Advanced Topics on Privacy Enhancing Technologies
CS-523
Location Privacy Exercises

1 Who can track me?

Note: This exercise has many possible solutions, we provide in the solutions examples of what can be correct. You can come up with your own solutions and discuss with the TAs.

A big mall has built an Android app which recommends shops to users based on their interests. Initially, the application used GPS to locate users inside the mall. Unfortunately, GPS signals are weak inside buildings and their accuracy is not adequate for locating users inside a mall. To solve this problem, the app developers decided to use nearby Bluetooth and wireless signals to localize users. In this approach, location service providers (LSPs), such as Google, gather the location of wireless access points and Bluetooth beacons into a dataset; Whenever smartphone users want to know their location, they scan for nearby devices and send this data to an LSP to pin-point their location.

After reaching high accuracy on the raw location, the application monitors the location in the background every 30 seconds and sends it to the server to detect which shop the user is visiting. The application always keeps Bluetooth and wireless sensors on for better accuracy. After getting privacy complaints from users, the mall decides to sample a fresh noise for each reading and perturb the location in the app before sending the raw location to the server.

Evaluate the privacy of users and the impact of the location perturbation mechanism against the following adversarial models:

1. The mall which controls the application.
2. The location service provider which localizes the user.
3. A shop which wants to detect frequent visitors.
4. A customer who jogs in the mall.

How can you improve the privacy for each adversary mentioned above?

2 k-anonymity cloaking

Clark is a resident of Smallville, a farming village with a low population density. He has developed SuperApp, a service to recommend sporting facilities to users based on their location. Clark decides to implement a k-anonymity cloaking scheme for SuperApp. The service works as follows: Users have SuperApp installed on their phones. SuperApp sends their location and query to a trusted third-party location anonymization service (LAS). The LAS computes a cloak. A cloak is an area based on the user's location that also contains $k - 1$ other users. The LAS sends the cloak and the query to an untrusted location service provider (LSP). The LSP computes the results for the query and sends it back to the app. Clark is the developer, and can modify the code of SuperApp as well as provide the k value required by the LAS. He cannot modify anything on the LAS or the LSP. Assume that Clark is not malicious.

1. Clark needs to choose an appropriate value of k for his service. What would be a potential issue if Clark chooses a very high value of k?
2. Clark finally picks an appropriate k value for his service. He then tells his friend Lois about it. Clark and Lois decide to attract more customers by introducing SuperApp in Lois' city, Metropolis. Lois, who is in charge of releasing it in Metropolis, looks at the k-anonymization logic used by Clark. She concludes that since Clark has spent so much time fine-tuning the k value, it would provide sufficient privacy for the users of Metropolis. Is Lois correct in her conclusion? Justify.
3. Clark decides that trusting the LAS is risky – if the LAS gets compromised, location data of the users could be revealed. He develops a workaround: the app creates dummy locations and sends them in addition to the real user location. This results in the cloaking area being calculated based on the user location and dummy locations (instead of other user locations in the system), and makes it harder for the LAS to determine the real location of a user. What are points that Clark has to keep in mind while implementing this change?

3 It's all a hoax

In a campaign to relax privacy regulations, a politician stated that "location privacy is just a big hoax". To prove his point, he publicly released the history of his location visits. As a firm supporter of strong privacy protections, you decide to prove him wrong and demonstrate in a short blog post how much sensitive information one can infer from an individual's location history. Describe shortly how you would go about inferring the following information from the politician's data public data:

1. Where does the politician live?

2. Where does he work?
3. Is he religious? If yes, what is his religion?
4. Which companies had dealing with the politician?
5. Does he have a healthy lifestyle?
6. Is he healthy?