relays are the nodes that form the Tor network, and they can have different flags. The "Guard" flag is assigned to relays that are suitable for use as the first hop in a Tor circuit (entry node), while the "Exit" flag is assigned to relays that are suitable for use as the last hop in a Tor circuit (exit node).

Hence, "Guard" flag relays play a crucial role in establishing the initial secure connection and are responsible for encrypting the data before it traverses the rest of the network. Because of this, "Guard" relays need to be stable, well-maintained, and have a higher level of uptime to ensure better performance and security of the Tor network.

(b)

For the 50 KiB file:

Download time for 50KiB file : 5.315 s

Download rate for 50KiB file : 9635.34 bytes per second (file size / download time)

For the 5 MiB file:

Download time for 5 MiB file: 28.58 s

Download rate for 5 MiB file: 179028.19 bytes per second (file size / download time)

Latency = 740 ms

Total Time = Latency + File Size / Transfer Rate

Total Load Time for 50 KiB file = 0.74 seconds + (50 * 1024 bytes) / 9635.34 bytes per second => 5.921 seconds

Total Load Time for 5 MiB file = 0.74 seconds + (5 * 1024 * 1024 bytes) / 179028.19 bytes per second => 30.035 seconds

For the 50 KiB file:

Percentage Due to Latency = 0.74 seconds / 6.682 seconds ≈ 0.125 s (about 12.5%)

Percentage Due to Transfer Rate = 1 - Percentage Due to Latency ≈ 0.875 s (about 87.5%)

For the 5 MiB file:

Percentage Due to Latency = 0.74 seconds / 28805 seconds ≈ 0.02464 s (about 2.464%)

Percentage Due to Transfer Rate = 1 - Percentage Due to Latency ≈ 0.97535 s (about 97.535%)

(c)

Traditionally usage of 2 nodes instead of 3 nodes should reduce the amount of time needed to load the files as Load Time = Transfer Time + Latency

Using the values from part (b)

2 Nodes Scenario:

Transfer Time (2 nodes) = File Size / Download Rate = 50 KiB / 9635.34 B/s ≈ 5.19 seconds

Load Time (2 nodes) = Transfer Time (2 nodes) + Latency (1 connection) = 5.19 s + 0.74 s ≈ 5.93 seconds

3 Nodes Scenario:

Transfer Time (3 nodes) = File Size / Download Rate = 50 KiB / 9635.34 B/s ≈ 5.19 seconds

Load Time (3 nodes) = Transfer Time (3 nodes) + Latency (2 connections) = 5.19 s + 2 * 0.74 s ≈ 6.67 seconds

Therefore, if all other conditions remain constant (same transfer rate and latency), it would take approximately 5.93 seconds to load the 50 KiB file using 2 nodes and approximately 6.67 seconds to load the same file using 3 nodes.

(d)

Using 4 nodes instead of 3 nodes in a Tor circuit could potentially improve performance by distributing the load and reducing congestion on individual nodes. However, it might also introduce additional latency as the data passes through more nodes.

Regarding privacy, adding more nodes might provide an extra layer of encryption and anonymity for the traffic, but it could also increase the chances of encountering a malicious or compromised node, which could potentially compromise user privacy.

Question 2

(a) Proposed Technique: Secure Multiparty Computation (SMPC)

Explanation: SMPC allows multiple parties (contestants in this case) to jointly compute a function over their private inputs without revealing those inputs to each other. In the context of the contest, contestants can collaborate to develop and test algorithms on the private data without directly accessing individual user information. This ensures privacy while allowing algorithm development.

Why k-anonymity is not suitable: K-anonymity involves modifying the dataset to ensure that each record is indistinguishable from at least 'k' other records. However, k-anonymity might not be well-suited for a contest scenario where contestants need to access and analyze the raw data. Applying k-anonymity would likely involve preprocessing the data and providing aggregated information, which could hinder the development of accurate and effective recommendation algorithms.

(b) Proposed Technique: Differential Privacy

Explanation: The goal is to protect individual users' location information while still providing useful insights on exercise patterns. By adding noise to the data before releasing it, the smart device company can ensure that no single user's movements can be accurately traced back. This maintains individual privacy while allowing the company to analyze aggregated trends for product improvement.

Why Private Information Retrieval (PIR) is not suitable: PIR is primarily used to retrieve specific data from a database without revealing which specific data point is being requested. It might not directly apply to this scenario, as the goal is not to retrieve specific data points from a remote database, but rather to add privacy to the data collected by the smart device.

(c) Proposed Technique: Private Information Retrieval (PIR)

Explanation: PIR allows you to retrieve information from a database without revealing which specific data you are requesting. In this case, you can check the availability of the web domain without revealing your specific interest to potential cybersquatters. This protects your privacy while ensuring that the domain's availability can be checked.

Why k-anonymity is not suitable: K-anonymity involves data modification to protect individuals' privacy. However, in the context of checking domain availability, k-anonymity might not be

directly applicable, as the goal is not to aggregate or modify data but to query a specific piece of information privately.

(d) Proposed Technique: Differential Privacy

Explanation: In this scenario where individuals want to know if someone in their apartment has been infected with a disease while maintaining their own privacy. By introducing noise to the query results, the hospital can provide aggregate information about potential infections without revealing specific individual data. This ensures privacy for the concerned individuals while providing them with valuable information.

Why Secure Multiparty Computation (SMPC) is not suitable: SMPC involves collaborative computation without revealing individual inputs. However, in this scenario, the primary goal is to aggregate infection information and share it with concerned individuals. Differential privacy is better suited to the task of protecting privacy while releasing aggregated insights.