JoLA: Job Landscape Aware Job Recommendation

RecSys in HR 2025

Solal Nathan, Guillaume Bied, Elia Perennes, Philippe Caillou, Bruno Crepon, Christophe Gaillac, Michèle Sebag.

26th of September, 2025

Introduction & Motivation

Position of the problem

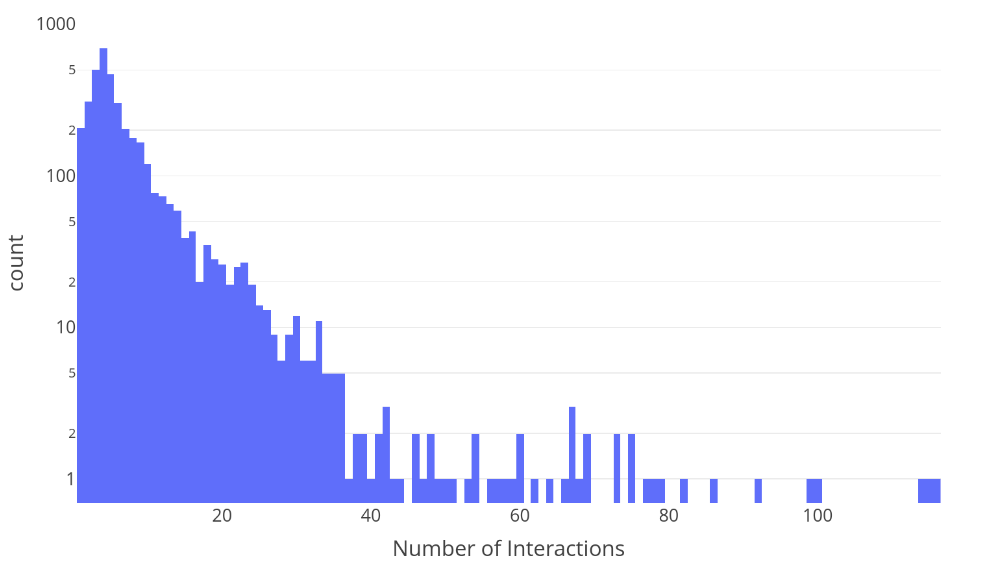


Figure 1: Job Market Imbalance example. Some jobs are extremely popular while others receive nearly no applications

Position of the problem

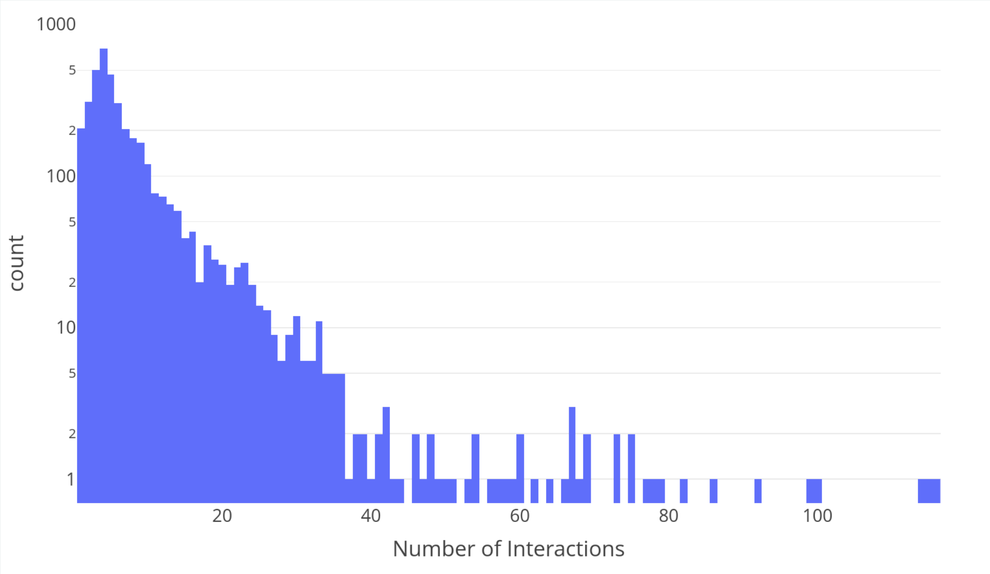


Figure 1: Job Market Imbalance example. Some jobs are extremely popular while others receive nearly no applications

Congestion Phenomenon: Popular jobs get flooded with applications, creating severe competition.

Position of the problem

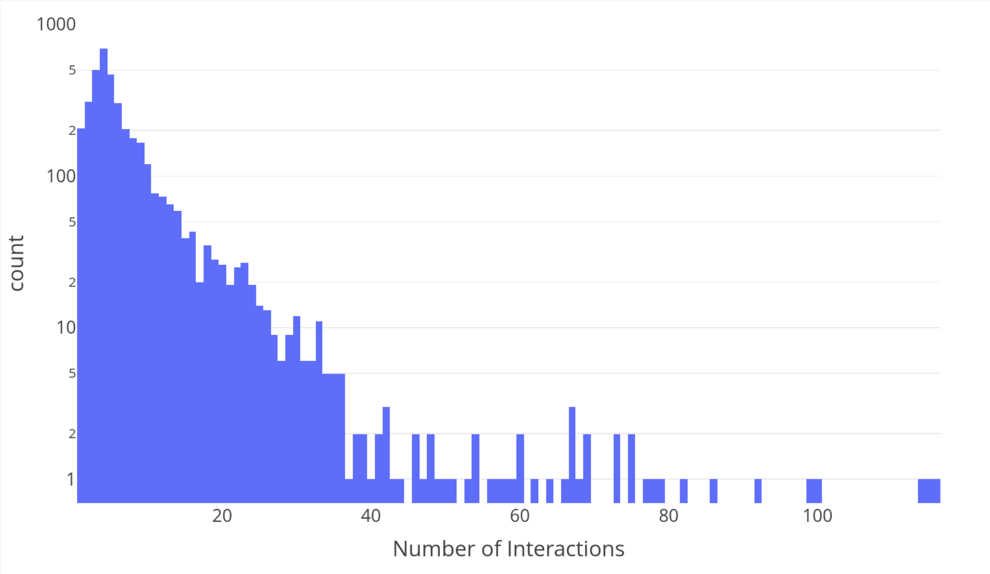


Figure 1: Job Market Imbalance example. Some jobs are extremely popular while others receive nearly no applications

Congestion Phenomenon: Popular jobs get flooded with applications, creating severe competition.

Orphan Jobs: Many job ads receive very few or no applications

- Inefficiency of the market → **frictional unemployement**

Objective:

- Efficient Job Recommender Systems (JRS) in HR / Recall
- In processing congestion avoidance

Novel Perspective: Congestion avoidance has been extensively studied [1][2],

Orphan-job phenomenon is largely overlooked in the literature [3].

[1]Y. Mashayekhi, B. Kang, J. Lijffijt, and T. De Bie, "ReCon: Reducing Congestion in Job Recommendation Using Optimal Transport," in *Proceedings of the 17th ACM Conference on Recommender Systems*, ACM, 2023, pp. 696–701. doi: 10.1145/3604915.3608817.

[2]H. Abdollahpouri, M. Mansoury, R. Burke, and B. Mobasher, "The Impact of Popularity Bias on Fairness and Calibration in Recommend Ation," *arXiv.org*, 2019.

[3]G. K. Patro, A. Biswas, N. Ganguly, K. P. Gummadi, and A. Chakraborty, "FairRec: Two-Sided Fairness for Personalized Recommendations in Two-Sided Platforms," in *Proceedings of The Web Conference 2020*, 2020, pp. 1194–1204.

JoLA: Job Landscape Aware Recommendation

Model Baseline

Items

Users	j_1	j_2	j_3	j_4	j_5
u_1	**1**	0	**1**	**1**	0
u_2	**1**	**1**	0	**1**	**1**
u_3	**1**	0	**1**	**1**	0
u_4	0	0	0	**1**	0

Interaction Matrix $M_{i,j}$

Model Baseline

Items

Users	j_1	j_2	j_3	j_4	j_5
u_1	**1**	0	**1**	**1**	0
u_2	**1**	**1**	0	**1**	**1**
u_3	**1**	0	**1**	**1**	0
u_4	0	0	0	**1**	0

Interaction Matrix $M_{i,j}$

$\longrightarrow$

-1	-0.3
-0.7	0
-0.4	0.3
-0.1	0.6

User Embeddings $\varphi(i)$

Model Baseline

Items

Users	j_1	j_2	j_3	j_4	j_5
u_1	**1**	0	**1**	**1**	0
u_2	**1**	**1**	0	**1**	**1**
u_3	**1**	0	**1**	**1**	0
u_4	0	0	0	**1**	0

Interaction Matrix $M_{i,j}$

-1	-0.3
-0.7	0
-0.4	0.3
-0.1	0.6

User Embeddings $\varphi(i)$

-1	-0.7	-0.4	-0.1	0.2
-0.5	-0.2	0.1	0.4	0.7

Item Embeddings $\psi(j)$

Model Baseline

Items

$$
\begin{array}{cc}
 & \begin{array}{ccccc} j_1 & j_2 & j_3 & j_4 & j_5 \end{array} \\
\text{Users} \begin{array}{c} u_1 \\ u_2 \\ u_3 \\ u_4 \end{array} & \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}
\end{array}
\longrightarrow
\begin{bmatrix} -1 & -0.3 \\ -0.7 & 0 \\ -0.4 & 0.3 \\ -0.1 & 0.6 \end{bmatrix}
$$

Interaction Matrix $M_{i,j}$

User Embeddings $\varphi(i)$

$$
\downarrow
$$

$$
\begin{bmatrix} -1 & -0.7 & -0.4 & -0.1 & 0.2 \\ -0.5 & -0.2 & 0.1 & 0.4 & 0.7 \end{bmatrix} \longrightarrow
$$

Item Embeddings $\psi(j)$

Score Computation

$$s(i, j) = \varphi(i) \cdot \psi(j)$$

Model Baseline

Score function (appeal between user i and job j)

$$s(i, j) = \sigma(\varphi(i) \cdot \psi(j) + b)$$

with $\varphi(i)$ and $\psi(j)$ the user and item embeddings and b is bias.

Model Baseline

Score function (appeal between user i and job j)

$$s(i,j) = \sigma(\varphi(i) \cdot \psi(j) + b)$$

with $\varphi(i)$ and $\psi(j)$ the user and item embeddings and b is bias.

Baseline Model uses a Binary Cross Entropy (BCE)

$$\mathcal{L}_{\text{BCE}}(s) = \left(1 - M_{i,j}\right) \log(1 - s(i,j)) + M_{i,j} \log(s(i,j))$$

with $M_{i,j}$ the interaction matrix.

Market Share Awareness

Market Share Definition: Number of job seekers recommended a specific job ad.

$$\mathrm{MS}(j) = \#\text{users where } s(i,j) \geq s(i,j_k)$$

where j_k is the k$^{\text{th}}$ recommended item to user i, for the score function s

Market Share Awareness

	j_1	j_2	j_3	j_4	j_5
u_1	−0.45	0.1	0.35	0.6	−0.15
u_2	−0.2	0.05	0.6	−0.15	0.1
u_3	0.05	0.3	−0.45	0.1	0.35
u_4	0.3	0.55	−0.2	0.05	0.6

Score Matrix $s(i, j)$

Market Share Awareness

	j_1	j_2	j_3	j_4	j_5
u_1	-0.45	0.1	0.35	0.6	-0.15
u_2	-0.2	0.05	0.6	-0.15	0.1
u_3	0.05	0.3	-0.45	0.1	0.35
u_4	0.3	0.55	-0.2	0.05	0.6

Score Matrix $s(i, j)$

$\longrightarrow$

0	1	1	1	0
0	1	1	0	1
0	1	0	1	1
1	1	0	0	1

Top-k Recommendations ($k = 3$)

Market Share Awareness

	j_1	j_2	j_3	j_4	j_5
u_1	−0.45	0.1	0.35	0.6	−0.15
u_2	−0.2	0.05	0.6	−0.15	0.1
u_3	0.05	0.3	−0.45	0.1	0.35
u_4	0.3	0.55	−0.2	0.05	0.6

Score Matrix $s(i, j)$

0	1	1	1	0
0	1	1	0	1
0	1	0	1	1
1	1	0	0	1

Top-k Recommendations ($k = 3$)

1	4	2	2	3

Item Market Share

Congestion & Orphan

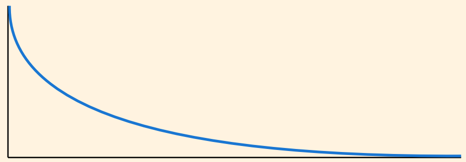
Item Market Share

Congestion & Orphan

Item Market Share

Congestion

$$C = -\sum \mathrm{MJ}_j \log \mathrm{MS}_j$$

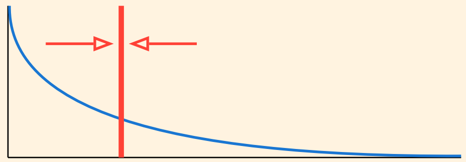


Congestion & Orphan

Item Market Share

Congestion

$$C = -\sum \mathrm{MJ}_j \log \mathrm{MS}_j$$

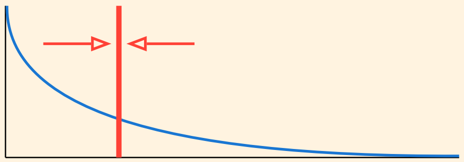


Congestion & Orphan

Item Market Share

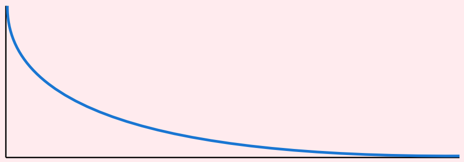
Congestion

$$C = -\sum \mathrm{MJ}_j \log \mathrm{MS}_j$$

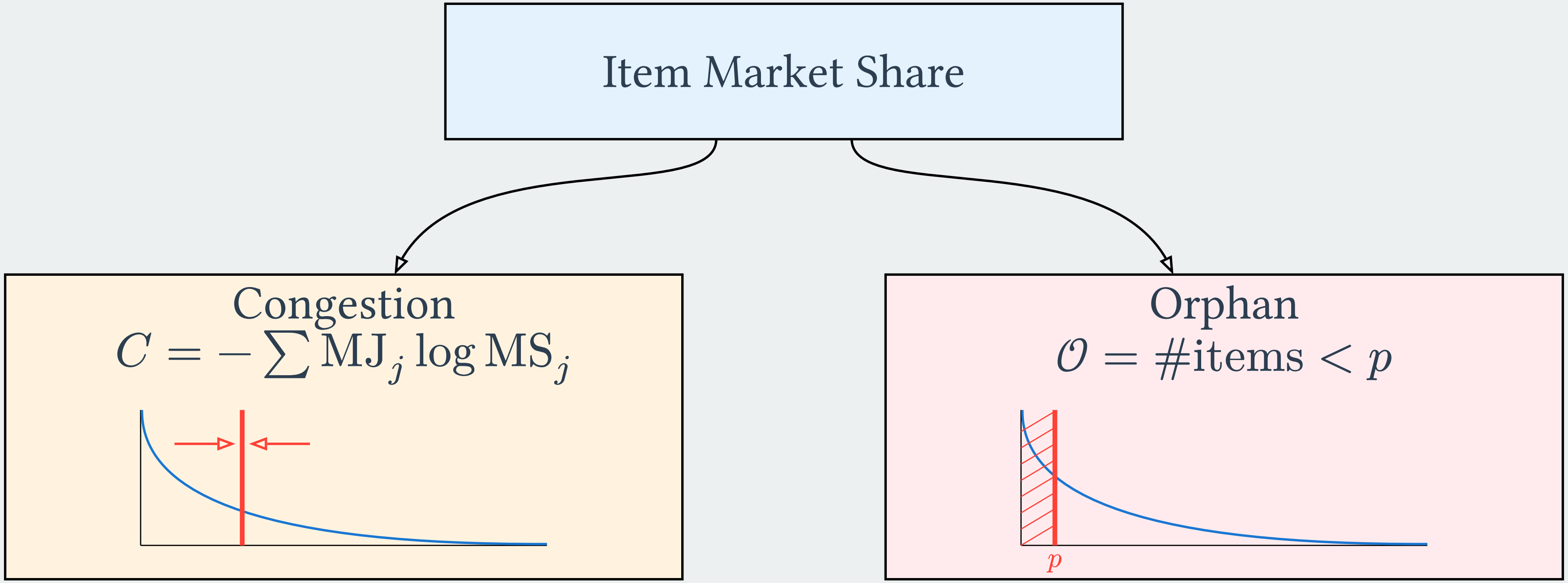


Orphan

$$\mathcal{O} = \#\text{items} < p$$



Congestion & Orphan



Differentiable Approximationon of the Market Share

Differentiable Approximation: gradient-based optimization:

$$\mathrm{MS}(j) = \frac{1}{\mu} \sum_i \max(0, s(i,j) - s(i,j_k))$$

where $\mu = \frac{1}{nk} \sum_{i,j} \max(0, s(i,j) - s(i,j_k))$ Frugal compared to a fully differentiable top-k using [1]

[1]F. Petersen, H. Kuehne, C. Borgelt, and O. Deussen, "Differentiable Top-k Classification Learning." Accessed: Jun. 13, 2024. [Online]. Available: http://arxiv.org/abs/2206.07290

JoLA-Congestion (JoLA-c)

Objective: Accuracy + Congestion

Loss Function:

$$\mathcal{L}_{\text{Congestion}}(s) = \mathcal{L}_{\text{BCE}}(s) + \alpha \cdot \text{Congestion}(s)$$

Congestion Metric: Entropy of normalized market shares

$$\text{Congestion}(s) = -\sum_{j \in [m]} \text{MS}_{n(j)} \log \text{MS}_{n(j)}$$

Remark: Encourages balanced exposure across all job ads.

JoLA-Orphan (JoLA-o)

Objective: Accuracy + Orphans (jobs with less than p applications)

Loss Function:

$$\mathcal{L}_{\mathrm{O}}(s) = \mathcal{L}_{\mathrm{BCE}}(s) + \alpha \cdot \mathcal{L}_{p,k}(s)$$

Orphan Loss: Focuses on jobs with market share < p

$$\mathcal{L}_{p,k}(s) = \frac{1}{\mathcal{O}_{p,k}(s)} \sum_{j \text{ where } \mathrm{MS}(j)<p} (p - \mathrm{MS}(j))^2$$

where $\mathcal{O}_{p,k}$ is the number of orphans. We choose $p_{\max} = \left\lfloor \frac{nm}{k} \right\rfloor$.

JoLA-Orphan-Compound (JoLA-oc)

Objective: Accuracy + Orphans + Popularity Avoidance

Loss Function:

$$\mathcal{L}_{\text{OC}}(s) = \mathcal{L}_{\text{BCE}}(s) + \alpha \cdot \mathcal{L}_{p,k}(s) + \beta \cdot \mathcal{L}^{c}_{p,k}(s)$$

Compound Term: Reduces market share of non-orphan jobs

$$\mathcal{L}^{c}_{p,k}(s) = \frac{1}{n - \mathcal{O}_{\text{p,k}}(s)} \sum_{j \text{ where } \text{MS}(j) \geq p} (\text{MS}(j) - p)$$

Remark: Handles invisible job ads (MS = 0) by linearly penalizing the rest.

Experimental Setup

Datasets and Baselines

Baselines:

- **BCE**: simple Binary Cross entropy loss
- **ReCon** State-of-the-art congestion reduction using Optimal Transport [1]

Datasets: Same setting as ReCon to have **heavy tail** and **warm start**

Dataset	#users	#jobs	#interactions
CareerBuilder-Small (CB-S) (10 days)	3 k	4 k	30 k
CareerBuilder-Large (CB-L) (90 days)	42 k	41 k	470 k

[1] Y. Mashayekhi, B. Kang, J. Lijffijt, and T. De Bie, "Scalable Job Recommendation With Lower Congestion Using Optimal Transport," *IEEE Access*, vol. 12, pp. 55491–55505, 2024.

Datasets and Baselines

Evaluation Metrics:

- **Accuracy:** HR@10, Recall@10
- **Congestion:** Entropy of market shares
- **Orphans:** Proportion of (p,k)-orphans (jobs with $< p$ applications when $k=10$, $p=8$)

Training Configuration

Hyperparameters:

- **Embedding dimension:** 512
- **Batch size:** 1,024 (CB-S), 4,096 (CB-L)
- **Learning rate:** 10^{-2} with AdamW optimizer
- **Loss weights:** $\alpha, \beta \in \{10^{-2}, \ldots, 10^{+2}\}$

Training Schedule:

- **Phase 1:** BCE-only warm-start
- **Phase 2:** JoLA-specific losses (congestion, orphan or compound)

Computational Resources:

- **Runtime:** 10 minutes (CB-S), 4.3 hours (CB-L) on consummer GPU (2070S)

Results

Main Findings on CB-S

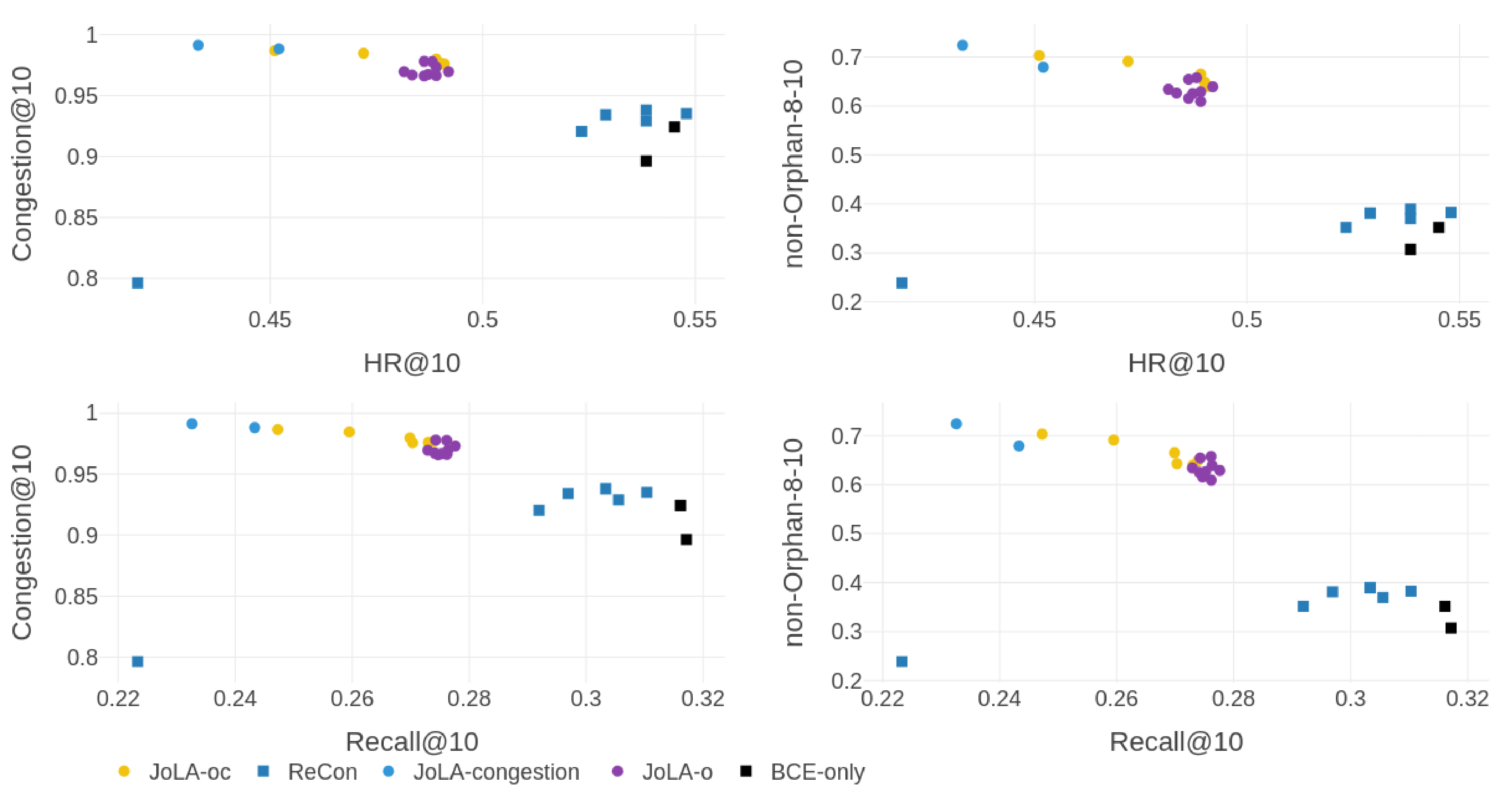


Figure 4: Pareto front on the CB-S dataset between performance and congestion metrics

Main Findings on CB-L

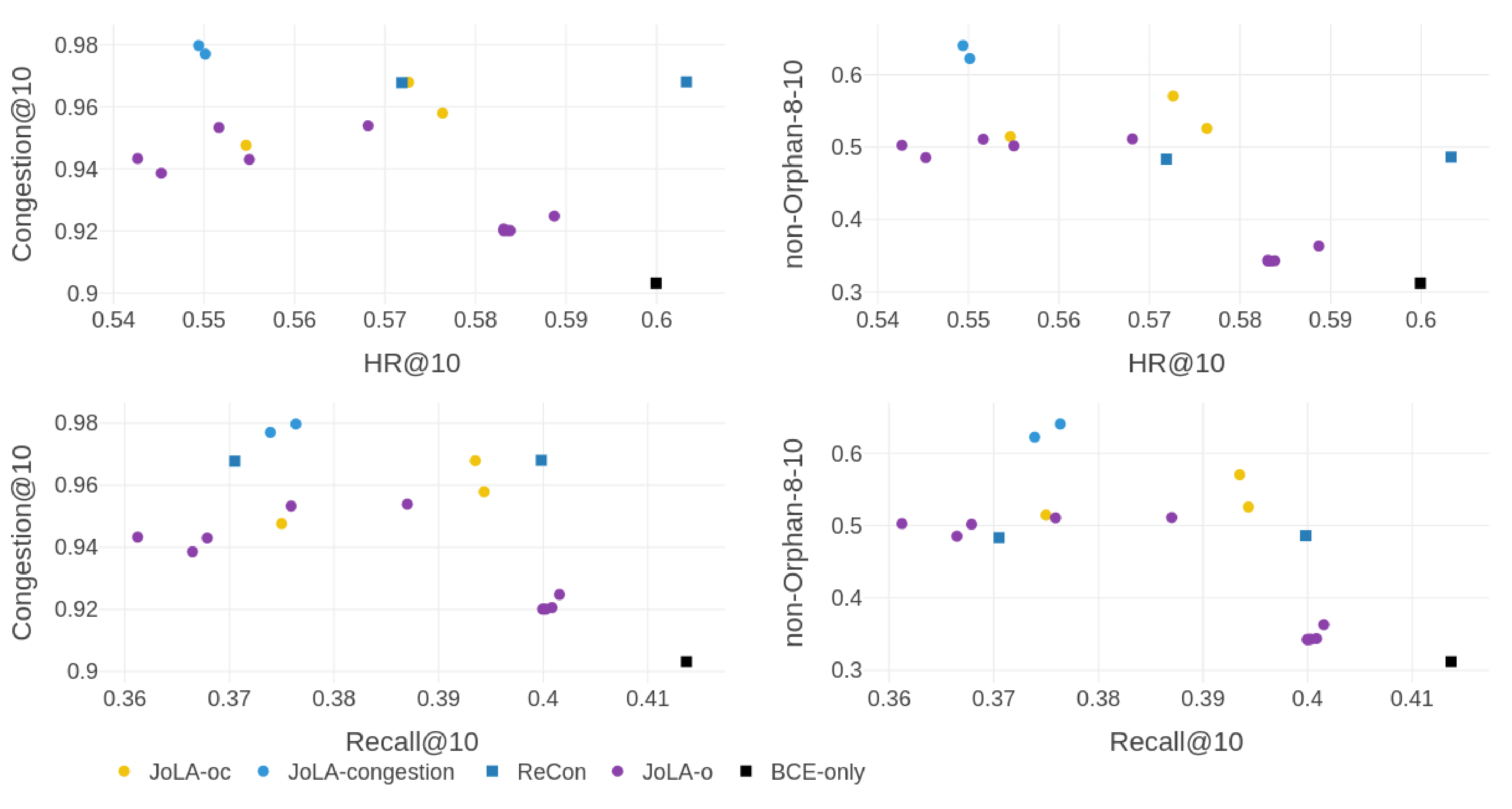


Figure 5: Pareto front on the CB-L dataset between performance and congestion metrics

Discussion

Orphan Metric Advantages: More fine-grained than traditional congestion metrics, directly addresses invisible job ads (MS = 0)

Discussion

Orphan Metric Advantages: More fine-grained than traditional congestion metrics, directly addresses invisible job ads (MS = 0)

Trade-off Patterns: JoLA-c provides best congestion reduction, JoLA-o offers balanced performance, JoLA-oc enables fine-grained control

Discussion

Orphan Metric Advantages: More fine-grained than traditional congestion metrics, directly addresses invisible job ads (MS = 0)

Trade-off Patterns: JoLA-c provides best congestion reduction, JoLA-o offers balanced performance, JoLA-oc enables fine-grained control

Scalability: Consistent performance across CB-S and CB-L datasets demonstrates method robustness

Discussion

Orphan Metric Advantages: More fine-grained than traditional congestion metrics, directly addresses invisible job ads (MS = 0)

Trade-off Patterns: JoLA-c provides best congestion reduction, JoLA-o offers balanced performance, JoLA-oc enables fine-grained control

Scalability: Consistent performance across CB-S and CB-L datasets demonstrates method robustness

Practical Implications: JoLA can be deployed as in-processing approach or even post-processing, making it compatible with existing JRS infrastructure.

Conclusion & Future Work

Contributions Summary

JoLA Framework: lessons learned on pulic datasets (CB-S, CB-L)

- Novel differentiable market share approximation for gradient-based optimization
- Three complementary loss functions addressing different market balance objectives
- Focus on the orphan jobs, which are under studied in the litterature

Future Research Directions

More datasets: apply JoLA on VDAB and France Travail

Future Research Directions

More datasets: apply JoLA on VDAB and France Travail

Feature-based dataset: Extend JoLA to handle features and not only interaction data

Future Research Directions

More datasets: apply JoLA on VDAB and France Travail

Feature-based dataset: Extend JoLA to handle features and not only interaction data

Real-world Deployment: Investigate JoLA performance in production job recommendation systems (AB testing)

Questions?

Paper

Source Code

Contact

solal.nathan@inria.fr

@solalnathan.com

@solalnathan@sigmoid.social