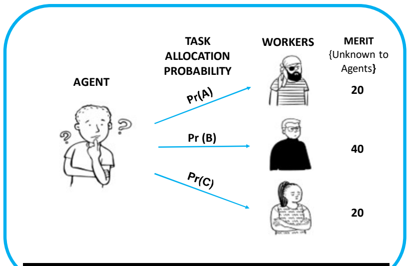
Fairness and Privacy Guarantees in Federated Contextual Bandits

Sambhav Solanki [1], Sujit Gujar [1], and Shweta Jain [2]

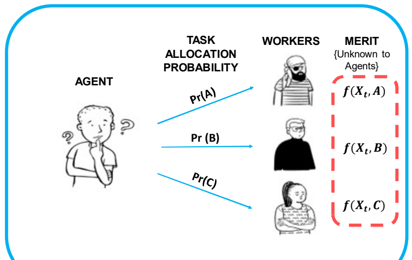
[1] IIIT, Hyderabad [2] IIT, Ropar

sambhav.solanki@research.iiit.ac.in sujit.gujar@iiit.ac.in shwetajain@iitrpr.ac.in

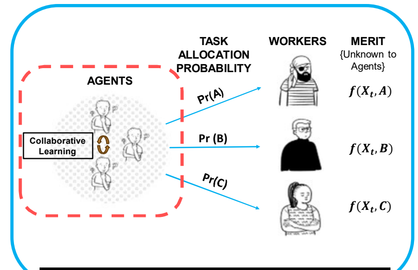
From 'Bandits' to 'Fair Federated Contextual Bandits'



Bandit Problem : Pick the worker with most merit every round

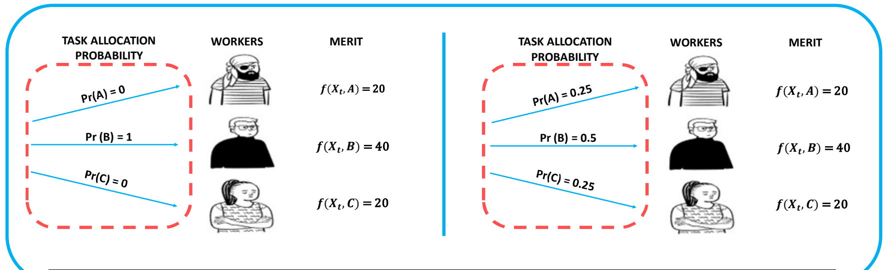


Contextual Bandits : Merit is dependent on the context, X_t, observed (ex. Task category)



Federated Learning : Agents working with same/similar workers can learn faster

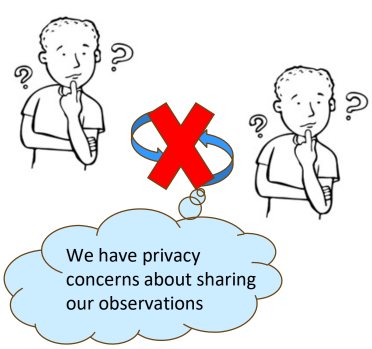
Fairness of exposure: The selection probability ratio of two arms should be equal to merit ratio of the two arms (here workers).



'To be greedy and loose workers to starvation' or 'To be fair and allocate proportional to merit'

Goal: To minimizes fairness regret for federated contextual bandits. Fairness regret is given by the difference norm between chosen and optimal (if merit wasn't unknown) selection probability vectors.

Differential Privacy



Given that to engage in federated learning, agents have to share their observations (both context and reward), privacy concerns are expected to arise

Privacy guarantees should be provided. Specifically, we ensure that our algorithm can provide differential privacy.

Theoretical Results

Fed-FairX-LinUCB (Non-private communication)

Idea: Use a novel communication protocol to periodically share observations. Find the optimal selection vector by constructing a confidence region around current estimate using shared gram matrices and reward vectors.

Theorem 1 (simplified) : With high probability, Fed-FairX—LinUCB achieves following fairness regret if the context norm is bounded by 1,

$$O\left(\frac{\sqrt{\beta_T}}{\gamma}\sqrt{mTd\log\left(1+\frac{T}{d}\right)+\ m^2d^3log^3\left(1+\frac{T}{d}\right)}\right)$$

Priv-FairX-LinUCB (Private communication)

Idea: Use privatizer routine (tree-based mechanism), to calculate noisy gram matrix and reward vectors, before sharing observations.

Theorem 2 (simplified) : With high probability, Priv-FairX—LinUCB achieves following fairness regret,

$$O\left(\frac{\sqrt{\beta_T}}{\gamma}\sqrt{mTd\log\left(\frac{\overline{\rho}}{\underline{\rho}}+\frac{T}{d\underline{\rho}}\right)+\ m^2d^3log^3\left(\frac{\overline{\rho}}{\underline{\rho}}+\frac{T}{d\underline{\rho}}\right)}\right)$$

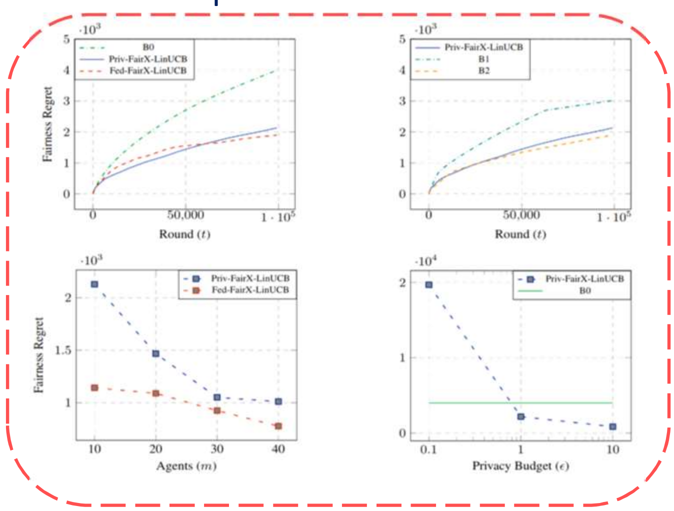
Communication Protocol

Phase 1
Gap doubles between communication rounds every time

Phase 2
Periodic communication with fixed bounded gaps

Bounded number of communication rounds and bounded gaps between any two communication rounds

Experimental Results



B0 – Single Agent Learning
B1 – Priv-FairX-LinUCB with Dubey et al 's [1] communication protocol
B2 – Priv-FairX-LinUCB with Solanki et al 's [2] communication protocol

Checkout the full paper

Download this poster

[1] -- Dubey, Abhimanyu, and AlexSandy Pentland. "Differentially-private federated linear bandits." Advances in Neural Information Processing Systems 33 (2020): 6003-6014.
[2] -- Solanki, Sambhav, et al. "Differentially private federated combinatorial bandits with constraints." Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham: Springer Nature Switzerland, 2022.