An Empirical Study of Local Differential Privacy-based Model against Inverting Gradient Attacks on Federated Learning

Xiaotong Wu Nanjing Normal University Nanjing, China wuxiaotong@njnu.edu.cn

ABSTRACT

federated learning.

CCS CONCEPTS

 \bullet Security and privacy \to Privacy-preserving protocols; \bullet Computing methodologies \to Distributed artificial intelligence.

KEYWORDS

neural networks, federated learning, attack and defense

ACM Reference Format:

1 INTRODUCTION

federated learning [1] differential privacy inverting gradient attacks

2 EVALUATION FRAMEWORK

- 2.1 Federated Learning Algorithms
- 2.2 Inverting Gradient Attacks
- 2.3 LDP-based Models
- 2.4 Evaluation Metrics

Prediction performance

• Accuracy.

Attack metrics

•

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY © 2018 Association for Computing Machinery. ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00 https://doi.org/XXXXXXXXXXXXX

3 EXPERIMENTAL SETTING

- 3.1 Datasets
- 3.2 Neural Networks
- 4 EXPERIMENTAL RESULTS
- 5 CONCLUSION

REFERENCES

 Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. 2017. Communication-Efficient Learning of Deep Networks from Decentralized Data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, AISTATS, Vol. 54. 1273–1282.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009