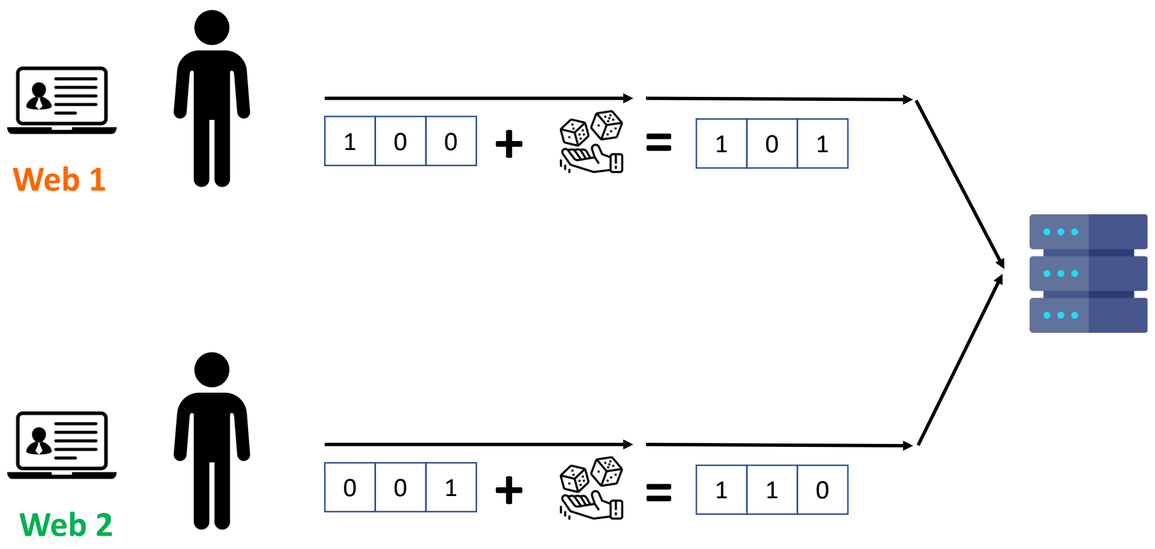
A model-based filter to improve local differential privacy

Juan M. Gutierrez[1], Valérie Gauthier-Umaña[2], Juan F.Pérez[2]

[1]Quantil, Colombia [2]Universidad de los Andes, Colombia

Local differential privacy



A set of users submit a value of interest (e.g., the URL of a website visited by the user) to a central server that accumulates this information to generate summary statistics. Before submission, the user value v is passed through an algorithm A that guarantees its (local differential) privacy, defined as follows:

Definition.(Local Differential Privacy - LDP) Given a privacy budget $\epsilon \geq 0$, an algorithm A satisfies ϵ-local differential privacy if and only if for any input values v_1 and v_2,

$$\Pr(A(v_1) = y) \leq e^{\epsilon} \Pr(A(v_2) = y),$$

For any $y \in Range(A)$, i.e, for all possible outputs of algorithm A.

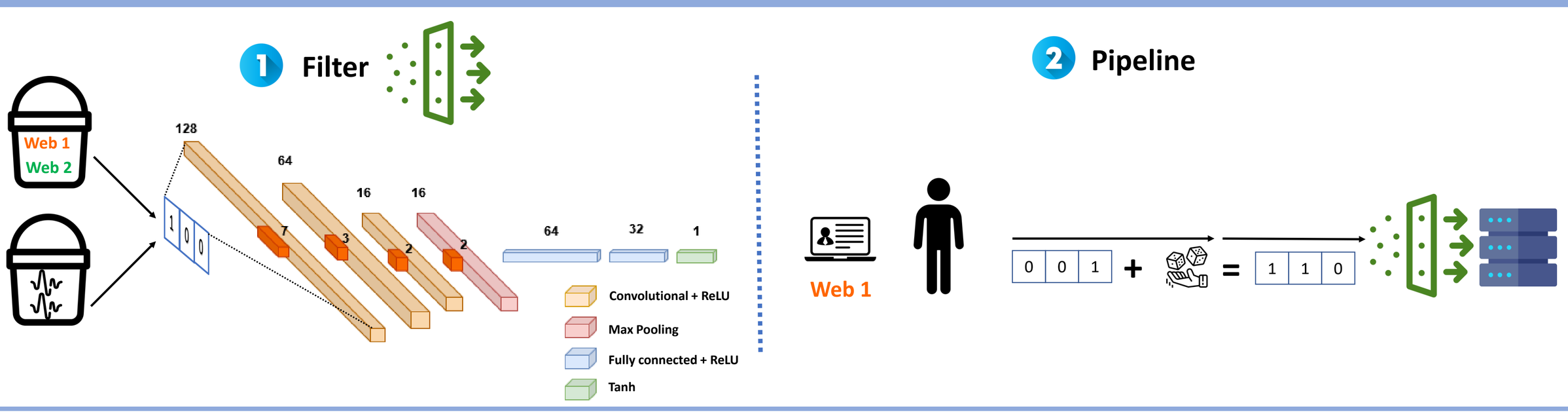
State of Art

A number of LDP protocols have been proposed by Google, Apple, and several researchers. Different protocols vary in their encoding, using local hashing or Hadamard transformations, the domain size reduction using sketches or Bloom Filters, and methods to identify heavy hitters or post-process the data to improve the estimation via rounding and projections.

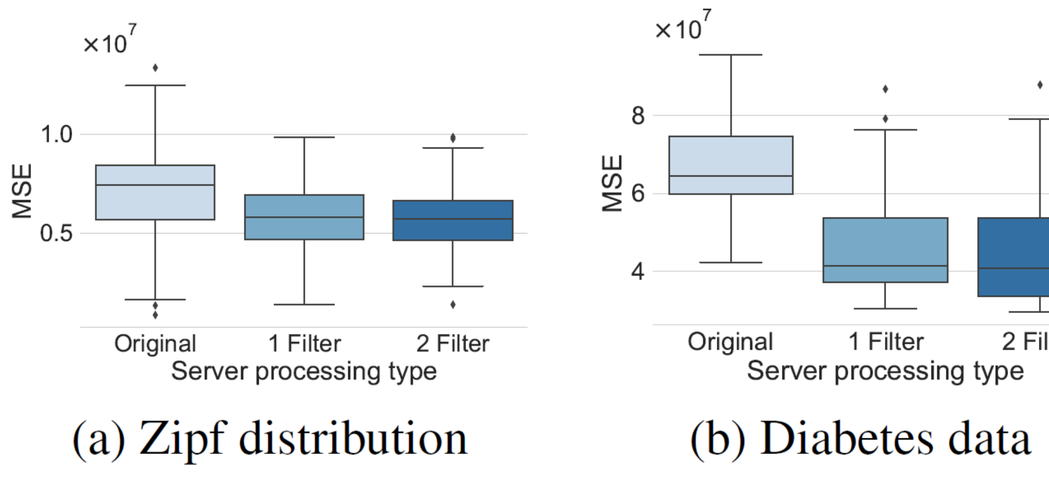
Proposed solution

- We propose to add a filtering step to improve the estimation by removing excessively noisy observations at the server.
- The filter is a classification model trained to identify highly noisy observations.
- **Architecture:** The model consists of a six-layer neural network, where the first three layers are 1D convolutional layers using ReLU activation functions with 128 (kernel size 7), 64 (kernel size 3), and 16 (kernel size 2) neurons, respectively. The last of these convolutional layers is processed by max polling (kernel size 2). The last three layers are dense with 64, 32, and 1 neurons, respectively, where the first two layers use a ReLU activation function, and the last layer uses a hyperbolic tangent activation function. Also, we use a dropout of 0.5 between layers.
- This technique can be combined with any of the existing methods, we choose RAPPOR as a baseline.

A model-based filter

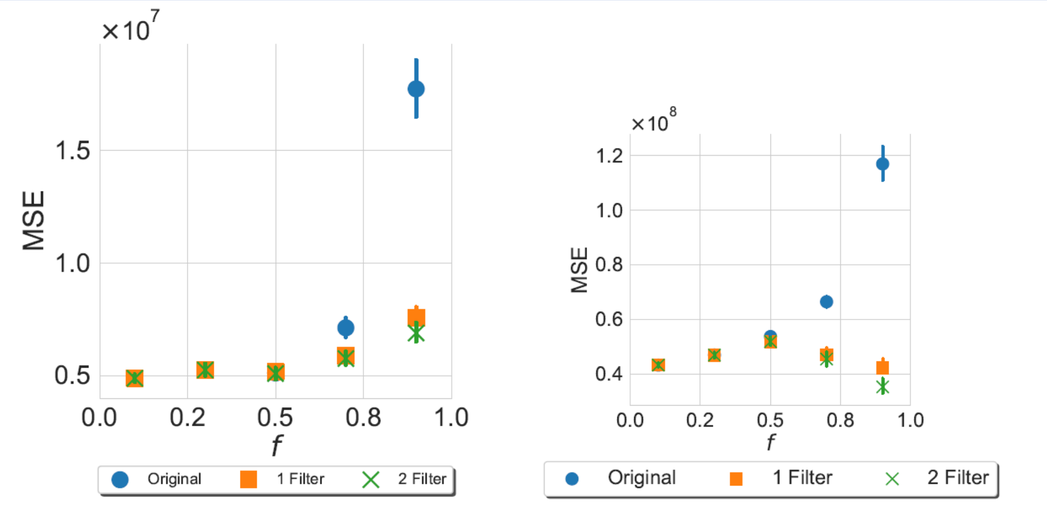


Results



(a) Zipf distribution (b) Diabetes data

Figure 1. MSE comparison of the baseline Rappor with 1 and 2 pre-filters. $f = 0.7, h = 2, k = 128, \tau = 0.8$



(a) Zipf Distribution (b) Diabetes data

Figure 2. Impact of noise probability f. $h = 2, k = 128, \tau = 0.8$

- h: Hash functions.
- k: Range of the hash functions.
- τ: Threshold to discard noisy observations.
- f: Added noise.

Discussion

The results illustrate the benefits of incorporating a model-based pre-filter in an LDP mechanism. The benefits are extracted mostly with a single pre-filter, although the addition of a second pre-filter provides further improvements, leading to a reduction of up to 31% in MSE. Also, the benefits are larger when the noise f is larger, which is related to scenarios with tighter privacy budgets.