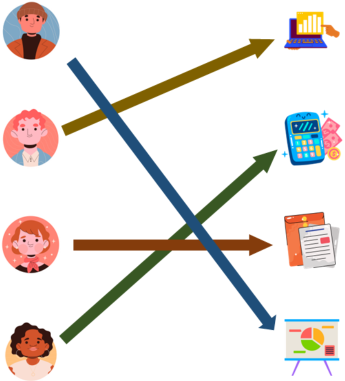
Stable Matching with Ties: Approximation Ratios and Learning

Shiyun Lin[1], Simon Mauras[2], Nadav Merlis[3], Vianney Perchet[4]

1. School of Mathematical Sciences, Peking University
2. INRIA, FairPlay Joint Team
3. Technion - Israel Institute of Technology
4. CREST, ENSAE, Criteo AI Lab, FairPlay Joint Team

November 6, 2025

Stable Matching (Matching Markets)



- Two sides of the market: Workers & Jobs
- Each participant has a preference ordering over the other side.
- Jobs over workers: $w \succ_a w' \Rightarrow$ worker w performs job a strictly better than w'.
- Workers over jobs: utility matrix $\boldsymbol{U}$ where $\boldsymbol{U}(w, a) > \boldsymbol{U}(w, a')$ indicates worker w prefers job a over a'.
- **Stable Matching:** No worker and job form a blocking pair, i.e., there is no worker-job pair who would both be better off by matching with each other instead of their current partners.
- **Worker-optimal stable matching**: a single stable matching that is utility-maximizing for all workers.

- In online marketplaces, the preferences of workers over jobs (the utility matrix $\boldsymbol{U}$) may be uncertain but could be learned through the iterative matching process.
- Multi-player Multi-armed Bandit problem: workers (players) and jobs (arms) interact over T rounds. Every round, we implement a matching μ and noisy rewards of the matched pairs $(w, \mu(w))$ could be observed.
- Worker-optimal Stable Regret:

$$Reg_i(T) = T \cdot \boldsymbol{U}^*(w_i) - \mathbb{E}\left[\sum_{t=1}^{T} X_i(t)\right].$$

Existing Work on Tie-free Markets

- Prior work on minimizing worker-optimal stable regret focus exclusively on tie-free markets.
- The results are inapplicable when preferences contain ties since existing regret bounds scale as $1/\Delta^2$, where Δ is the minimum utility gap across all workers and jobs.
- **Fundamental reason**: There is not a single worker-optimal stable matching in a matching market with indifferent preferences, i.e., $\boldsymbol{U}^*(w_i)$ is achieved in different stable matchings for different worker w_i.

> To balance competing interests of different workers, we should consider approximation.

Back to the Offline Problem...

- **Optimal Stable Share (OSS)**: The maximum utility a worker could receive among all possible stable matchings, i.e., $\boldsymbol{U}^*(w) = \boldsymbol{U}(w, \mu^*(w)) = \max_{\mu \in \mathcal{S}} \boldsymbol{U}(w, \mu(w))$, where $\mathcal{S}$ denotes the class of stable matchings.
- **OSS-Ratio**:

$$R_{\mathcal{C}} := \min_{D \in \Delta(\mathcal{C})} \max_{w \in \mathcal{W}} \frac{\boldsymbol{U}^*(w)}{\boldsymbol{U}_D(w)}.$$

- The OSS-ratio quantifies the satisfaction level of any worker w in the market when we consider a distribution D over a set of matchings $\mathcal{C}$.
- The OSS-ratio ensures every worker receives a fair share of their optimal stable utility.
- If $\max_{\boldsymbol{U}} R_{\mathcal{C}} \leq \alpha$, every worker is guaranteed at least $(1/\alpha)\boldsymbol{U}^*(w_i)$ in expectation.

Upper and Lower Bounds for the OSS-ratio

- We have the following definition

$$\mathcal{S} := \{\mu : \mu \text{ is a stable matching}\}, \mathcal{I} := \{\mu : \mu \text{ is an internally stable matching}\},$$
$$\mathcal{M} := \{\mu : \mu \text{ is a matching}\}$$

- A first observation: $\mathcal{S} \subset \mathcal{I} \subset \mathcal{M} \Rightarrow R_{\mathcal{M}} \leq R_{\mathcal{I}} \leq R_{\mathcal{S}}$.

Theorem (Lower bound on the set of stable matchings)

There exists an instance such that $R_{\mathcal{S}} = \Omega(N)$. (This is tight since $R_{\mathcal{S}} = \mathcal{O}(N)$.)

Theorem (Lower bound on the set of matchings)

There exists an instance such that $R_{\mathcal{M}} = \Omega(\log N)$.

Theorem (Upper bound on the set of internally stable matchings)

For every instance, it holds that $R_{\mathcal{I}} = \mathcal{O}(\log N)$. (We construct an approximation oracle.)

Bandit Learning Regret Bounds

- Building on the offline results, we introduce α-approximation stable regret:

$$Reg_i^{\alpha}(T) := \alpha T \cdot \boldsymbol{U}^*(w_i) - \mathbb{E}\left[\sum_{t=1}^{T} X_i(t)\right].$$

- We propose an adaptive algorithm *ETCO* handles both strict and tied preferences.

Theorem (Regret upper bound)

The ETCO algorithm with exploration phase of length T_0 ensures that, for every worker $w_i \in \mathcal{W}$,

$$Reg_i(T) = \mathcal{O}\left(\frac{K \ln T}{\Delta_{\min}^2}\right), \text{ if } \Delta_{\min} = \Omega\left(\sqrt{\frac{K \ln T}{T_0}}\right), \text{(Optimal)}$$

$$Reg_i^{\alpha}(T) = \mathcal{O}\left(T_0 + T\sqrt{\frac{K \ln T}{T_0}}\right), \text{ otherwise.}$$

Bandit Learning Regret Bounds, Cont'd

- There is a fundamental trade-off between the market with and without ties: no algorithm can simultaneously achieve optimal regret in both large-gap (standard regret) and small / no-gap (approximation regret) regimes.

Theorem (Regret lower bound)

Assume that an algorithm guarantees sublinear regret for all workers in all instances with $\Delta_{rel} = \Omega(T^{-1/2+\delta})$, then there exists an instance such that this algorithm suffers $\Omega(T^{1-2\delta})$ approximation regret for some worker when $\Delta_{rel} = 0$.

Thanks for listening!