

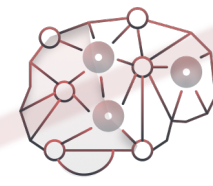


POLITECNICO
MILANO 1863

VARIANCE-AWARE OPTIMAL RANKING IN LOG-CONCAVE RANDOM UTILITY MODELS

DIEGO ALOVISETTI, MARCO MUSSI, ALBERTO MARIA METELLI

{diego.alovisetti, marco.mussi, albertomaria.metelli}@polimi.it

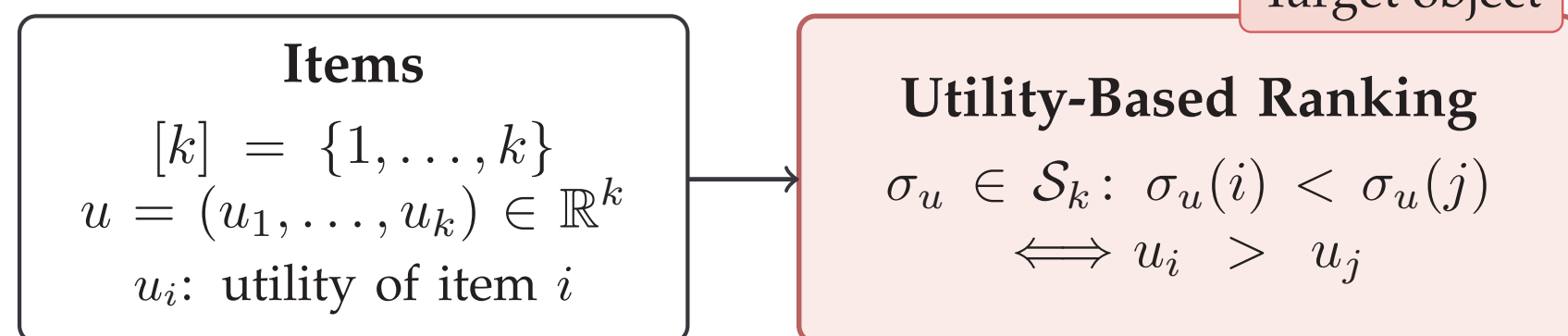


RL³

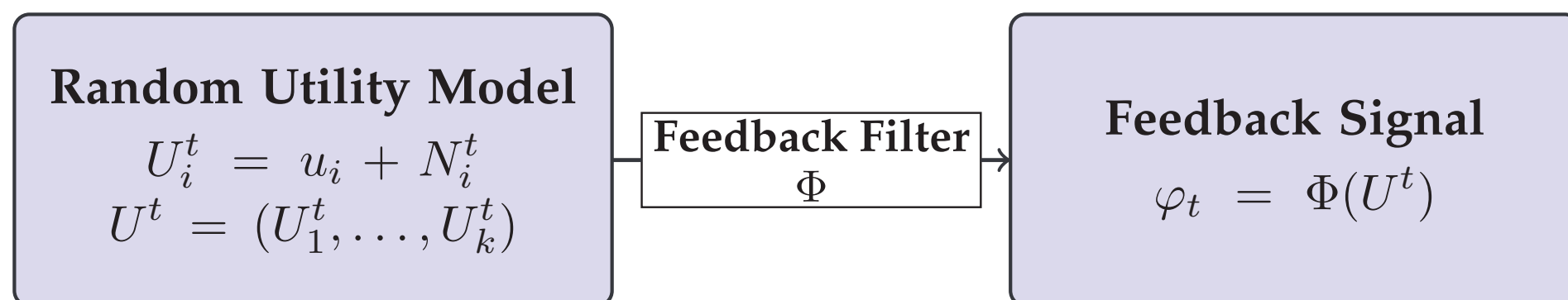


THE LEARNING PROBLEM (SEQUENTIAL DECISION MAKING)

Latent Quantities



Interaction Protocol (time $t = 1, 2, \dots$)



Research question (informal): For a given feedback filter, how many observations (feedback signals) does a learner need to recover the utility ranking?

PERFORMANCE INDEX: (ε, δ) -PAC RANKING RECOVERY

Close utilities: $|u_i - u_j| < \varepsilon \Rightarrow$ swapping i and j is tolerated.

Separated utilities: $|u_i - u_j| \geq \varepsilon \Rightarrow$ their order must be correct.

Key Novelty: Utility-Sensitive ε -Accuracy

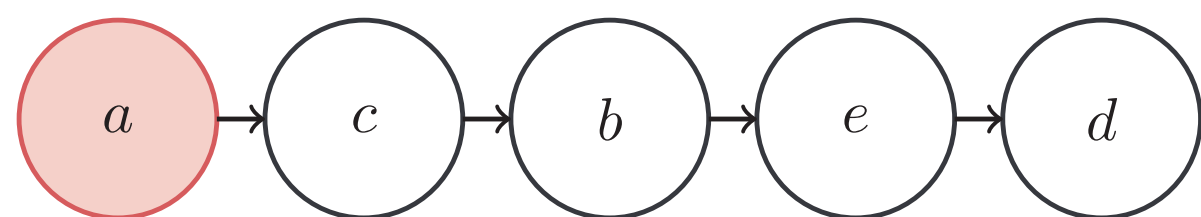
A ranking $\hat{\sigma}$ is ε -accurate if, for every $i \neq j$,

$$(\sigma_u(i) - \sigma_u(j))(\hat{\sigma}(i) - \hat{\sigma}(j)) < 0 \implies |u_i - u_j| < \varepsilon.$$

PAC requirement (3): $\mathbb{P}(\hat{\sigma} \text{ is } \varepsilon\text{-accurate}) \geq 1 - \delta.$

TWO FEEDBACK MODELS: OBSERVATIONS, ALGORITHMS, AND GUARANTEES

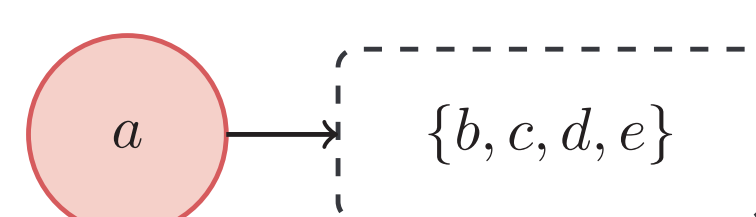
Full-Ranking (FR)



Rank-breaking (1) converts this into $\binom{k}{2}$ **pairwise relations**.

Winner-Only feedback *hides* comparisons among non-winners, contributing to **additional learning difficulty**.

Winner-Only (WO)



Only $k - 1$ **winner-loser relations** are observed.

Assume: The noise distribution ν is log-concave (LC) and its variance is upper bounded by $V > 0$ (V).

Assume (beyond (LC) and (V)): The utility range is smaller than $|\text{Supp}(\nu)|$ (WO1); the reversed-hazard-rate ratio is non-increasing (WO2).

FR-VUR Algorithm

Parameters: δ, ε, V

- **Rank breaking** turns rankings into pairwise comparison outcomes.
- The **pairwise-preferences** $P_{ij} := \mathbb{P}(U_i > U_j)$ are estimated.
- The stopping rule uses a **Bernstein-type confidence radius** (2).

WO-VUR Algorithm

Parameters: δ, ε, V

- Winner observations are turned into **item-wise Bernoulli outcomes**.
- The **winning probabilities** $P_i := \mathbb{P}(U_i > U_j \forall j \neq i)$ are estimated.
- The stopping rule uses a **Bernstein-type confidence radius**.

Optimal Sample Complexity

up to logarithmic factors (Theorems 1 & 2)

$$\mathbb{E}[T] = \tilde{O}\left(\frac{V}{\varepsilon^2} \log \frac{1}{\delta}\right)$$

Optimal Sample Complexity

up to logarithmic factors (Theorems 3 & 4)

$$\mathbb{E}[T] = \tilde{O}\left(\frac{V}{\varepsilon^2 \min_{i \in [k]} P_i} \log \frac{1}{\delta}\right)$$

ONE ALGORITHMIC ENGINE: UTILITY-TO-PROBABILITY MAPPING

Small-gap regime: $\Delta_{ij} := u_i - u_j \lesssim \sqrt{V}$. Universal constants and logarithmic factors are suppressed.

Pairwise-Preference Probabilities (FR)

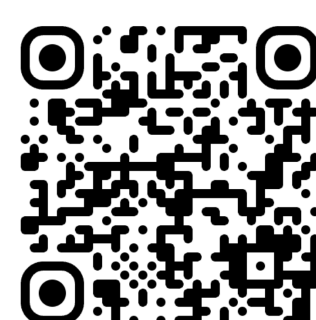
$$P_{ij} - \frac{1}{2} = \frac{1}{2} \eta(\Delta_{i,j}), \quad \text{where } \eta(\Delta_{i,j}) \gtrsim \frac{\Delta_{i,j}}{\sqrt{V}}.$$

Winning Probabilities (WO)

$$P_i - P_j \geq P_i \zeta(\Delta_{i,j}), \quad \text{where } \zeta(\Delta_{i,j}) \gtrsim \frac{\Delta_{i,j}}{\sqrt{V}}.$$

REFERENCES

- [1] H. Azari Soufiani, D. C. Parkes, and L. Xia. Computing parametric ranking models via rank-breaking. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2014.
- [2] A. Maurer and M. Pontil. Empirical Bernstein bounds and sample-variance penalization. In *Annual Conference on Learning Theory (COLT)*, 2009.
- [3] L. G. Valiant. A theory of the learnable. *Communications of the ACM*, 1984.



Paper