Supplementary Material - ReCon: Reducing Congestion in Job Recommendation using Optimal Transport

Yoosof Mashayekhi
yoosof.mashayekhi@ugent.be
IDLAB - Department of Electronics and Information Systems, Ghent University
Ghent, Belgium

Bo Kang
bo.kang@ugent.be
IDLAB - Department of Electronics and Information Systems, Ghent University
Gent, Belgium

Jefrey Lijffijt
jefrey.lijffijt@ugent.be
IDLAB - Department of Electronics and Information Systems, Ghent University
Gent, Belgium

Tijl De Bie
tijl.debie@ugent.be
IDLAB - Department of Electronics and Information Systems, Ghent University
Gent, Belgium

ACM Reference Format:
Yoosof Mashayekhi, Bo Kang, Jefrey Lijffijt, and Tijl De Bie. 2023. Supplementary Material - ReCon: Reducing Congestion in Job Recommendation using Optimal Transport. In *Seventeenth ACM Conference on Recommender Systems (RecSys '23), September 18–22, 2023, Singapore, Singapore.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3604915.3608817

A EXPERIMENTAL EVALUATION

A.1 Details of hyper-parameters of the recommendation models

In this section we describe the hyper-parameters used.

For the base CNE model, we use the degree prior (for more information please see [1]) and use dot product as the distance function.

We optimize both CNE and ReCon with AdamW optimizer. We use early stopping based on validation loss for CNE and based on Congestion@1 for ReCon in the training phase. We also perform hyper-parameter tuning using Bayesian optimization based on validation Hit Rate@10 for CNE. For ReCon, we select a set of hyper-parameters that result in a good trade-off between Congestion@1 and validation Hit Rate@10, by first using Bayesian optimization based on Congestion@1 and selecting a model based on both measures. Table 1 shows the selected hyper-parameters after hyper-parameter tuning.

A.2 Baseline comparison

Here we compare the methods in terms of the desirability measures and congestion-related measures (**Q1**). Figures 1, 2, 3 and Figures 4, 5, 6 show the performance of all methods for VDAB and CareerBuilder datasets, respectively. They all compare a desirability measure (NDCG, Recall, or Hit Rate) and a congestion-related measure (Congestion, Coverage, or Gini Index). We can observe that for some selections of hyper-parameters, ReCon usually finds a good trade-off between both measures.

A.3 Hyper-parameter sensitivity analysis for λ

In this section, we analyze the performance of ReCon for different values of λ (the weight of the optimal transport in ReCon). Figure 7 shows different measures for different values of λ for both datasets. As expected, we can observe that congestion-related measures are mostly improved by increasing λ. Surprisingly, decreasing the value of λ does not necessarily result in an improvement in desirability measures. Difficulties in the training phase for very small values of λ could be the reason for this observation. However, more experiments are needed to validate this hypothesis.

RecSys '23, September 18–22, 2023, Singapore, Singapore

ACM ISBN 979-8-4007-0241-9/23/09.
https://doi.org/10.1145/3604915.3608817

Table 1: The selected hyper-parameters after hyper-parameter tuning.

Dataset&Method	VDAB-CNE	CareerBuilder-CNE	VDAB-ReCon	CareerBuilder-ReCon
Batch size	512	4096	256	512
Learning rate	0.001	0.0005	0.0005	0.0005
Weight decay	0.01	0.1	0	0.001
Embedding dimension	128	32	128	128

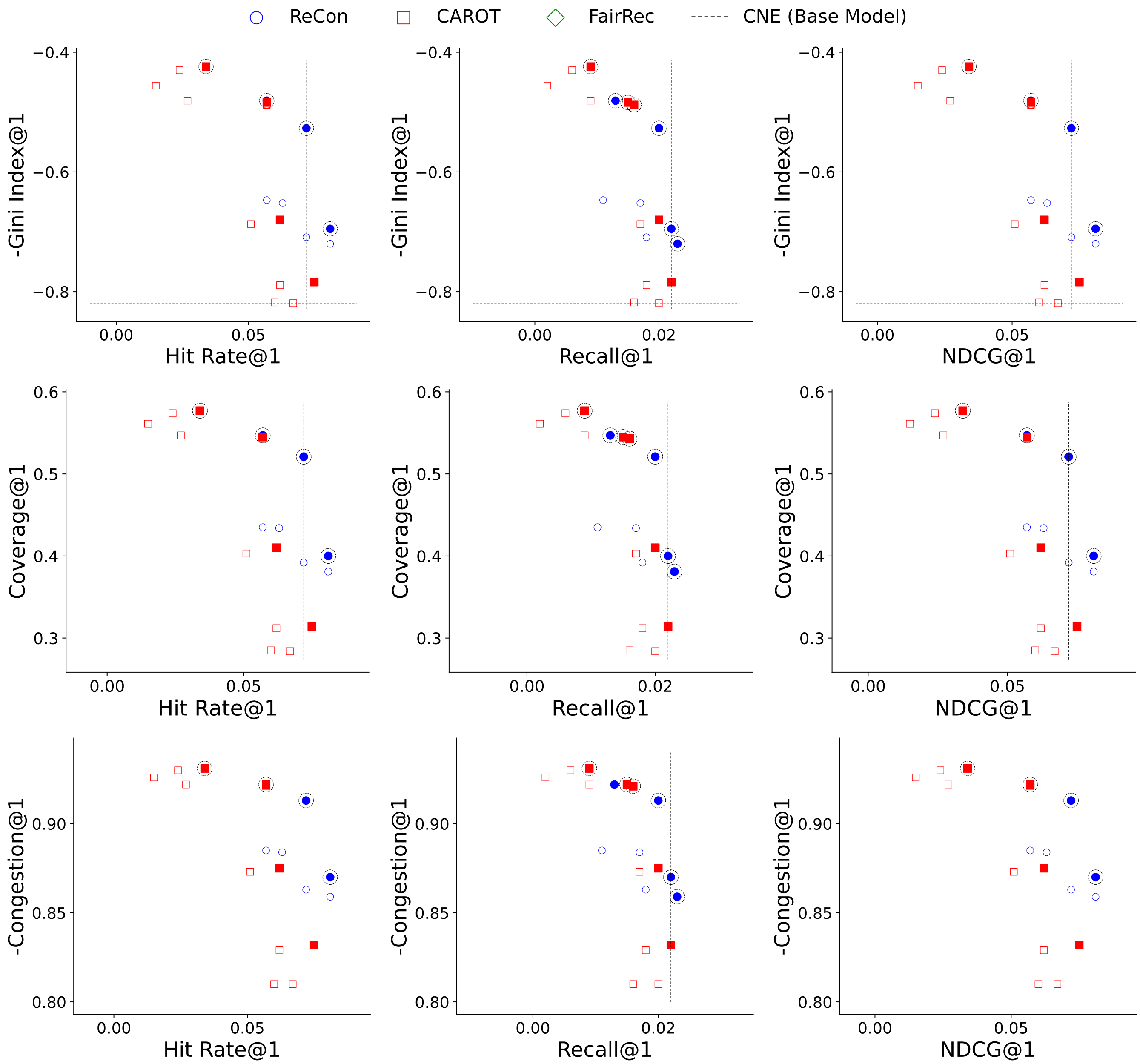


Figure 1: Desirability versus congestion-related measures in VDAB dataset for top-1 recommendation (higher values are better). Points represent different hyper-parameter combinations. Pareto optimal points per method are filled. Pareto optimal points across methods have a circle around.

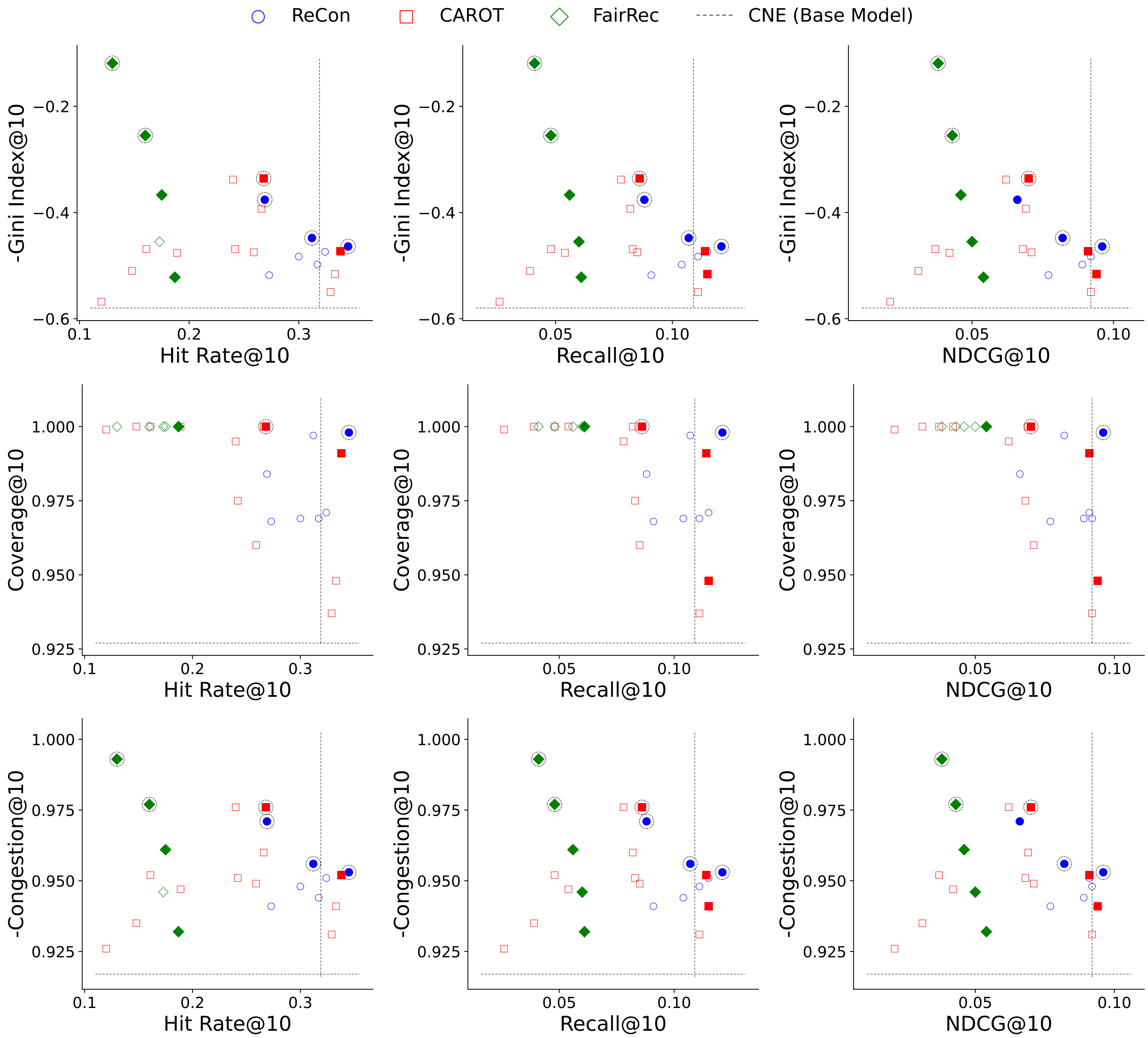


Figure 2: Desirability versus congestion-related measures in VDAB dataset for top-10 recommendation (higher values are better). Points represent different hyper-parameter combinations. Pareto optimal points per method are filled. Pareto optimal points across methods have a circle around.

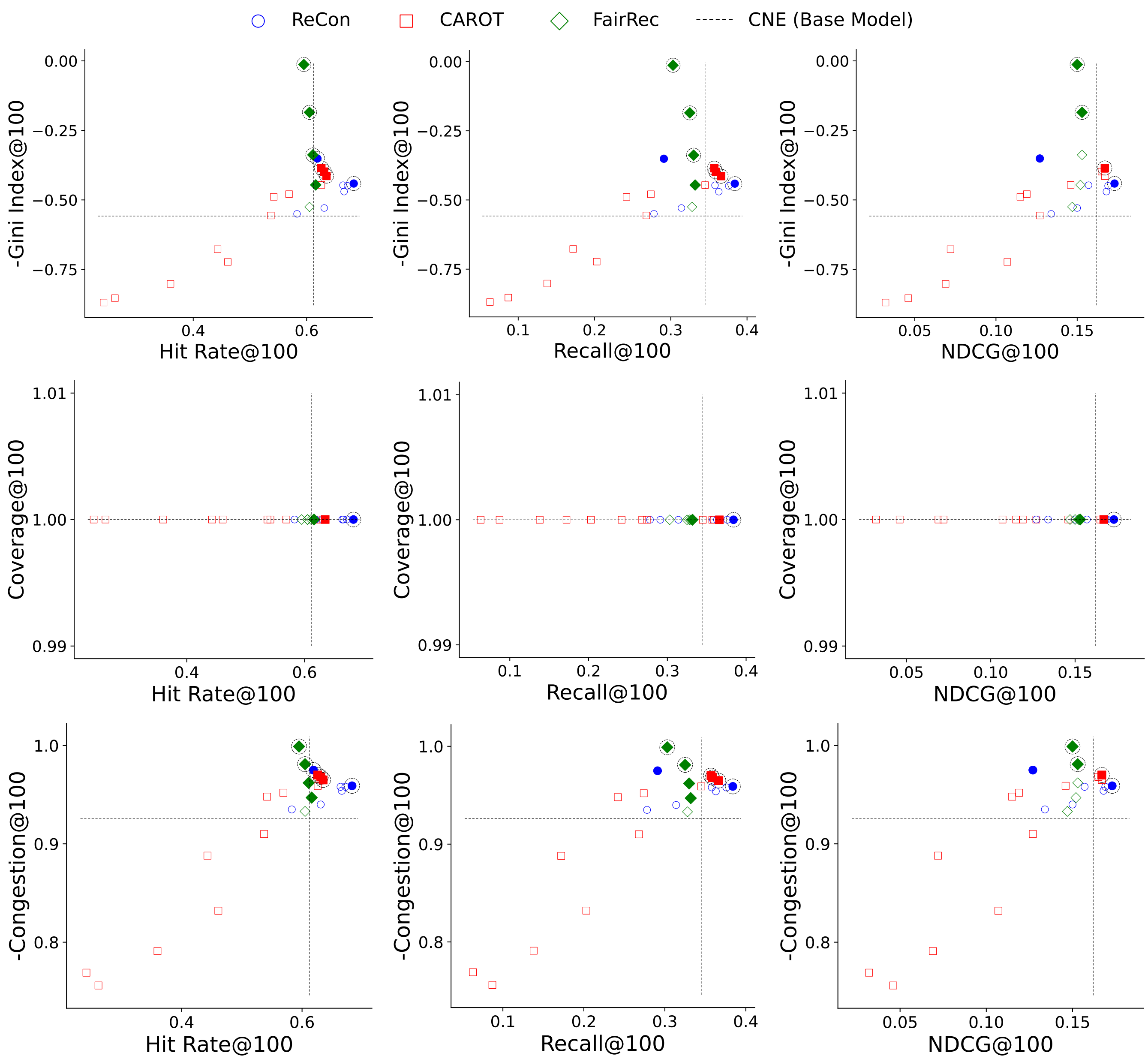


Figure 3: Desirability versus congestion-related measures in VDAB dataset for top-100 recommendation (higher values are better). Points represent different hyper-parameter combinations. Pareto optimal points per method are filled. Pareto optimal points across methods have a circle around.

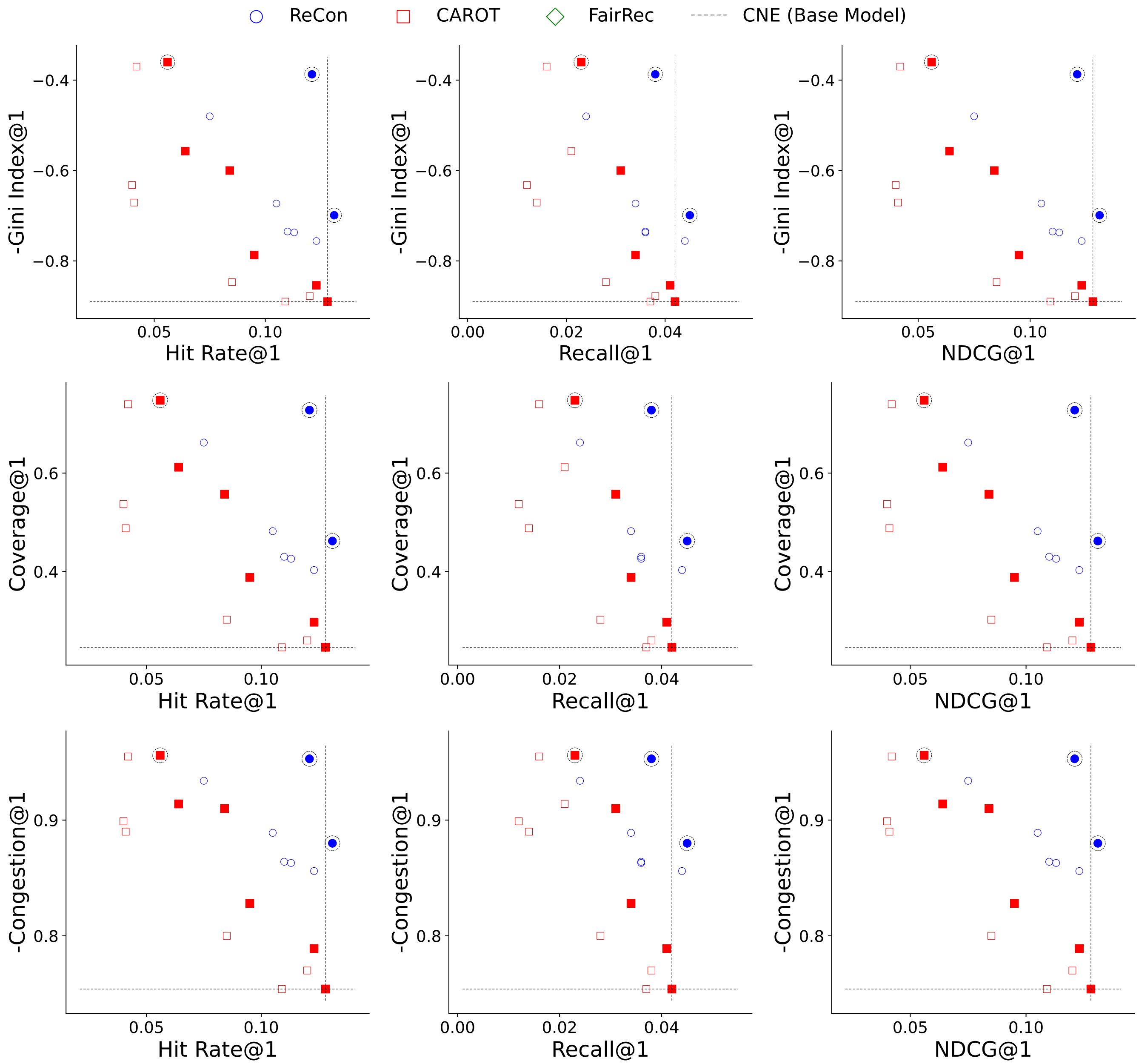


Figure 4: Desirability versus congestion-related measures in CareerBuilder dataset for top-1 recommendation (higher values are better). Points represent different hyper-parameter combinations. Pareto optimal points per method are filled. Pareto optimal points across methods have a circle around.

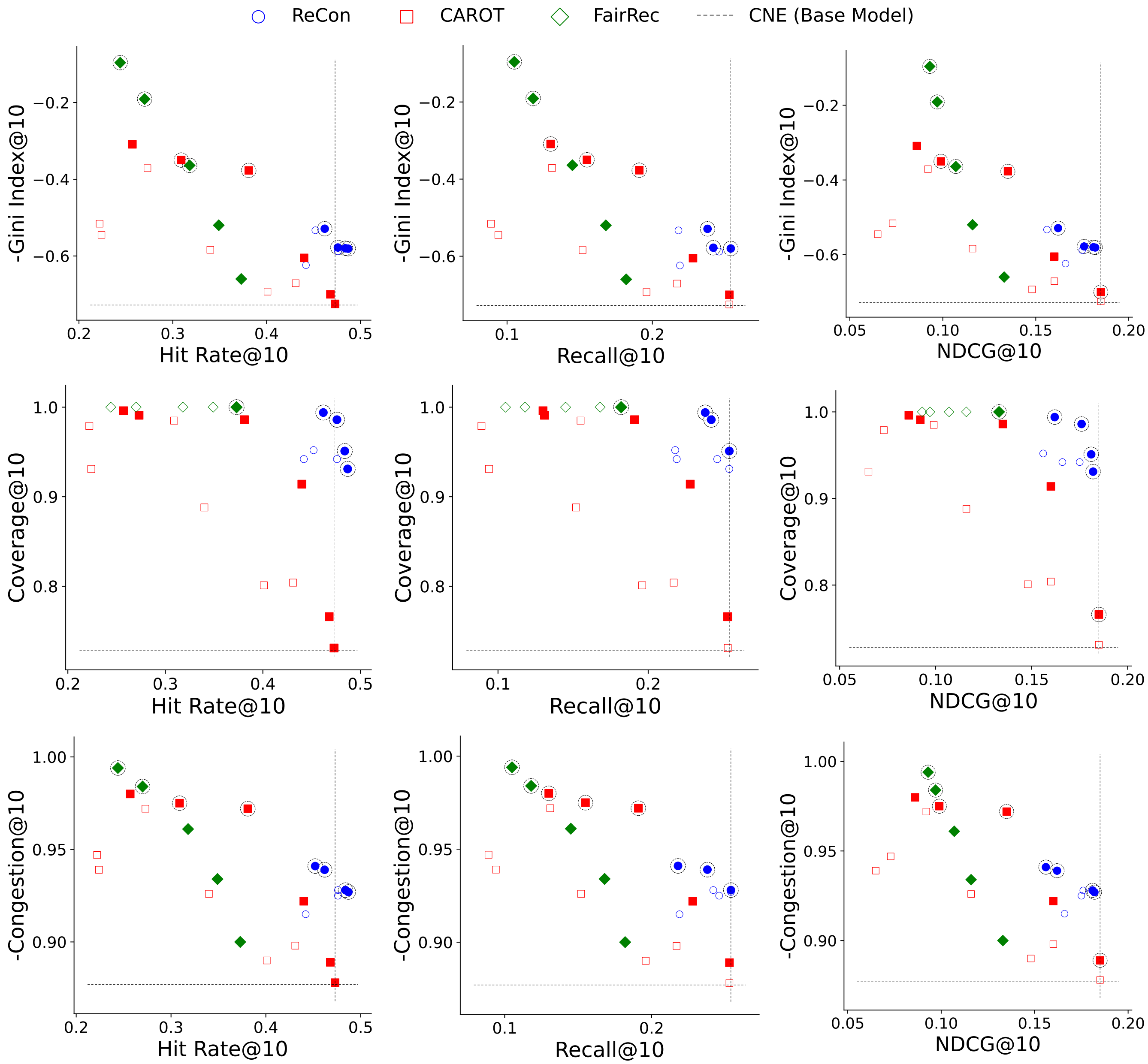


Figure 5: Desirability versus congestion-related measures in CareerBuilder dataset for top-10 recommendation (higher values are better). Points represent different hyper-parameter combinations. Pareto optimal points per method are filled. Pareto optimal points across methods have a circle around.

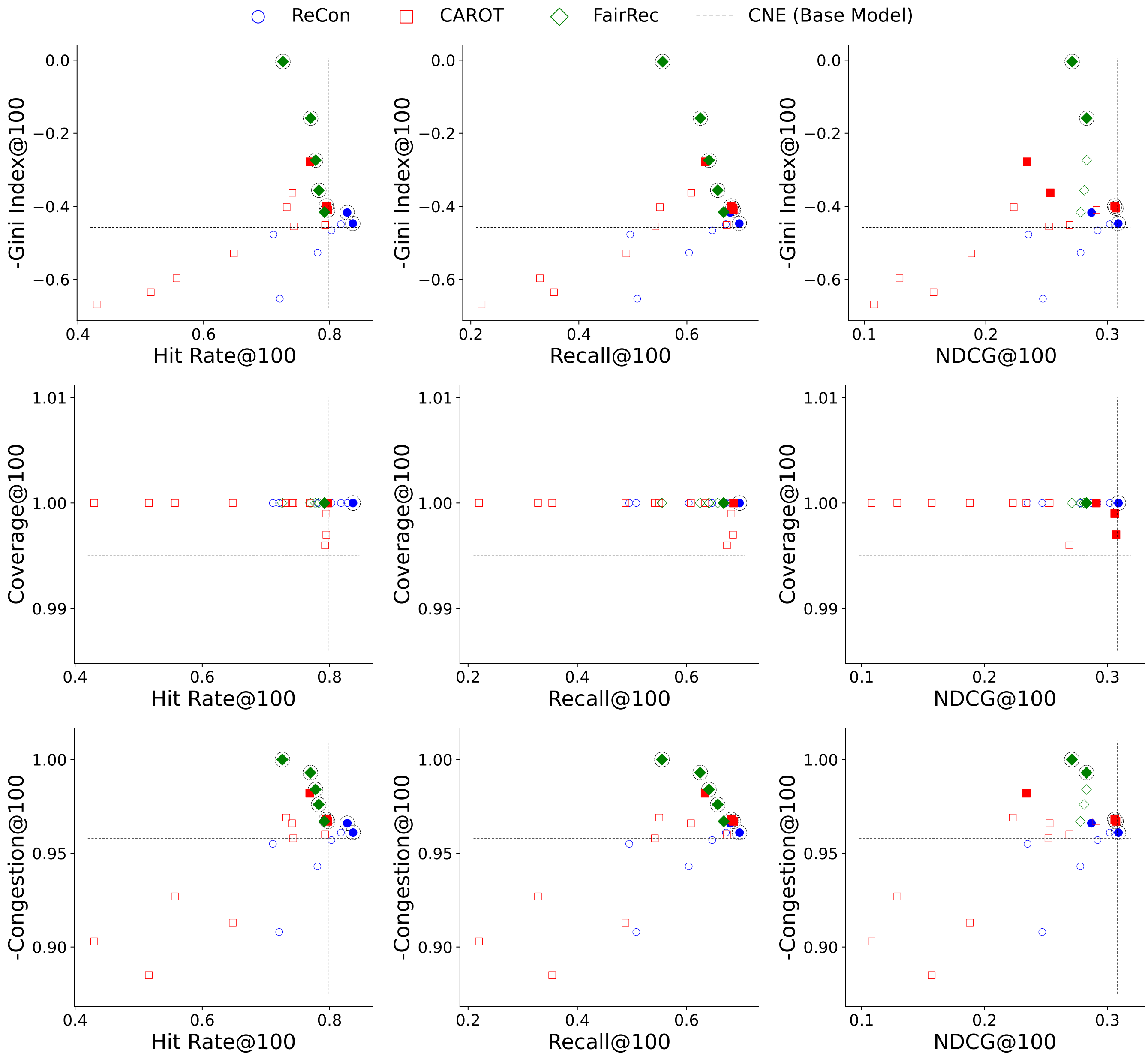
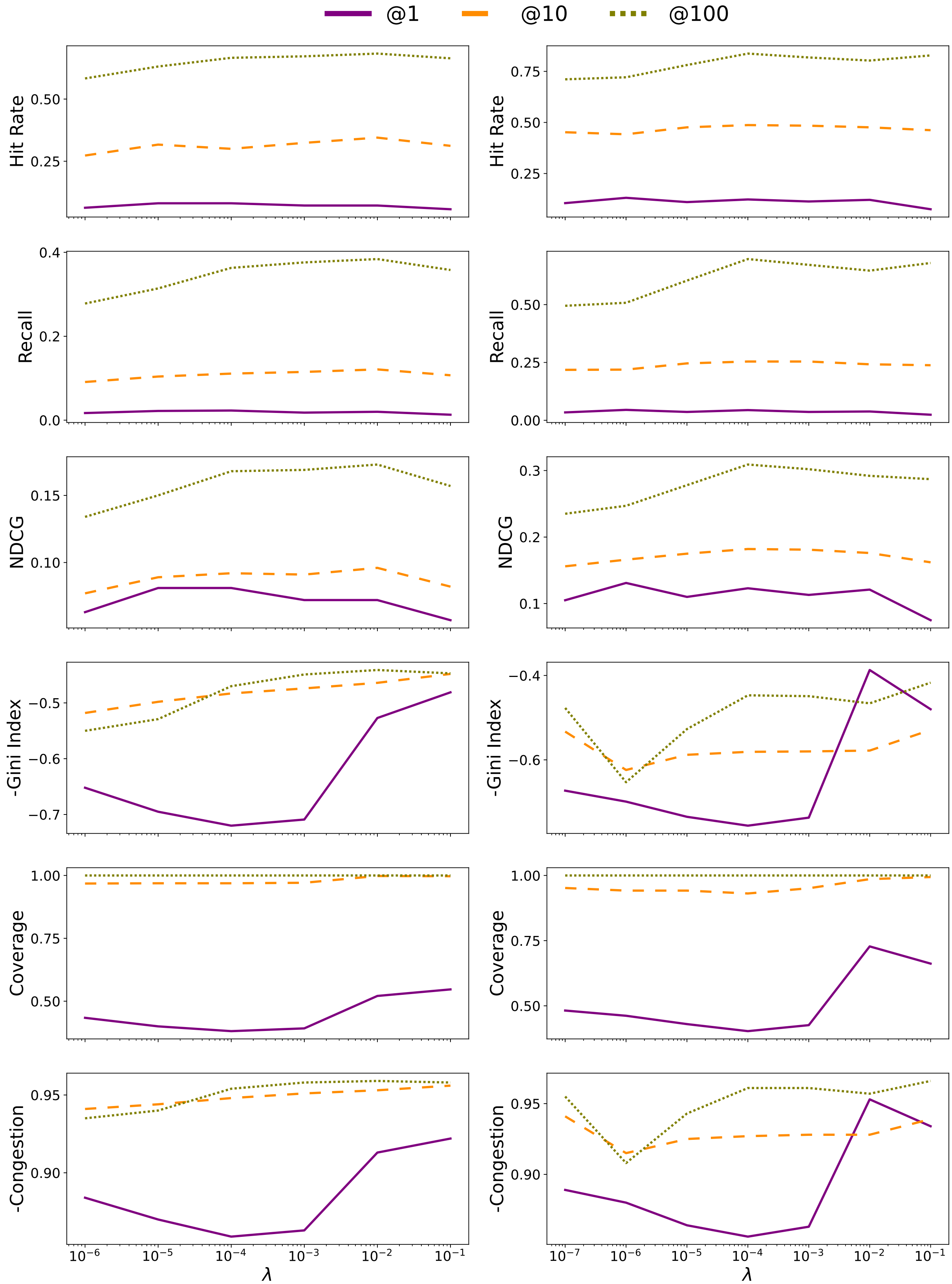


Figure 6: Desirability versus congestion-related measures in CareerBuilder dataset for top-100 recommendation (higher values are better). Points represent different hyper-parameter combinations. Pareto optimal points per method are filled. Pareto optimal points across methods have a circle around.



(a) VDAB

(b) CareerBuilder

Figure 7: Performance of ReCon for different values of λ (the weight of the optimal transport in ReCon).

REFERENCES

[1] Bo Kang, Jefrey Lijffijt, and Tijl De Bie. 2019. Conditional network embeddings. In *7th International Conference on Learning Representations, ICLR 2019*.