2 Aid distribution

There is a vast sea with an island located at the geographical point (0, 0). The island calls for help.

Some ships are floating somewhere around the island. Any of the ships is able to bring help. A table `shipData(name, latitude, longitude, speed)` lists these ships with their locations and speeds. There are no obstacles, and each ship can move in a straight direction towards the island.

It is sufficient to get aid from one ship, so it is reasonable that only one of the ships starts moving. This should be the ship that would arrive first. The island needs to know when this ship will arrive to be prepared to receive it. For this, the following queries should be computed:

Q1: When does the first ship arrive?

\[
\text{SELECT MIN}\left(\frac{\sqrt{\text{latitude}^2 + \text{longitude}^2}}{\text{speed}}\right) \text{ FROM shipData;}
\]

Q2: Which ship arrives first?

\[
\text{SELECT name FROM shipData ORDER BY } \frac{\sqrt{\text{latitude}^2 + \text{longitude}^2}}{\text{speed}} \text{ ASC LIMIT 1;}
\]

The ships do not want to reveal their location neither to the island nor the other ships. They want differential privacy.

2.1 A Distance on Database

There are different ways to define distance between two databases. Using standard metric like adding/removing a ship from the table would result in too high impact on the output, and hence in high noise. We can do better knowing that a ship does not care about leaking its name or speed, but only cares about location. Assume that it is fine for a ship if the attacker may only determine its location within 10 miles of precision. Define a suitable distance on tables for differential privacy.

2.2 Adding Differential Privacy

Let \( d \) be the distance on database defined in the previous point.

1. Make the query \( Q1 \) \( \epsilon \)-differentially private w.r.t. \( d \), using Laplace mechanism.

2. Make the query \( Q2 \) \( \epsilon \)-differentially private w.r.t. \( d \), using exponential mechanism.
2.3 Composition

Since the potential attacker observes outputs of both queries, we need to apply composition to ensure $\epsilon$-DP of the system of two queries.

1. Do we need to apply parallel or sequential composition of $Q_1$ and $Q_2$?

2. Assuming that the queries $Q_1$ and $Q_2$ have the same privacy level $\epsilon_1 = \epsilon_2$, estimate how different is their noise level. That is, using inequality $\Pr[\text{noise} \geq \alpha] \leq \beta$ for the particular DP mechanisms in use, compare accuracy $\alpha_1$ of $Q_1$ and $\alpha_2$ of $Q_2$ for the same $\beta$. How should we distribute $\epsilon$ among the two queries to get $\alpha_1 = \alpha_2$ for $\beta = e^{-3} \approx 0.05$? Give particular values for $\epsilon_1$ and $\epsilon_2$, assuming that the whole system should enjoy $\epsilon = 1.0$ differential privacy.