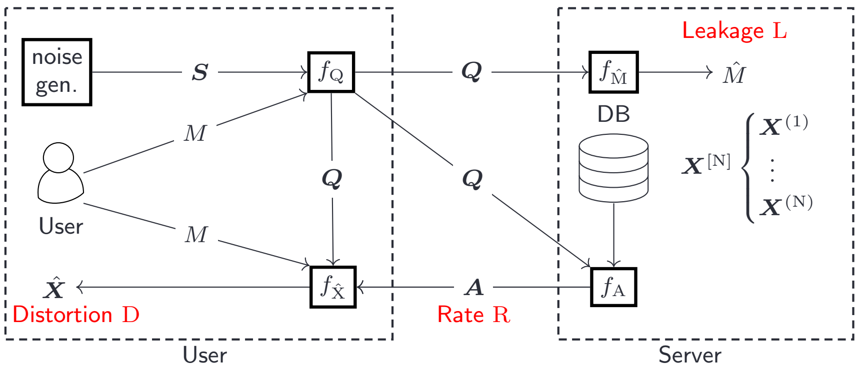
Generative Adversarial User Privacy in Lossy Single-Server Information Retrieval

Chung-Wei Weng*†, Yauhen Yakimenka*, Hsuan-Yin Lin*, Eirik Rosnes*, and Jörg Kliewer‡

*Simula UiB, Bergen, Norway †University of Bergen, Norway ‡New Jersey Institute of Technology, USA

Single-Server Information Retrieval (Lossy, Weakly-Private)

- Private information retrieval (PIR) problem: retrieve the M-th file $\boldsymbol{X}^{(M)}$ from a database $\boldsymbol{X}^{[\mathrm{N}]} = \{\boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(\mathrm{N})}\}$, but keep the index M secret



$$\min_{f_{\mathrm{Q}}} \max_{f_{\hat{\mathrm{M}}}} \quad \mathsf{E}_{M,\boldsymbol{Q}}\Big[f_{\mathrm{Loss}}(M, \hat{M})\Big],$$

$$\text{subject to:} \quad \mathsf{E}_{M,\boldsymbol{Q}}\Big[d(\boldsymbol{X}^{(M)}, \hat{\boldsymbol{X}})\Big] \le \mathrm{D}, \quad \mathrm{R}(f_{\mathrm{Q}}, f_{\mathrm{A}}) \le \mathrm{R}$$

Experiments

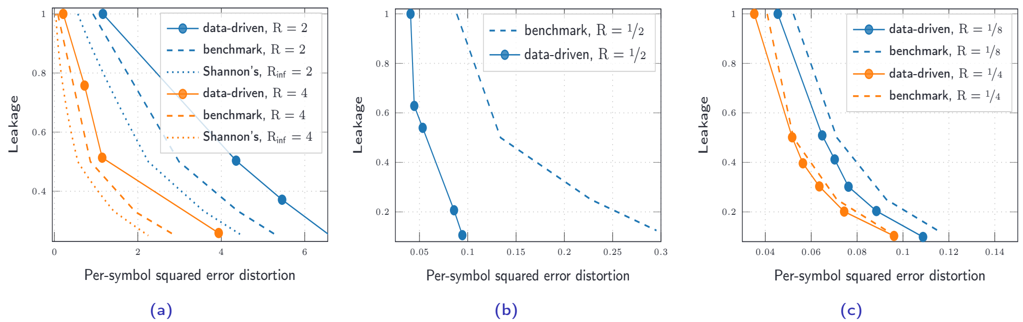


Figure: Leakage versus per-symbol squared error distortion for both the data-driven approach and the schemes from theoretical approximation. (a) Synthetic Gaussian dataset. (b) MNIST. (c) CIFAR-10

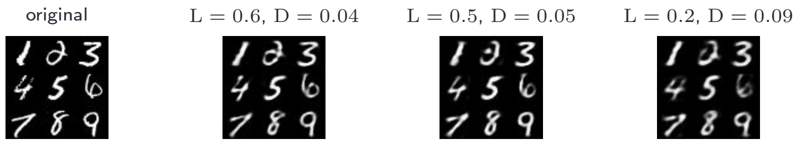
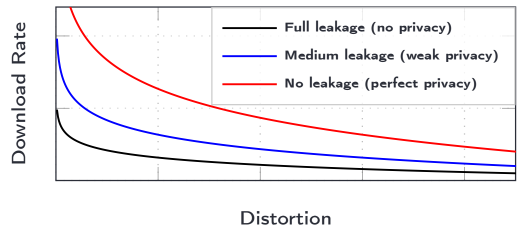


Figure: MNIST reconstruction example, $\mathrm{R} = 1/2$ bits per pixel

Motivation

- Make PIR more practical: improved download cost (or rate) by relaxing conditions:
  - relaxed perfect privacy → leaky (or weakly-private) protocol
  - relaxed perfect reconstruction → distorted reconstruction



- Limitation: unknown statistical properties of real-world datasets

Contribution

- Study the download rate, distortion, and user privacy leakage trade-off under a generative adversarial network (GAN) based approach for a single server
- Evaluate the performance for synthetic (Gaussian) and real-world (MNIST, CIFAR-10) datasets
- Compare the approaches:
  - data-driven: GAN-based
  - theoretical approximation (benchmark): lossy compression of a random subset of the files in the database
  - Shannon's asymptotic limit of rate-distortion (requires knowledge of the database's probability distribution)

Explainable Results

- Analyze (statistical) dependence of file indices used to produce an answer $\boldsymbol{A}$ for a requested file index M

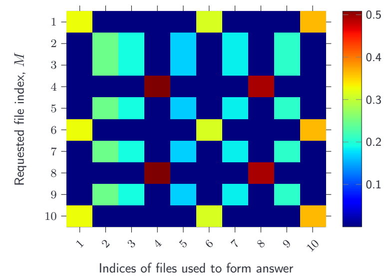


Figure: Heat map with CIFAR-10 for leakage $\mathrm{L} = 0.30$, distortion $\mathrm{D} = 0.064$, and rate $\mathrm{R} = 1/4$

- Explainable behavior: trained f_{A} splits files into subsets and form the answer $\boldsymbol{A}$ similarly for each file in the subset (in the example above, randomized queries for files $\boldsymbol{X}^{(1)}$, $\boldsymbol{X}^{(6)}$, and $\boldsymbol{X}^{(10)}$ are processed in the same manner, cf. the rows 1, 6, and 10 in the heat map)