CS-523 Final 2023

Location privacy

The EPFL administration wants to improve space utilization on campus. As a first step, EPFL decides to build a mobile application that collects data on how students and staff move on campus. The application works as follows: after installation, each user gets assigned a permanent unique identifier. Every 30 minutes, the application checks whether its user is on campus and if so gets its exact GPS coordinate. To protect privacy, the application generalizes the exact GPS coordinates to the nearest EPFL building before sending it to a central EPFL server. If the user's GPS coordinates are inside a building, the application assigns the user to the building, if the GPS coordinates indicate a point outside a building, the nearest building (measured via the Euclidean distance) is used.

Unique student ID	GPS coordinates	Timestamp
22af7827	BC building	05/06/2023 09:30
9bd657fa	Rolex building	05/06/2023 09:30
5623fab7	CO building	05/06/2023 09:30
22af7827	BC building	05/06/2023 10:00

Table 1: example of 4 anonymized records

Q1: To demonstrate the usefulness of the collected data, the IT team decides to give EPFL professors access to the pseudoanonymized records of 10 random students. The IT team thinks this does not have any privacy implications for the students. Do you agree? If yes, explain why, if not, describe an attack the professors could launch that allows the professor to learn information about the students the professor did not have before the records were available.

Q2: After multiple months, the number of users who opt-in to data collection remains low because rumours have been spreading that the application is not as privacy-friendly as claimed. EPFL thus decides to apply two additional privacy-preserving mechanisms to the GPS data before sending it to the server:

1. Spatial obfuscation: pick, for each user of the application, a random building at EPFL and with probability ¼ assigns the GPS coordinates of the random building instead of the correct building coordinate at every epoch.
2. Hiding: randomly hide (i.e., delete) one location every four epochs per student. Do the additional measures the application more privacy-friendly compared to using only generalization as in Q1? Justify.

Machine Learning /2

To earn some money during your studies, you take on a side job as a developer but quickly realise that it takes up too much of your study time. To save time but still do the job, you want to train a generative machine learning (ML) model that given a function name and a short description writes code for you.

To train the model with as many training samples as possible, you use all the data available (i.e., all your code and other documents) on your laptop. In addition, you offer a classmate 30% of your earnings if they allow you to use the data on their laptop as well. Since they are not willing to give you all their data in clear, you agree to use a collaborative learning algorithm.

In particular, you decide to use decentralized learning, a variant of collaborative learning in which participants communicate model updates directly to each other and locally aggregate and apply model updates. More precisely, at each round, you and your colleague each locally compute a gradient update using simple stochastic gradient descent on your own data and exchange the produced gradient updates with each other. You and your colleague then locally compute the average of these two gradients and apply it on the current model to improve it. You repeat this process for 10 rounds.

Q1: Privacy Attack /1
A while ago, you sent this classmate your precious cs-523 lecture notes, under the condition that they would delete them after reading. You suspect that they actually kept the notes.

(a) Using the information available to you during the decentralized learning protocol, propose an attack that would allow you to infer whether your classmate still has the notes on their laptop. Detail which information you use as input of your attack, and how you use it.

(b) Explain the conditions and the threat model (i.e., knowledge and capabilities) necessary for your attack to be successful, and why they are realistic.

Q2: Privacy Defense /1
As you perform the attack, you are worried your own sensitive data might be at risk. Propose two defenses that would reduce the risk of your colleague inferring the presence of some private content on your laptop:

(a) one that is ML-based (i.e., the ML algorithm is modified), and

(b) another one that is system-based (i.e., the ML algorithm is not modified).

Anonymous comms/censorship

SandCave is a journalist association specialized in leaks of government-sensitive information. To avoid state prosecution for sharing leaks, SandCave’s journalists communicate exclusively using an anonymous messaging service called Blend. Blend is a free and publicly available messaging service that has been audited by experts who confirm that the cryptographic implementation of the system has no vulnerabilities. Blend protect its users from traffic analysis attacks as follows:

1) When a sender sends an encrypted message to the Blend server, the message is put into a queue
2) Every day at precisely 18:00 Blend takes the first 1000 messages in the queue and sends them to their recipients. Before sending, Blend re-encrypts the messages and shuffles them so that they are output on a different order than they arrived. This ensures unlinkability between incoming and outgoing messages.

Given the volume of traffic Blend receives, messages usually do not stay in the system longer than two days. Among all traffic, on average 10 messages per day are from SandCave's journalists.

Malory is a government agent that has infiltrated SandCave. She suspects Alex, a member of SandCave, to directly receive documents from a whistleblower in her governmental department. Malory wants to uncover the whistleblower. To do so, she sets up the following trap to gather evidence that Alex is the one receiving documents from her government department::

1) A colleague from her governmental department creates fake information that they believe the whistleblower will leak to Alex
2) Malory asks SandCave's members if anyone has heard of a government leak and if so whether they could send the leak to her.

Assume that the government can monitor the internet traffic of all SandCave members, and that it has enough money to rent a large number of servers (>100) across the globe.

Q1. If Alex sends the leaked information to Malory using Blend on the same day he received it, can Malory and the government gather evidence that Alex was the one sending the leak through Blend to then interrogate him about the whistleblower? Justify.

Q2. The government wants to disrupt SandCave operations, and finds that the best way is to disrupt the Blend service to prevent SandCave members from talking to each other. However, the government does not want to be associated with freedom of speech violations. Propose a technical censorship mechanism which performs censorship in a covert way and does not require controlling ISP infrastructure.

Online Tracking

Alice uses a browser which prevents cross-domain tracking by blocking third-party cookies. The browser also implements protection against redirect tracking. Additionally, to prevent stateless tracking, Alice modifies the browser to run a low-runtime dynamic analysis of JavaScript code to detect and block all tracking scripts. Assume that the dynamic analysis technique is robust enough to not result in any false negatives.

1) Assume that Alice browses a number of webpages over HTTP. What can an ISP-level adversary do to track Alice across websites/applications? Justify your answer.
2) In spite of using several measures to prevent being tracked, Alice starts seeing some personalised ads.
    a) She suspects that the ISP was tracking her and wishes to take legal action. To whom can she complain about the ISP?
    b) She further switches to using HTTPS for all her communications. Does the attack you proposed in the previous question still work? Justify your answer.