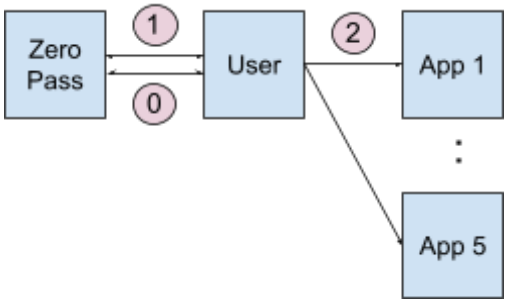
Midterm 2022

Question 1: ABCs

The startup ZeroPass offers a solution for user authentication that removes the need for passwords. The ZeroPass startup supports 5 applications. The diagram below shows the system architecture:



The ZeroPass protocol works in three steps.

Step 0 - Verification: The user provides an email address and verifies they are the owner of this email by clicking on a link sent to the provided address.
Step 1 - Registration: The user and the ZeroPass server engage in an Attribute Based Credential (ABC) issuance protocol, where the issuer-defined attribute is the user's email address. There are no user-defined attributes.
Step 2 - Connection: After registration, the user can connect to the application of its choice by using their credential. To do so, the user runs the ABC disclosure proof and reveals their email address to the application. The application verifies that the user holds a valid ZeroPass credential and uses the email address to identify the user.

Part 1: ZeroPass proudly claims that they guarantee **non-impersonation**, i.e., that no user can successfully connect to an application as another user.
Do you agree with ZeroPass' claim? If yes, argue why the ZeroPass solution guarantees non impersonation. If not, describe an attack that would violate this claim.

Part 2: You have learned a lot from the CS-523 class and notice that revealing the user's email address to all applications is quite privacy invasive. You believe that there are two privacy properties that the ZeroPass solution should provide.

Unlinkability across applications*: when two colluding applications each receive a connection request that uses the ZeroPass credential (step (2)), they cannot tell whether those two connections come from the same user or different users.*
Partial information: *an application cannot learn what other applications a user that uses ZeroPass to authenticate is registered with.*

Obviously, the ZeroPass protocol in its current state does not fulfil either. You want to provide the ZeroPass designers with a protocol that provides both properties in addition to non impersonation.
Notes:

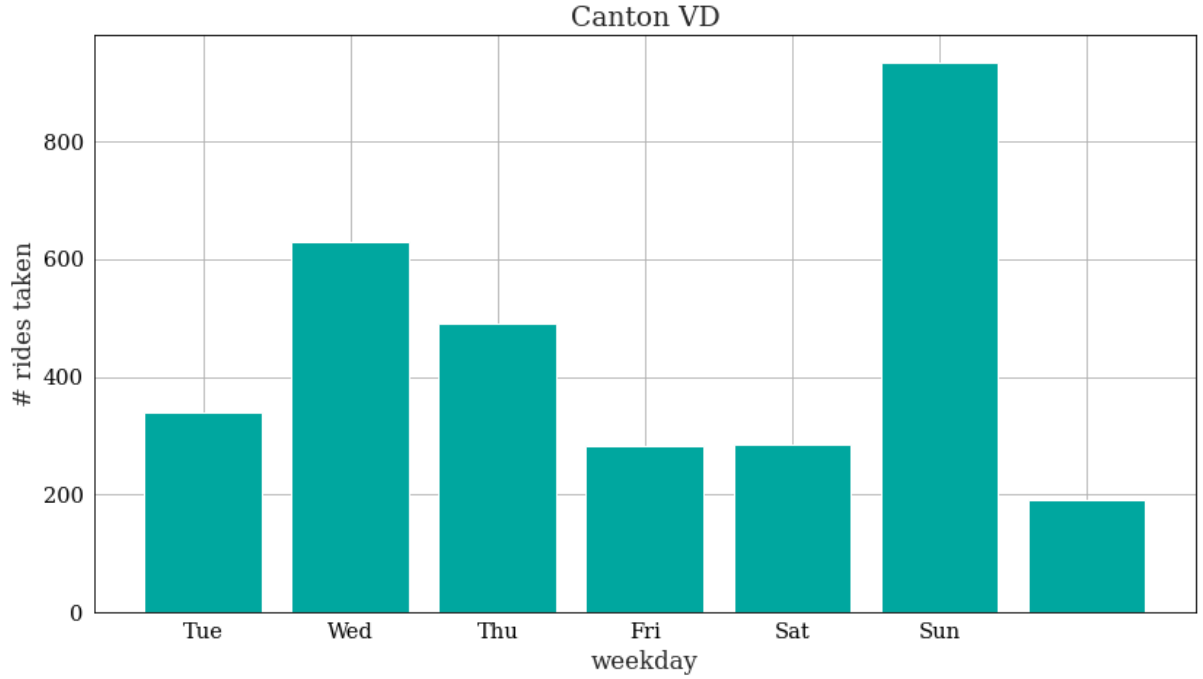
- Applications need to identify their users, but you can change what they are identified with.

- You can assume that network connections are anonymous.

Part 2a: Describe your protocol in the style of steps 0,1,2 above. Specify what kind of ABC scheme you want to use, and explicitly detail what are the attributes, who provides them, and which are hidden and which are disclosed.
Part 2b: Argue why your solution provides non impersonation, unlinkability across applications and partial information.

Question 2: Data Publishing

Paula has been hired as a data analyst by Swüber, a Swiss-German ride-sharing app. Her first task at Swüber is to analyse how many trips that last longer than an hour have been booked in each of the Swiss cantons over the past two weeks. Swüber asks Paula to produce a histogram, like the one shown below, for each canton and for each week. Paula does not have direct access to the trips database. She can only query the data through a standard SQL-style interface.



Part 1: Describe a privacy concern related to the publication of the histogram data and a threat model under which this concern might materialize.

Part 2: Describe a privacy mechanism that Paula could implement which allows her to publish the data in the requested format but mitigates the privacy risk you identified in Part 1. Your answer should sketchan algorithm and specify the details that would be needed to implement this algorithm. Justify why your chosen mechanism addresses the risk you identified in Part 1 within the threat model you defined.

Part 3: Swüber has partnered with the Cantonal Authority (CA) in Vaud to investigate voter participation across communes in the last referendum. The CA asks Swüber to count the number of trips with a polling station as their destination taken on the referendum day. Because the data on whether a user has voted or not is considered highly sensitive, Swüber asks Paula to come up with a method for the app to collect the data from users and produce the statistics in a privacy-preserving way. Neither Swüber nor the CA should be able to learn whether a user has taken a trip to the polling station or not. Describe a privacy mechanism that Paula could implement which fulfils these requirements. Your answer should provide an algorithm and specify the details that would be needed for you to implement this algorithm; and explain why it fulfils the privacy requirements

Question 3: Anonymous Communication

PocketCampus, the developer of the EPFL Campus mobile application, wants to improve the app with a feature that allows students to complain about exam grading. Since students may hesitate to issue complaints under their real name, PocketCampus wants to provide anonymity to the sender of a complaint . They read about Crowds and consider implementing it, treating students as participants of the network.

The PocketCampus Crowd would work as follows: Every student that is online at the moment of complaint submission, automatically participates in the Crowds network. The Crowd is initialized at the beginning of the submission with probability of forwarding p_f equal to 3/4. For simplicity, assume that all participants remain connected until the complaint is received by an EPFL server and that Crowd participants can communicate directly.

Part 1: Assuming the absence of malicious users, does increasing p_f to 7/8 improve sender anonymity towards the EPFL server? Does it degrade the network performance? Justify your answer.

Before implementing the suggested solution, PocketCampus submits the system design to the student representatives for feedback. The student representatives reject the design. They argue that because the EPFL administration controls the Internet infrastructure on campus, the sender of a complaint could still be identified.

Part 2: Do you agree with the students' concern? If yes, describe an attack to identify the sender of a complaint. If not, argue why the suggested design preserves sender anonymity towards the EPFL administration. Assume that all app users are on campus and are using Wi-Fi.

PocketCampus is in a hurry to roll out their new system. They cannot wait for your answer, so they come up with a new design that relies less on EPFL's internet infrastructure in small settings. In this design, PocketCampus replaces the Internet connection between devices with Bluetooth. The Crowds network is formed dynamically between all devices in Bluetooth range. Assume each device is in range of all other

devices. Devices forward messages via Bluetooth. If a device does not forward the package, it uploads it to the server via the Internet. If there is only one participant, submission is not possible.
Part 3: Does this new design provide sender anonymity towards EPFL? If yes, motivate why. If not, describe an attack to identify the sender of a complaint.

Part 4:
PocketCampus realizes that they cannot implement the Crowds protocol with Bluetooth on a campus scale. In the standard Crowd protocol, a package that is forwarded (rather than submitted to the server) will be sent to a randomly chosen node. The range of Bluetooth, however, is not large enough that all devices can directly communicate with other devices. Suppose you modify the Crowds protocol so that forwarding of packages only happens to nodes within Bluetooth range. Does this modification to Crowds degrade sender anonymity towards the server? Justify your answer.

Question 4: Location Privacy
Alice uses a location-obfuscating mechanism when interacting with a location service. Instead of sending her actual location x, she sends an obfuscated location using a randomized obfuscation mechanism $M(x)$ when querying the service.

Alice wants to measure the privacy properties of her mechanism. One relevant privacy notion is distortion privacy, which uses the inference error of an adversary as a privacy metric. The notion of distortion privacy can be formalised as follows: An adversary $\mathcal{A}$, given some background knowledge π, and Alice's obfuscated location $M(x)$, aims to guess her original location, $\hat x = \matcal{A}(\pi, M(x))$. Formally, given some background knowledge π, the privacy metric is the error of the adversary over the randomness of the obfuscation mechanism and over the distribution of Alice's possible locations, $L = E_{x, \mathcal{A}}[(\mathcal{A}(\pi, M(x)) - x)^2]$, where we use a standard mean squared loss to quantify the adversary's error.

Another relevant notion of privacy is geo-indistinguishability. A mechanism satisfies ε-geo-indistinguishability if the following holds for any two locations x and x', and for any subset of locations R:
$$\frac{P(M(x) \in R)}{P(M(x') \in R)} \leq \varepsilon \cdot d(x, x'),$$
where we take $d(a, b)$ to be the Euclidean distance between locations (coordinates).

Part 1. Describe the adversarial models under which these notions quantify privacy? Are they the same? Justify your answer.

Part 2. Is it possible that the mechanism that Alice uses has good (low ε) geo-indistinguishability yet bad (low error L) distortion privacy?

What about bad geo-indistinguishability and good distortion privacy? Justify your answer.

Part 3. Under what threat model should Alice choose distortion over geoindistinguishability to analyze the privacy of her obfuscation mechanism $M(x)$? Justify.