COM 402 exercises 2024, session 12: Privacy: definitions and properties

Exercise 12.1

- Are the following statements true or false? Justify.
  1. It is possible to deploy surveillance only on end-users of systems.
  2. Privacy as control ensures that only the minimal amount of information is provided to the service.
  3. Fine-grained accountability and auditability make it difficult to implement systems with strong privacy protection.

Exercise 12.2

- Consider a privacy-preserving forum to ask questions in the class. To provide privacy, when a student posts a question, instead of publishing the student's name, it chooses uniformly at random another name in the class that starts by the same letter. For the following students, discuss what is the privacy this mechanism gives in terms of error (probability) the professor will face when guessing who wrote a question. Who has more protection?
  - Charlie, who is in the class with Celia, Carla, Constantin, and Colin.
  - Louisa, who is in the class with Lorenz, Lex, and other two Louisas.

Exercise 12.3

- Aggregation is a privacy-protection technique consisting in regrouping data before processing (e.g.: by binning and taking the mean of the data instead of the data itself). Discuss what kind of privacy this is from the point of view of the paradigms (confidentiality, control, practice) or adversary (social, institutional, anti-surveillance) when:
  1. the aggregation is made locally by the user before releasing her data.
  2. the aggregation by all users is made by a third party.

Solutions to the Exercises

Solution 12.1

1. False. Developers and CEOs of companies, Government employees, and in general everyone is at the end of the day an end-user. Once the surveillance infrastructure is deployed, everyone will be under surveillance.
2. False. The paradigm of privacy as control does not really focus on quantity. It focuses on the user knowing how the information is going to be used, but not on minimizing the amount of information disclosed.
3. True. accountability and auditability mainly rely on logging actions. These logs typically record all actions in the system, becoming an extra source of information that can be used to infer private information about users.

Solution 12.2

In both cases the students enjoy anonymity among the other 5 students. The professor has 1/5 probability of guessing correctly, and 4/5 of making an error. In the case of Louisa, the professor succeeds in guessing the correct name 3/5 of the time, but he still cannot know with certainty which of the two Louisas wrote the question. Both students have the same protection

Solution 12.3

We have two cases here:

1. If aggregation is local, from the point of view of the paradigms, aggregation can be seen as an obfuscation mechanism that aims at achieving *privacy as confidentiality.* The idea is to not give the adversary any information about individuals. As such, we can also categorize it as *anti-surveillance* privacy.
2. When aggregation is on a third party, then, with respect to this party that sees all the data the protection is *privacy as control*: we give the data to this party to only perform the aggregation, and only share with other parties the aggregated value. We could still say this is an *anti-surveillance* privacy mechanism from the point of view of the final entity that receives the data, but with respect to the aggregator we would be under *institutional* privacy assuming that this aggregator is semi-trusted and will do what is agreed upon and nothing else.