Federated Latent Dirichlet Allocation: A Local Differential Privacy Based Framework

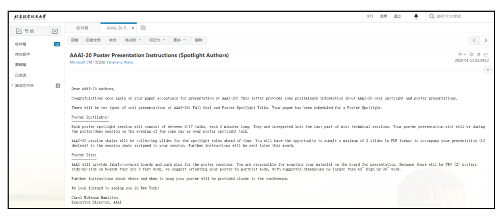
Yansheng Wang, Yongxin Tong , Dingyuan Shi

SKLSDE Lab, BDBC, School of Computer Science and Engineering and IRI, Beihang University, China

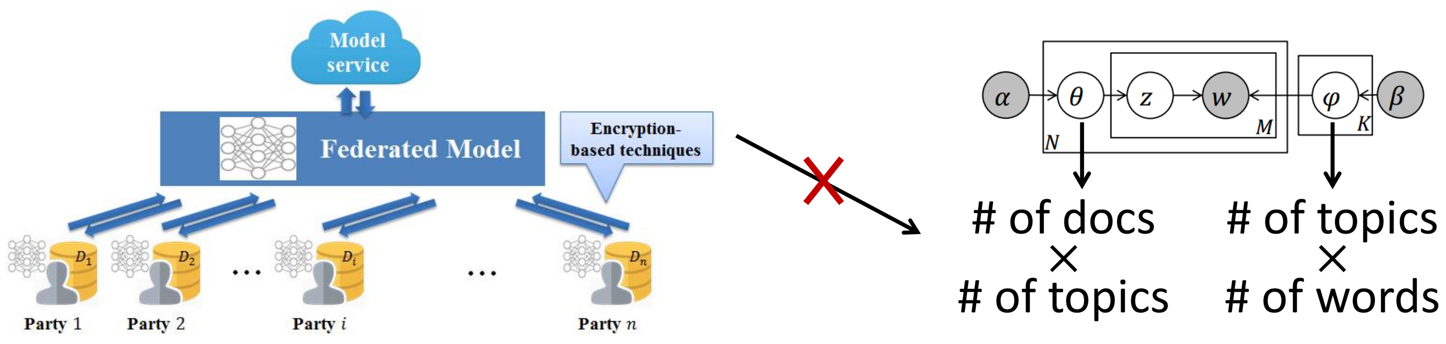
{arthur_wang, yxtong, chnsdy}@buaa.edu.cn

Introduction

- Latent Dirichlet Allocation (LDA) is often used for text mining and has been a fundamental building block for many Internet services, but privacy leak in text data is a problem.



- Federated learning (FL) can be a potential solution, but existing techniques can hardly be applied in LDA.

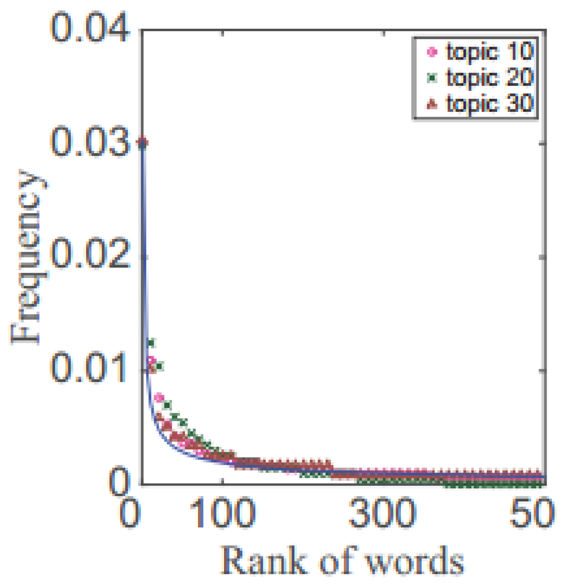


- Comparison between privacy-preserving techniques

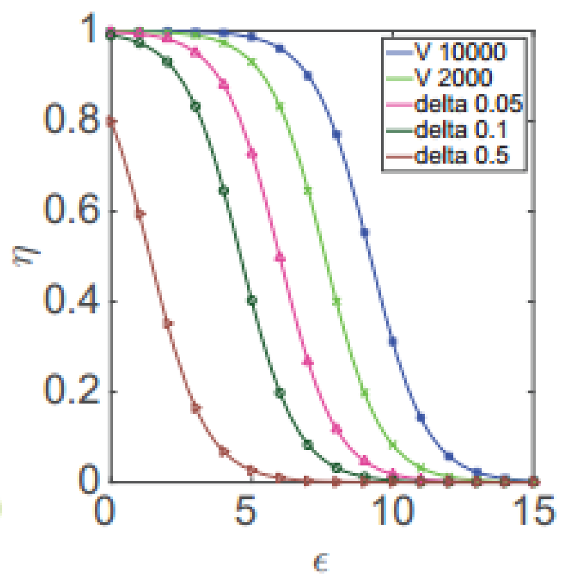
Methods	Computation cost	Communication cost	Threat model	Results
Homomorphic encryption	High	Very High	Semi-honest	Accurate
Secret sharing	High	High	Semi-honest	Accurate
Garbled Circuit	Very high	High	Semi-honest	Accurate
Local differential privacy	Low	Low	Malicious	With noise

Theoretical Analysis

- Assume frequency of words follow the *Zipf's law* (a)
- Use the failure rate δ to control the remainders with low frequency (b)



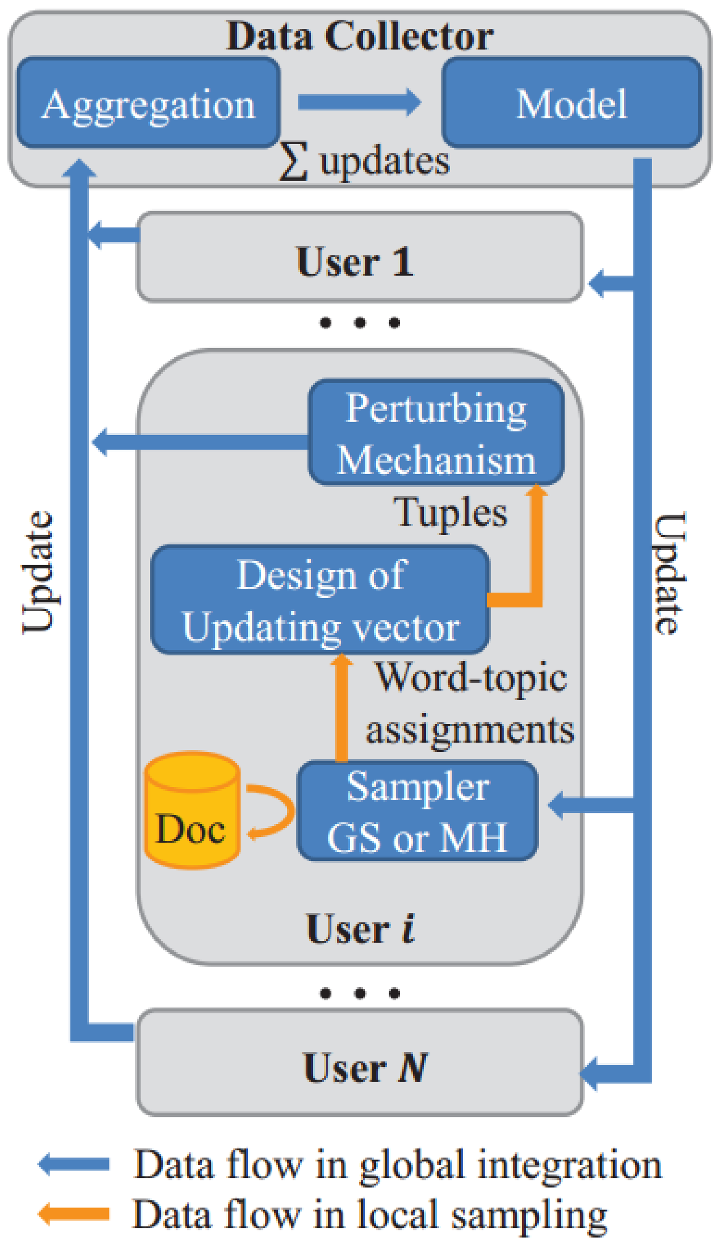
(a)



(b)

- Theorem 1 (Privacy guarantee)
  - By setting $\eta = \frac{1}{\delta\delta_0 e^{\epsilon}+1}$, the proposed mechanism satisfies $(\epsilon, 2\delta)$-***LDP***, where $\delta_0 = \delta - (\delta^{-\frac{1}{\gamma}} + 1)^{-\gamma}, \gamma \geq 1$ is a constant.
- Theorem 2 (Utility guarantee)
  - Given a fixed topic, the expected relative error of the model parameter ϕ_w after perturbation is bounded by $O(\eta k^2)$ where k is the rank of w by sorting ϕ_w in descending order.

FedLDA Overview



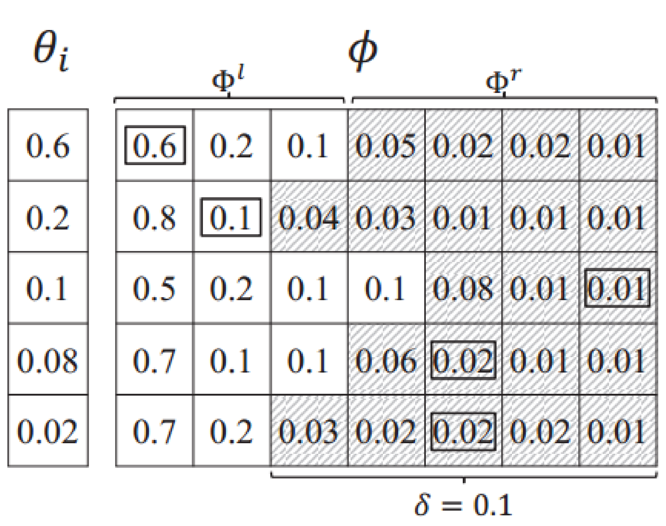
Workflow of FedLDA

System components

- Local sampling
  - Each user samples new topic-word assignment from the global φ and the local θ and submits them to the server.
- Global integration
  - The server updates the global φ while θ is updated locally

Privacy components

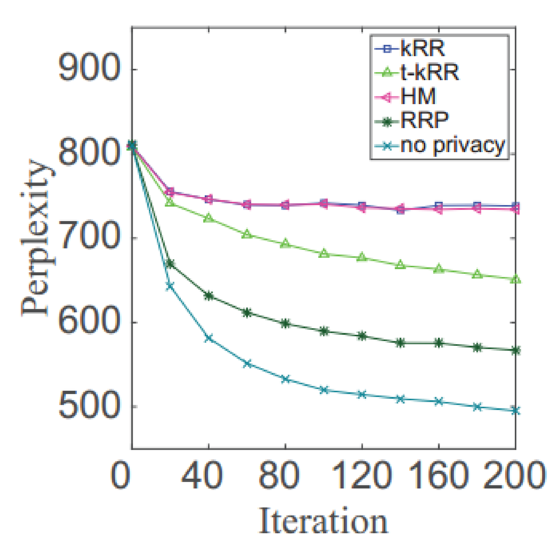
- Design of updating vector
  - Dense representation of updates.
  - Padding and sampling to reduce communication cost
- Perturbing mechanism
  - With a probability η to perturb a topic-word assignment
  - The perturbation of a word will be another word sampled from the current LDA model
  - We use a trick to exclude rare words (shaded parts)



The probability that the perturbed word is the one in the rectangle is 0.6×0.6+0.2×0.1 = 0.38

Experimental Evaluation

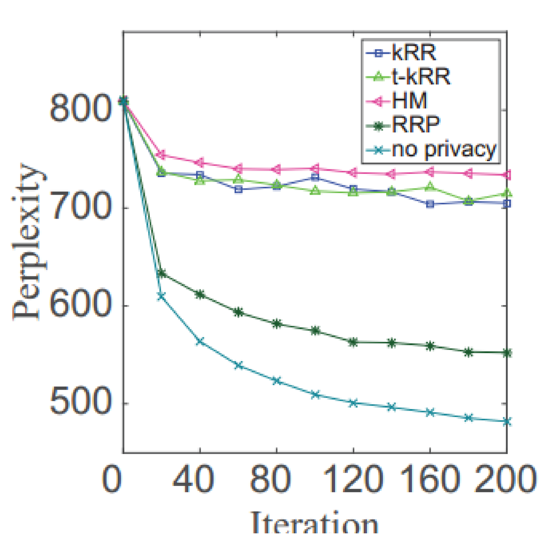
Results of GS

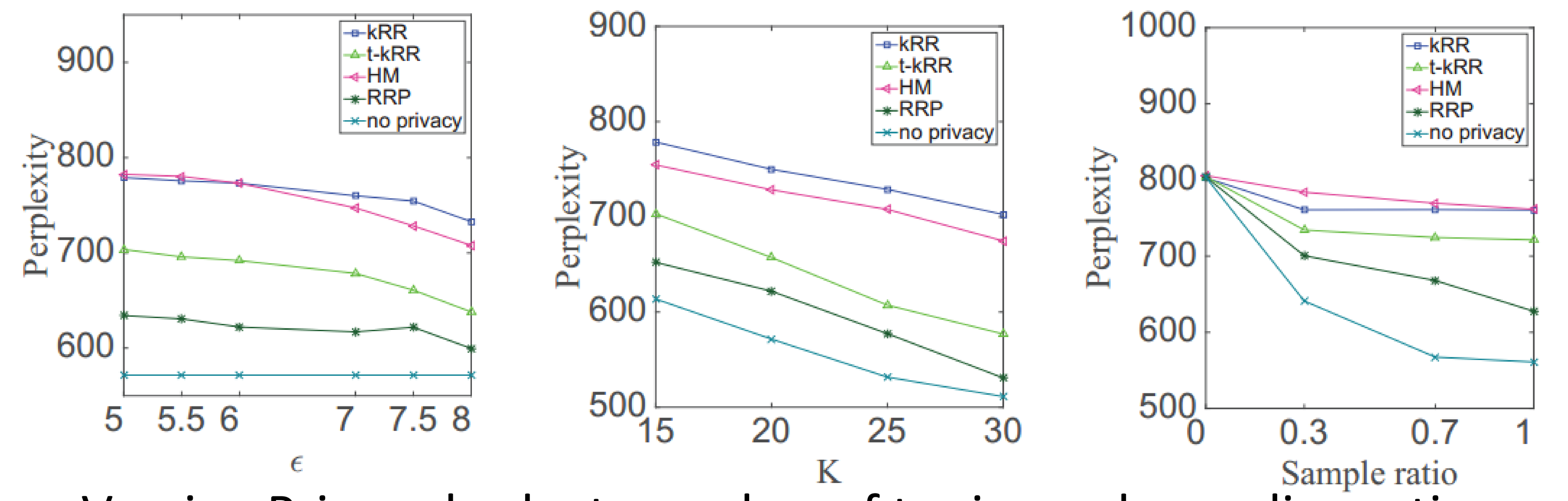


Experimental Settings
Data: Reviews and emails
\# of users: 1K, 3K, 5K
\# of topics: 20, 30, 50
Privacy budget: 5 ~ 8
Confidence: 0.5 ~ 1

Our method
No privacy

Results of MH





Varying Privacy budget, number of topics and sampling ratio

		LDA	FedLDA $_{7.5}$	FedLDA $_{5.0}$
SF	Precision	0.868	0.781	0.736
	Recall	0.708	0.767	0.760
	F1 score	0.780	0.774	0.748
	AUC score	0.798	0.771	0.738
SA	Precision	0.777	0.774	0.761
	Recall	0.814	0.776	0.766
	F1 score	0.795	0.775	0.764
	AUC score	0.794	0.778	0.767

AUC loss is less than *3%* compared with non federated model

Acknowledgment

We are grateful to anonymous reviewers for their constructive comments. This work is partially supported by the National Science Foundation of China (NSFC) under Grant No. 61822201 and U1811463. Yongxin Tong is the corresponding author of this paper.