Location Leakage in Federated Signal Maps

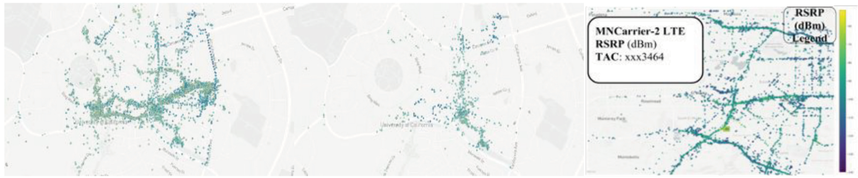
E. Bakopoulou, J. Zhang, M. Yang, J. Ley, K. Psounis, A. Markopoulou

UCI USC ProperData

SCAN ME

Motivation

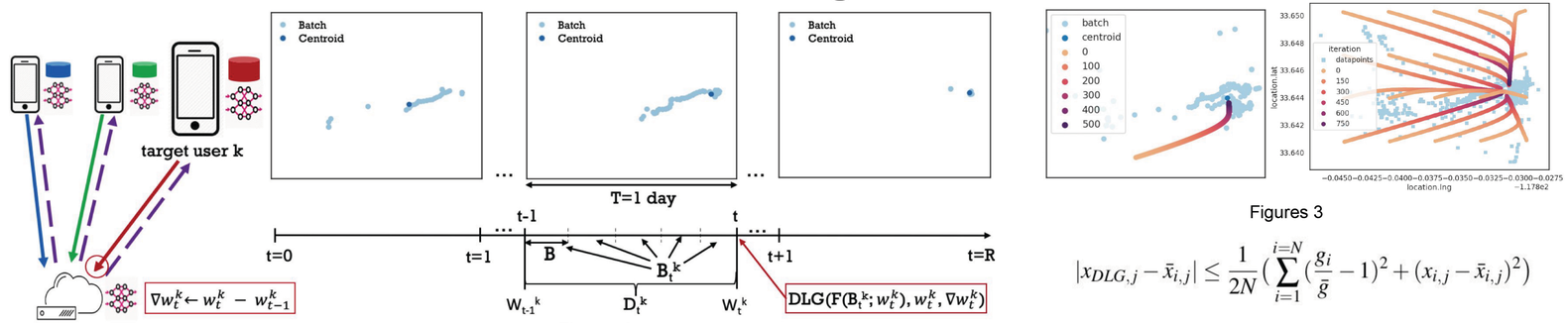
- **Background**: Mobile crowdsourcing is widely used to collect data from a large number of mobile devices. We consider the problem of predicting cellular network performance (signal maps) using online federated learning from measurements collected by several mobile devices.
- **Gradient leakage**: When applying federated learning in the signal strength predictions tasks, the model updates could cause the leakage of users' private information like locations.



Contributions

- Formulate the signal strength prediction problem within an online federated learning framework.
- We study, for the first time, a **DLG-based location privacy attack** launched by an honest-but-curious server.

Gradient Leakage Attack



Figures 2

Figures 3

$$|x_{DLG,j} - \bar{x}_{i,j}| \leq \frac{1}{2N}\left(\sum_{i=1}^{i=N}\left(\frac{g_i}{\bar{g}} - 1\right)^2 + \left(x_{i,j} - \bar{x}_{i,j}\right)^2\right)$$

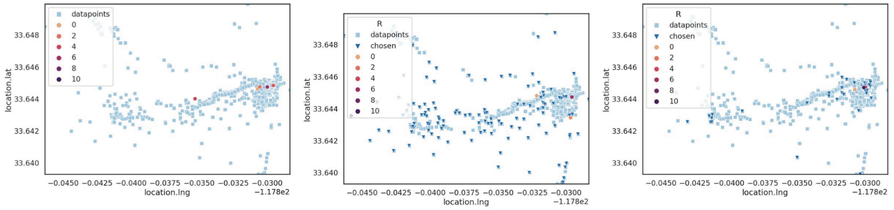
Fig 2 show the Online Federated Learning Framework with DLG Attack, using data from the campus. The target user k collects data in an online fashion, and processes them in intervals of duration T =1 day.

The server observes the model parameter update w at time t, computes the gradient and launches a DLG attack. For each day t, it manages to reconstruct the centroid (average location) of the points. During the last day, where the user did not move much, the centroid conveys quite a a lot of information.

Key observation O1: DLG on one batch: DLG converges to the average location xregardless of the initialization point.

Key observation O2: DLG across several rounds: Inferring the average location in successive rounds essentially reveals the trajectory at the granularity of interval T. And there is inherent clustering around these important locations. When successfully inferred, these can reveal sensitive information and help identify the user.

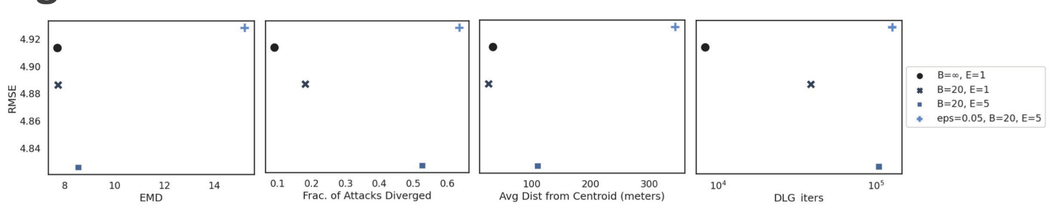
Defenses Against Attacks



(a) FedAvg: B=20, E=5: EMD=9.7, RMSE=4.83. (b) Diverse Batch: EMD=15.23, RMSE=4.93. (c) Random baseline to Diverse Batch: EMD=7.6.

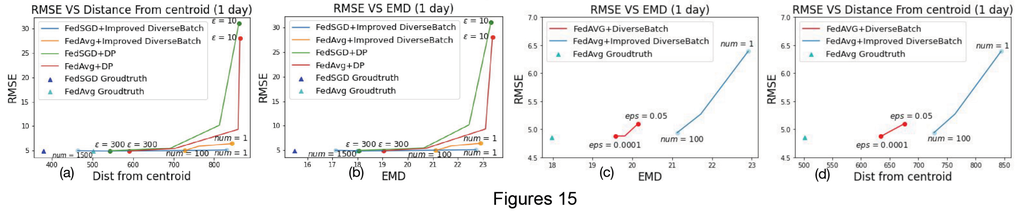
Figures 11

Fig 12 shows the privacy-utility trade-offs for all approaches. Our proposed defense mechanism FedAvg with Diverse Batch improves privacy (doubles EMD, increases divergence above 60%, and distance from 50 to 350m), without significantly hurting utility. Fig 15 (a,b) shows the comparison between differential privacy and Improved Diverse Batch for 1-day interval and (c,d) shows the comparison between Diverse Batch and Improved Diverse Batch for 1-day interval using FedAvg.



(a) EMD. (b) Diverged Attacks. (c) Distance from centroid. (d) Avg DLG iterations.

Figures 12



Figures 15

Broader Impact

Invited Talks:

- ITA Workshop, San Diego, CA, May 2022.
- Google's Federated Learning Workshop, Nov. 2021.

Future Directions

- We will continue to explore different designs and settings of DLG attacks and defenses.
- There are also some other privacy-preserving mechanisms like secure aggregation that can be applied to defense against DLG attack.