9.1 Large-Scale Differentially Private Data Analysis

9.1.a Scenario

Describe your scenario, i.e., the company or institution, the type of data it wants to analyze, and the goals of this analysis. What assumptions do you need to roll out your method?

In this scenario we would like to know whether an advertisement on YouPipe, which is an online media/video sharing platform. In the beginning of every video an ad is shown in order to generate ad revenue. The website wants to determine whether these leaf to the direct viewing of the associated website or wether the advertisement is dismissed (i.e. the ad is skipped via a button) in order to map certain advertisements to certain videos such that they provide the greatest effect for the advertised company. To be on the safe side the user can agree or disagree to send this sensitive data.

9.1.b Aspects of GOOGLE or APPLE

Which aspects can you adopt from the solutions used by Google or Apple?

YouPipe could adapt the methodology of Apple by first locally privatise the data, like which video was clicked, which ad was shown and whether the advertisement led to any further research of the advertised product or whether it was skipped. This would ensure that impeded attackers cannot gain access to the mainfram and the personal sensitive data, because only a very general obscure form of the data is actually stored on the server.
Also the data can be stripped of the IP from which it was send from, in order to ensure this data cannot be linked to the person that generated it. From these use-case data-centers YouPipe can then analyse the data and then conclude which advertisements do have the most effect for which video.
Furthermore it would be a of advantage to send the locally privatized data at a random time as Apple does in order to mask what videos were watched in which order to ensure this private information is not leaked and accessible at all.

9.1.c Differences of the Deployments

What differs from these deployments?

Because YouPipe is not a hardware manufacturer like Apple or Google but a website it can only access the information that is available via the browser so it might be harder or impossible to work around the seemingly random time at which the data is sent, because for example Apple creates the random time based on system variables and settings to which the website might not have access so YouPipe has to be aware that this might lead to somewhat possible attackpoint in reworking the time at which the data was generated from the point it was sent.