RUTGERS
School of Arts and Sciences

The ACM Conference Series on
Recommender Systems
2022

Fairness-aware Federated Matrix Factorization

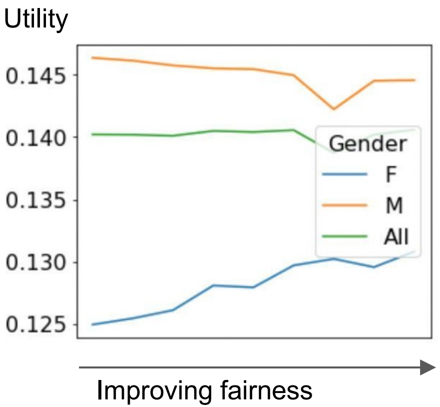
Shuchang Liu, Yingqiang Ge, Shuyuan Xu
Yongfeng Zhang, Amélie Marian

Motivation

User fairness in recommender systems.

- Should not be biased towards certain sensitive user group.
- Treatment equality by group recommendation unfairness [1]:
  - Performance(G0) = Performance(G1)

$$\mathcal{L}_{\text{fair}}(G_0, G_1, \mathcal{F}) = \left| \frac{1}{|G_0|} \sum_{u \in G_0} \mathcal{F}(u) - \frac{1}{|G_1|} \sum_{u \in G_1} \mathcal{F}(u) \right|^{\rho}$$



In reality, user group features that **require fairness control** may also be **sensitive ones that require privacy protection**.

> Gender, age, sexual orientation, …

Motivation

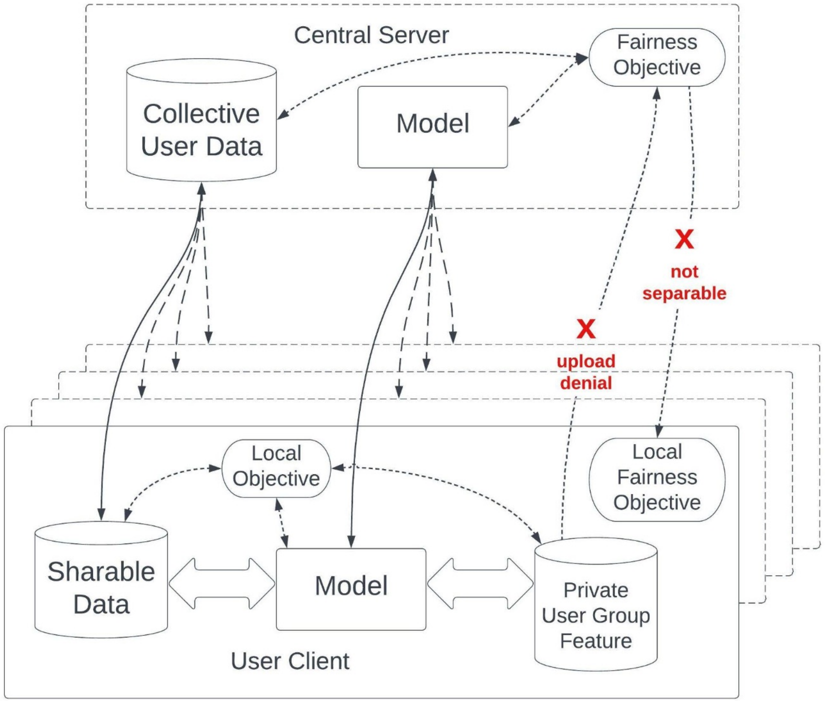
Privacy protection by federated learning:

- Leaving sensitive data on the users' devices without upload.
- Communicate model parameters and public data between user devices and central server.

In RS: federated recommender systems.

However, the fairness objective correspond to a global metric that requires the collective knowledge of user groups during optimization.

> A natural conflict in fair federated learning [2]



Related Work

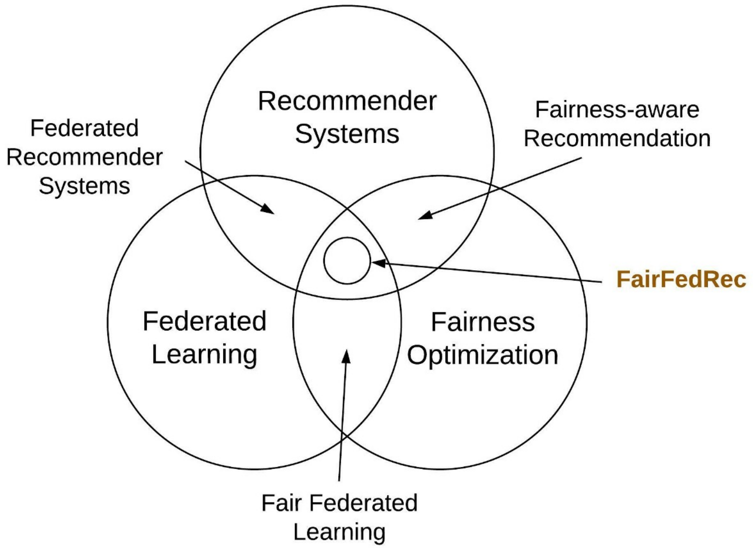
Federated recommender systems [3]

Fairness-aware recommendation [4,5]

Fair Federated Learning (FairFL) [2]:

- Several concurrent work that studied on vertical (cross-silo) federated scenarios in other machine learning tasks [6,7].

- Our goal: achieve user group fairness in horizontal FL system.



Solution

Given the overall objective $\mathcal{L} = \mathcal{L}_{\text{rec}} + \lambda \mathcal{L}_{\text{fair}}$ where the fairness objective:

$$\mathcal{L}_{\text{fair}}(G_0, G_1, \mathcal{F}) = \left| \underbrace{\boxed{\frac{1}{|G_0|} \sum_{u \in G_0} \mathcal{F}(u)}}_{A} - \underbrace{\boxed{\frac{1}{|G_1|} \sum_{u \in G_1} \mathcal{F}(u)}}_{B} \right|^{\rho}$$

Challenges:

- The fairness objective is not directly separable by users, so it does not accommodate FL.
- Utility function F(u) might be indifferentiable
  - E.g. Recall, F1, NDCG
- There is no universal metric of F(u) that also controls other metrics.

Solution

Given the overall objective $\mathcal{L} = \mathcal{L}_{\text{rec}} + \lambda \mathcal{L}_{\text{fair}}$ where the fairness objective:

$$\mathcal{L}_{\text{fair}}(G_0, G_1, \mathcal{F}) = \left| \overset{A}{\boxed{\frac{1}{|G_0|} \sum_{u \in G_0} \mathcal{F}(u)}} - \overset{B}{\boxed{\frac{1}{|G_1|} \sum_{u \in G_1} \mathcal{F}(u)}} \right|^{\rho}$$

Assume $\mathcal{F}(u) = -\mathcal{L}_{\text{rec}}^{(u)}$, then each user's local gradient becomes:

$$\nabla \Theta_u = D \frac{\partial}{\partial \Theta_u} \mathcal{L}_{\text{rec}}^{(u)}, \text{ where } D = 1 - \lambda C \, |A - B|^{\rho - 1}$$

$$C = \rho(-1)^{\mathbb{1}(A<B)}(-1)^{\mathbb{1}(u \notin G_0)}$$

Solution

Given the overall objective $\mathcal{L} = \mathcal{L}_{\text{rec}} + \lambda \mathcal{L}_{\text{fair}}$ where the fairness objective:

$$\mathcal{L}_{\text{fair}}(G_0, G_1, \mathcal{F}) = \left| \underbrace{\boxed{\frac{1}{|G_0|} \sum_{u \in G_0} \mathcal{F}(u)}}_{A} - \underbrace{\boxed{\frac{1}{|G_1|} \sum_{u \in G_1} \mathcal{F}(u)}}_{B} \right|^{\rho}$$

Assume $\mathcal{F}(u) = -\mathcal{L}_{\text{rec}}^{(u)}$, then each user's local gradient becomes:

$$\nabla \Theta_u = D \frac{\partial}{\partial \Theta_u} \mathcal{L}_{\text{rec}}^{(u)}, \text{ where } D = 1 - \lambda C \, |A - B|^{\rho - 1}$$

$$C = \rho(-1)^{\mathbb{1}(A<B)} (-1)^{\mathbb{1}(u \notin G_0)}$$

Intuitive explanation:

- C > 0 ⇒ D < 1: slow down training if user belongs to the advantage group.
- C < 0 ⇒ D > 1: speed up training if user belongs to the disadvantage group.

Solution

Still need to synchronize these aggregated statistics of F(u) and group membership counts |G|.

Given the overall objective $\mathcal{L} = \mathcal{L}_{\text{rec}} + \lambda \mathcal{L}_{\text{fair}}$ where the fairness objective:

$$\mathcal{L}_{\text{fair}}(G_0, G_1, \mathcal{F}) = \left| \underbrace{\boxed{\frac{1}{|G_0|} \sum_{u \in G_0} \mathcal{F}(u)}}_{A} - \underbrace{\boxed{\frac{1}{|G_1|} \sum_{u \in G_1} \mathcal{F}(u)}}_{B} \right|^{\rho}$$

Assume $\mathcal{F}(u) = -\mathcal{L}_{\text{rec}}^{(u)}$, then each user's local gradient becomes:

$$\nabla \Theta_u = D \frac{\partial}{\partial \Theta_u} \mathcal{L}_{\text{rec}}^{(u)}, \text{ where } D = 1 - \lambda C \, |A - B|^{\rho - 1}$$

$$C = \rho(-1)^{\mathbb{1}(A<B)} (-1)^{\mathbb{1}(u \notin G_0)}$$

Intuitive explanation:

- C > 0 ⇒ D < 1: slow down training if user belongs to the advantage group.
- C < 0 ⇒ D > 1: speed up training if user belongs to the disadvantage group.

Solution

The fairness objective only needs the correct aggregated group information instead of the group label of each individual user:

$$\mathcal{L}_{\text{fair}}(G_0, G_1, \mathcal{F}) = \left| \boxed{\frac{1}{|G_0|} \sum_{u \in G_0} \mathcal{F}(u)} - \boxed{\frac{1}{|G_1|} \sum_{u \in G_1} \mathcal{F}(u)} \right|^{\rho}$$

This opens up the choice of differential privacy:

- Disguise each user's label while keeping the aggregated info accurate.

Solution

The fairness objective only needs the correct aggregated group information instead of the group label of each individual user:

$$\mathcal{L}_{\text{fair}}(G_0, G_1, \mathcal{F}) = \left| \frac{1}{|G_0|} \sum_{u \in G_0} \mathcal{F}(u) - \frac{1}{|G_1|} \sum_{u \in G_1} \mathcal{F}(u) \right|^{\rho}$$

This opens up the choice of differential privacy:

- Disguise each user's label while keeping the aggregated info accurate.

Challenges:

- F(u) changes across epochs, so adding a single noise may still expose the user's group feature.

> Solution: user-wise noise + epoch-wise noise

Solution

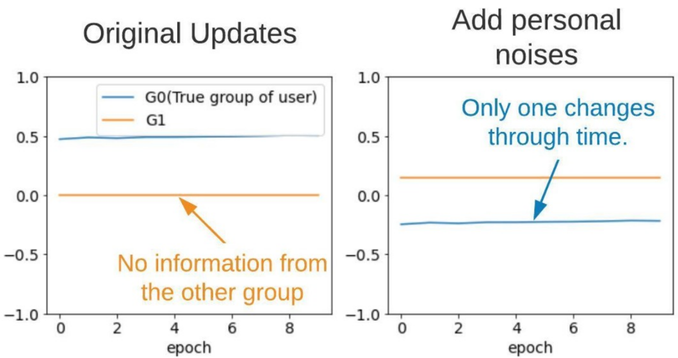
Users still need to upload F(u) and which group they belong to, but with disguise:

- Option 1: Random noise.
  - Outsiders can figure F(u) with continuous observation since $\Pr(|\lim_{N\to\infty}\bar{\epsilon} - \mathbb{E}[\epsilon]| < \delta) = 1$

Solution

Users still need to upload F(u) and which group they belong to, but with disguise:

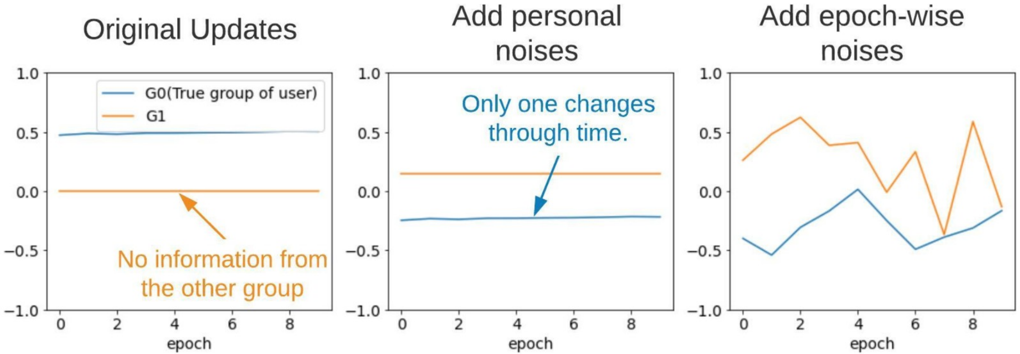
- Option 1: Random noise.
  - Outsiders can figure F(u) with continuous observation since $\Pr(|\lim_{N\to\infty} \bar{\epsilon} - \mathbb{E}[\epsilon]| < \delta) = 1$
- Option 2: User-wise noise.
  - Random noise across users, but fixed after intialization.
  - Information of only one group changes through time, and the group membership is exposed.



Solution

Users still need to upload F(u) and which group they belong to, but with disguise:

- Option 1: Random noise.
  - Outsiders can figure F(u) with continuous observation since $\Pr(|\lim_{N\to\infty} \bar{\epsilon} - \mathbb{E}[\epsilon]| < \delta) = 1$
- Option 2: User-wise noise.
  - Random noise across users, but fixed after intialization.
  - Information of only one group changes through time, and the group membership is exposed.
- Option 3 ✔: User-wise noise + epoch-wise random noise



Information to upload:

$$\nabla A_{\text{sum}}|u = \mathbb{1}(u \in G_0)\mathcal{F}_u + \epsilon_{1,u} + \epsilon_{A,t}$$

$$\nabla B_{\text{sum}}|u = \mathbb{1}(u \in G_1)\mathcal{F}_u + \epsilon_{2,u} + \epsilon_{B,t}$$

$$\nabla A_{\text{count}}|u = \mathbb{1}(u \in G_0) + \epsilon_{3,u}$$

$$\nabla B_{\text{count}}|u = \mathbb{1}(u \in G_1) + \epsilon_{4,u}$$

Solution

Users still need to upload F(u) and which group they belong to, but with disguise:

- Option 1: Random noi
  - Outsiders can figure F(
- Option 2: User-wise n
  - Random noise across
  - Information of only one
- Option 3 ✓: User-wise noise + epoch-wise random noise

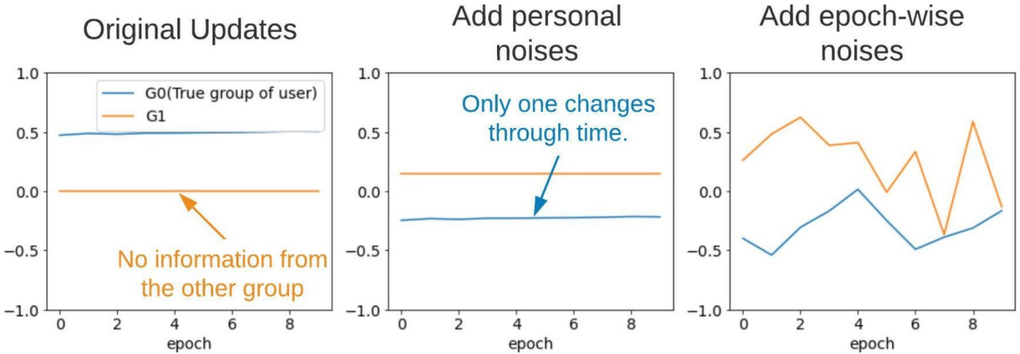
Central server aggregation:

$$\text{Update } \Theta^{(t)} \longleftarrow \text{Aggregation}(\nabla\Theta|u, \forall u \in \mathcal{U}_{\text{subset}}).$$

$$A^{(t)} \leftarrow \frac{\sum_{u\in\mathcal{U}_{\text{subset}}} \nabla A_{\text{sum}}|u}{\sum_{u\in\mathcal{U}_{\text{subset}}} \nabla A_{\text{count}}|u}$$

$$B^{(t)} \leftarrow \frac{\sum_{u\in\mathcal{U}_{\text{subset}}} \nabla B_{\text{sum}}|u}{\sum_{u\in\mathcal{U}_{\text{subset}}} \nabla B_{\text{count}}|u}$$

Aggregated A and B will be used to determine the scalar D in local optimization. The communication overhead is O(NK).



Information to upload:

$$\nabla A_{\text{sum}}|u = \mathbb{1}(u \in G_0)\mathcal{F}_u + \epsilon_{1,u} + \epsilon_{A,t}$$

$$\nabla B_{\text{sum}}|u = \mathbb{1}(u \in G_1)\mathcal{F}_u + \epsilon_{2,u} + \epsilon_{B,t}$$

$$\nabla A_{\text{count}}|u = \mathbb{1}(u \in G_0) + \epsilon_{3,u}$$

$$\nabla B_{\text{count}}|u = \mathbb{1}(u \in G_1) + \epsilon_{4,u}$$

Experiments

Model: Matrix factorization as base recommendation model.

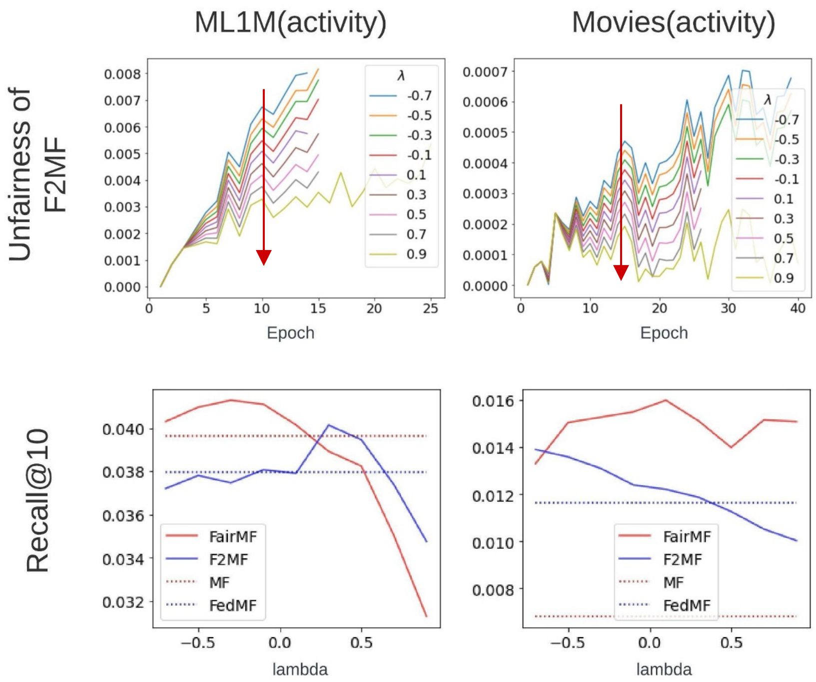
Shared information: interacted items.

User group information:

- Totally private (F2MF): gender, age (5 group).
- Partially private (F3MF): activity level.
    - Noise ← 0

Dataset (80-10-10):

Dataset	$\lvert\mathcal{U}\rvert$	$\lvert\mathcal{I}\rvert$	#record	sparsity	user feature	#group
ML-1M	6,022	3,043	995,154	0.9457	gender	2
					activity	2
					age	5
Movies	5,515	13,509	484,141	0.9935	activity	2



FairMF: Centralized counterpart of F2MF

Experiments

Threshold for effective fairness control:

Increase lambda

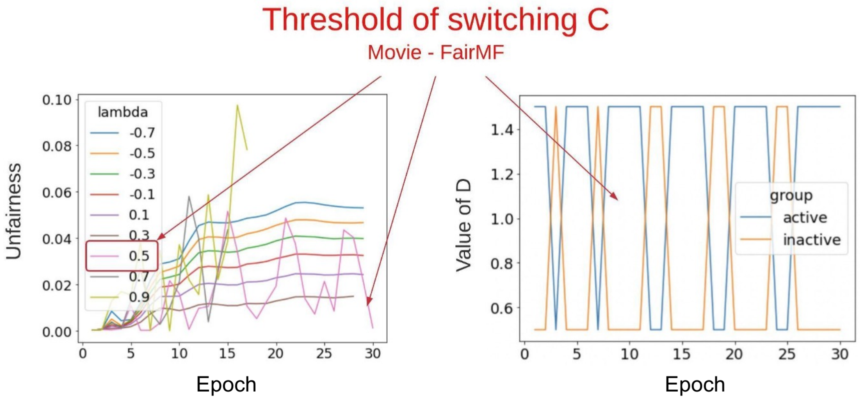
→ Smaller group difference

→ Higher chance observing switching C (i.e. advantage group ←→ disadvantage group)

→ Unstable fairness control

Note:

Stable fairness control below the threshold.



Experiments

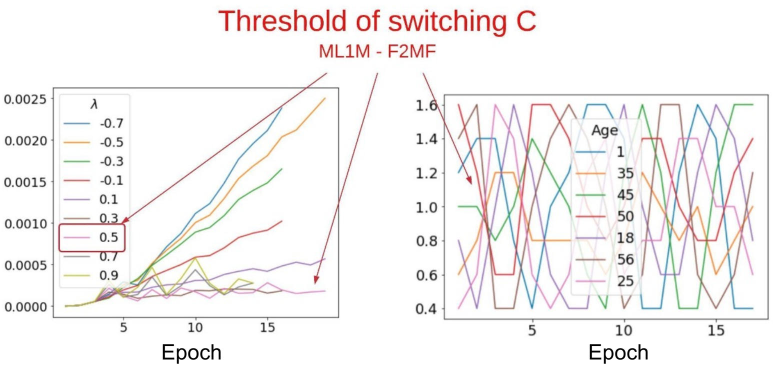
Threshold for effective fairness control:

Increase lambda or increase number of group

→ Smaller group difference

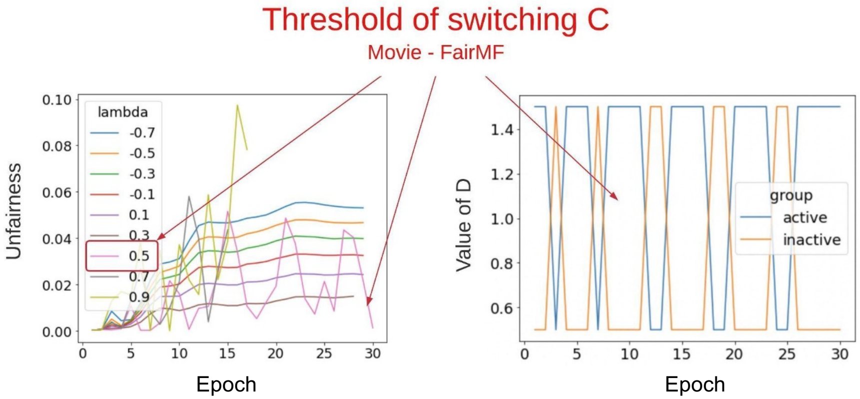
→ Higher chance observing switching C (i.e. advantage group ←→ disadvantage group)

→ Unstable fairness control



Note:

Stable fairness control below the threshold.

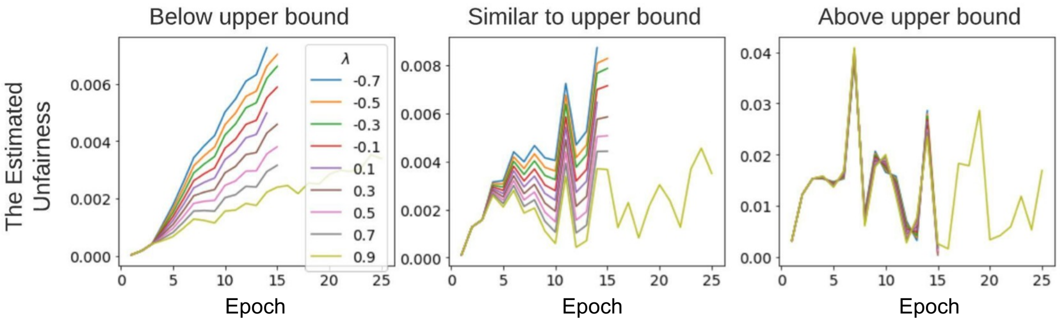


Experiments

Adequate noise magnitude for F2MF:

- The noise should be large enough to disguise ground truth information.
- The aggregated noise should be small enough to maintain accurate estimation of unfairness.

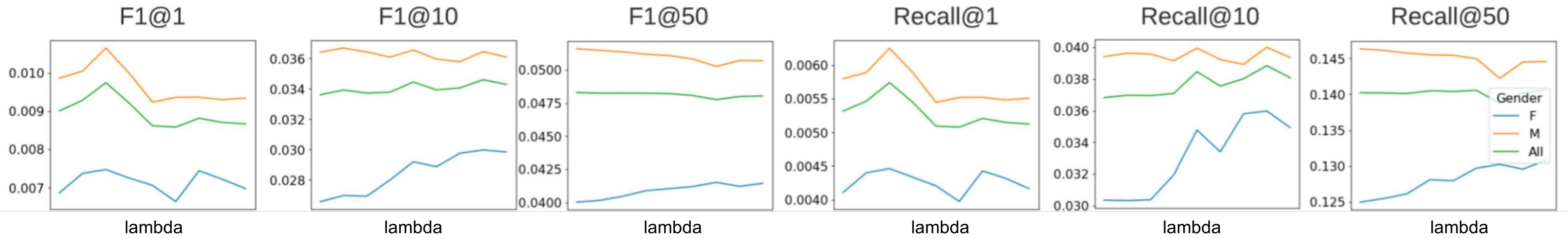
$$\sigma \leq H|\bar{X}_{\text{actual}}|\sqrt{N\delta_2}$$



Experiments

Correlation between metrics in unfairness evaluation:

There are cases when different metrics are consistent:



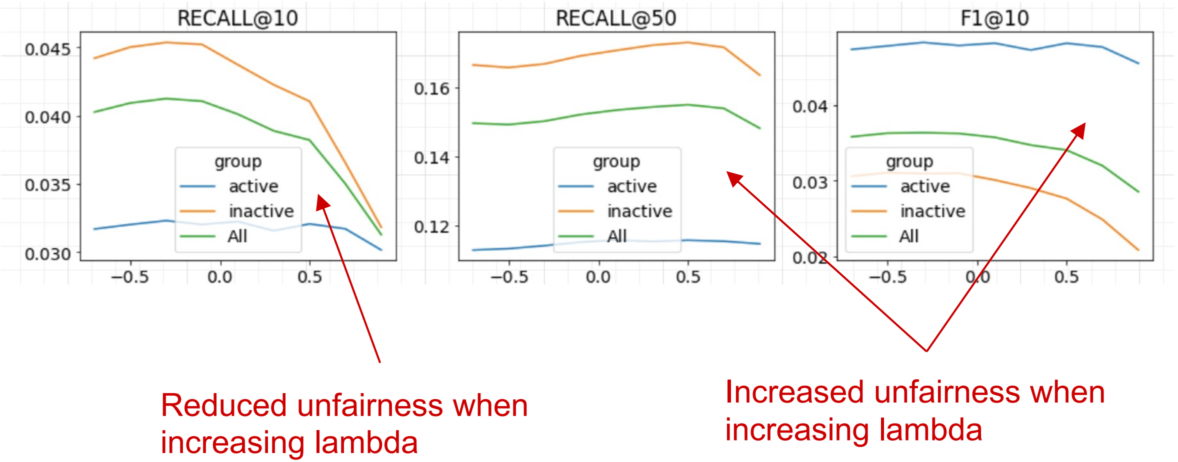
Improving fairness on one metric does not mean improving fairness on another.

Experiments

Correlation between metrics in unfairness evaluation:

There are cases when different metrics are consistent.

There are also cases where metrics are inconsistent, and improving fairness on one metric does not induce improving fairness on another.

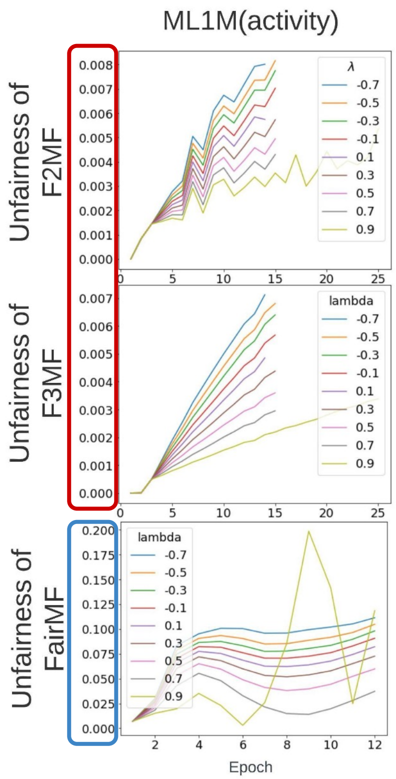


Experiments

Horizontal federated learning may systematically improves user fairness:

The estimated unfairness of federated solutions (F2MF and F3MF) are significantly smaller than their centralized counterpart (FairMF).

There are similar observations in other fair FL task [3].



Summary

- Goal: engage user group fairness control in horizontal federated recommender systems.
- F2MF solution framework:
  - Effective control through loss-based unfairness metric.
  - Little communication overhead from differential privacy module.
  - Works for both partially private and totally private scenarios.
- Some insights:
  - FL with FedAvg may naturally improves fairness.
  - Performance-based fairness may behave differently according to the chosen metric.

Implementation: https://github.com/CharlieMat/FedFairRec.git

Thanks!

References

[1] Chen, Le, et al. "Investigating the impact of gender on rank in resume search engines." *Proceedings of the 2018 chi conference on human factors in computing systems*. 2018.

[2] Zhou, Zirui, et al. "Towards fair federated learning." *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 2021. Tutorial: https://www.cas.mcmaster.ca/~chul9/Contents/KDD_2021_Tutorial/Towards_FairFL_Final.pdf

[3] Yang, Liu, et al. "Federated recommendation systems." *Federated Learning*. Springer, Cham, 2020. 225-239.

[4] Wang, Yifan, et al. "A Survey on the Fairness of Recommender Systems." *ACM Journal of the ACM (JACM)* (2022).

[5] Li, Yunqi, Yingqiang Ge, and Yongfeng Zhang. "Tutorial on fairness of machine learning in recommender systems." *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2021.

[6] Chu, Lingyang, et al. "Fedfair: Training fair models in cross-silo federated learning." *arXiv preprint arXiv:2109.05662* (2021).

[7] Ezzeldin, Yahya H., et al. "Fairfed: Enabling group fairness in federated learning." *arXiv preprint arXiv:2110.00857* (2021).

[8] Zeng, Yuchen, Hongxu Chen, and Kangwook Lee. "Improving fairness via federated learning." *arXiv preprint arXiv:2110.15545* (2021).