A Hybrid Decentralised Topology for Recommendations with Improved Privacy

Diarmuid O’Reilly-Morgan, Elias Tragos, James Geraci, Qinqin Wang, Neil Hurley, Barry Smyth, Aonghus Lawlor
Background and Motivation

● Recommender systems using matrix factorisation to update factors $P$ and $Q$.

● Distributed approaches promise increased privacy.
  ○ Update locally, only share $Q$ (item factors), and then aggregate on a server or with neighbours.

● However, sharing $Q$ can leak information about the user profile.

(Source - https://www.linkedin.com/pulse/fundamental-matrix-factorization-recommender-system-saurav-kumar)
Distributed Learning Topologies

- **Federated Learning (FL)**
  - Clients communicate only with a central server.

- **Anonymous Random Walks (FL-ARW)**
  - Clients communicate in sequential walks before communicating with server.
  - Small (Beta) probability of not updating the model weights.

- **Gossip-learning ARW (GL-ARW)**
  - ARW, but with no central server.
Privacy Attacks

- Distance correlation.
  - Measure mutual information between profiles and updates
    \[ dCorr(X, Y) := \frac{dCov(X, Y)}{\sqrt{dVar(X)dVar(Y)}} \]

- Profile reconstruction.
  - PCA on updates can easily reconstruct profiles from updates.

- Membership inference.
  - Linear Regression method + prior knowledge can find who contributed to an update.

Algorithm 2: Estimate rated items of client \( k \)

1. Require: Updated local item factors \( Q^k \), previous global item factors \( Q \);
2. Compute \( D = Q^k - Q \);
3. Select \( C \), the sub-matrix of non-zeros rows \( D \);
4. Compute covariance matrix \( G \) from \( C \);
5. Compute principal eigenvector \( e \) with largest \( G \) eigenvalue;
6. Return \( e^T D \): estimation of user’s rating preferences.

Figure 2: (A) PCA vectors and plotted items in a 2d matrix factorisation update. (B) Plotted representations of items in an update to which multiple users have contributed
Results

- ARW converges faster when measured in communication cost (fig. 3).
- ARW leaks less information when measured via distance correlation (fig. 4).
- ARW variants are more robust to profile reconstruction attack (fig. 6).
- ARW becomes more robust to membership inference as walk length increased (fig. 8).

Figure 3: Convergence for the various topologies on three, measuring HitRatio@10 against communication cost. The number after ARW indicates the ratio of random walks to clients.

Figure 4: Average distance correlation under different topologies (lower value is better).

Figure 6: Profile Reconstruction Attack success rate

Figure 8: Membership inference varying the walk length.