Ho Chi Minh City University of Technology
Faculty of Computer Science and Engineering · VNU-HCM

MAY 2026

PROJECT CODE • HK252-DATN-070 · MAJOR: COMPUTER SCIENCE

Diversity Approach to Knowledge-Graph Based Recommendation Systems

STUDENTS Nguyen Ngoc Song Thuong (2252803) · Nguyen Huu Thanh (2552742) SUPERVISORS Assoc. Prof. Dr. Thoai Nam · Dr. Diep Thanh Dang REVIEWER MEng. Nguyen Tuan Khoi

**Popularity bias in KG-based conversational recommenders is a training problem, not architecture.
Two training-time objectives lift catalog coverage up to +81% while preserving recall.**

Problem

Popularity bias is unfair to both sides of the recommendation.

USERS Stated topics get ignored when they fall outside the popular cluster. Discovery narrows to a tiny head of the catalog.

PROVIDERS Long-tail items are systematically denied exposure. A tiny share of the catalog wins most of the recommendation traffic.

KG-CRS (KBRD, KGSF, CR-Walker, TREA) inherit this directly — they train with cross-entropy alone, with no diversity term in the loss.

Why a quick fix fails. List-level diversity losses (ILD, DPP) require selecting a discrete top-K set first, and top-K is non-differentiable. The gradient never reaches the model, so the lever has to be the training objective, not the architecture.

<25% of catalog reached at top-10 on full-catalog KG-CRS

65% of ReDial movies appear ≤ 1× in training

≥0.80 slate Gini on both benchmarks (head-heavy)

0 surveyed KG-CRS models with a diversity-aware loss

GOAL
Add a diversity signal to training *without changing the architecture*, and check whether it generalises across multiple KG-CRS backbones and two benchmarks.

Approach TWO TRAINING-TIME OBJECTIVES

Both methods modify the training objective alone, leaving architecture and inference path intact. They share a temperature-scaled softmax primitive but operate at different optimisation stages — the second method picks up exactly where the first stops applying.

METHOD 1 SRD: Soft-rank loss

A differentiable surrogate for intra-list diversity, plugged into supervised training.

1. **Soft-select.** Replace the hard top-K with a temperature-scaled softmax over all M candidates.
2. **Score pairwise diversity.** Compute the weight-quadratic over the cosine similarity matrix `S` of item embeddings.
3. **Factorise.** Materialise `HᵀW`, never the full `M×M`. Cost is `O(dMB)`.

SOFT-RANK DIVERSITY

$$w_i = \mathrm{softmax}(s_i/\tau)$$
$$L_{\mathrm{div}} = w^\top S\, w, \quad S_{ij} = \cos(e_i, e_j)$$

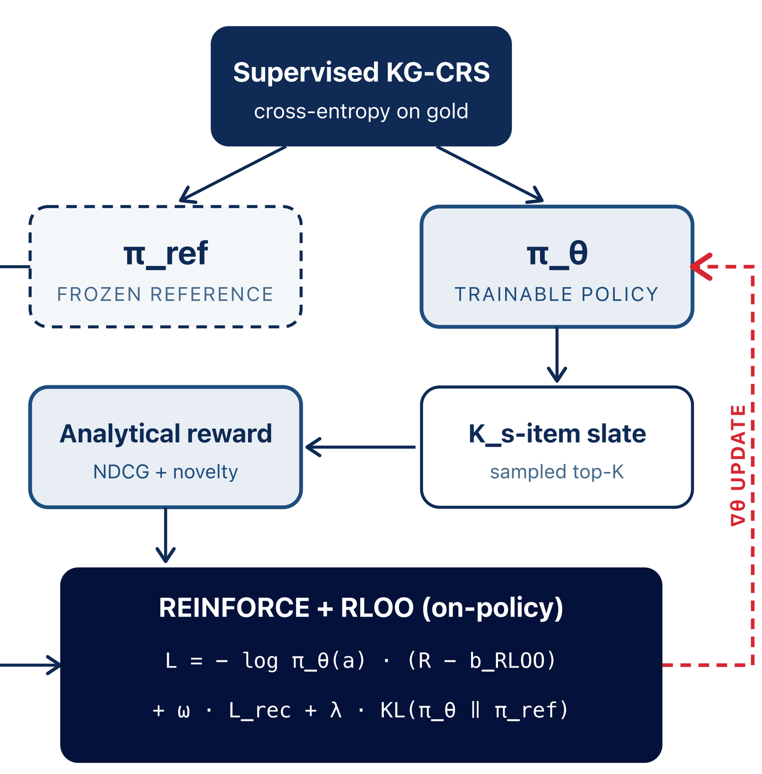
Applies to: backbones whose scoring is restricted to a KG-related candidate set — CR-Walker hop-2.

SRD: LIMITS Where Method 1 stops

- **Pairwise-cosine breaks** on full catalogs.
- **Recall trades for coverage** (KBRD R@1 3.70→3.02, KGSF 3.97→3.50).

METHOD 2 DivKG: REINFORCE fine-tuning

Optimise the non-differentiable diversity reward directly, after supervised pre-training.



DivKG pipeline. Supervised pre-training yields a frozen reference π_{ref}; the policy samples slates of K_s=50 items, the composite reward scores them, and a policy-gradient update with a KL anchor keeps π_θ close to π_{ref}.

SAME SKELETON AS RLHF DivKG borrows the supervised pre-training → KL-anchored policy-gradient pipeline that aligned modern large language models. Treat the trained KG-CRS as a **stochastic policy** and fine-tune with **REINFORCE** on an **analytical reward** composed of **accuracy** (NDCG), **novelty** (long-tail boost), and **relevance-weighted novelty** (diversity tied to user intent).

NDCG	NOVELTY	REL-NOV
$\frac{1}{\log_2(r_{\mathrm{gold}}+2)}$	$\left(1-\frac{\mathrm{pop}}{\mathrm{pop}_{\max}}\right)\cdot\lambda$	$R_{\mathrm{nov}}\cdot[\cos(e_a,e_g)]_+$

DIVKG TOTAL LOSS

$$L = \omega\, L_{\mathrm{rec}} + L_{\mathrm{RL}} + \lambda\, L_{\mathrm{KL}}$$
$$R = \alpha_1 R_{\mathrm{NDCG}} + \alpha_2 R_{\mathrm{nov}} + \alpha_3 R_{\mathrm{rel\text{-}nov}}$$

DIVKG: RESOLUTIONS How Method 2 closes the gaps

- **REINFORCE propagates non-differentiable reward** — works on the full catalog, no candidate-set assumption.
- **KL anchor + cross-entropy mix-in** preserve R@1 across KBRD, KGSF, TREA.

Results

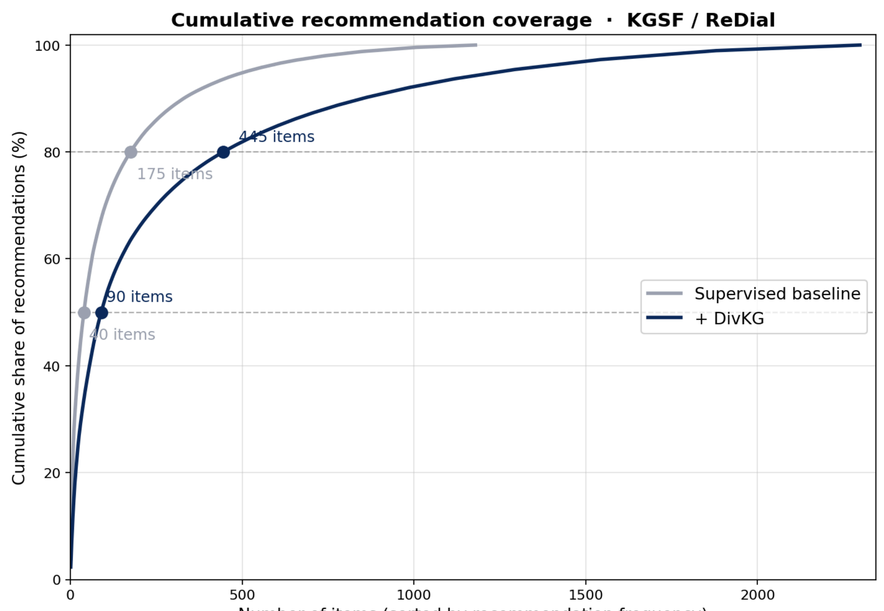
Quantitative — coverage gains

Both methods lift catalog coverage substantially across four KG-CRS backbones on ReDial. DivKG additionally generalises to TG-ReDial.

Model	R@1	R@10	Cov@10	Cov@50
CR-Walker	3.65	15.37	31.76	62.19
+ SRD	3.47	16.17	↑ 34.92	↑ 65.38
KBRD	3.70	18.79	18.37	36.37
+ SRD	↓ 3.02	↓ 16.68	↑ 29.10	↑ 50.51
+ DivKG	3.77	18.79	↑ 32.74	↑ 56.09
KGSF	3.97	19.36	17.03	29.92
+ SRD	↓ 3.50	↓ 18.79	↑ 23.83	↑ 39.90
+ DivKG	4.10	21.16	↑ 30.88	↑ 51.21
TREA	4.84	21.37	23.87	45.26
+ DivKG	4.64	20.10	↑ 41.75	↑ 71.16

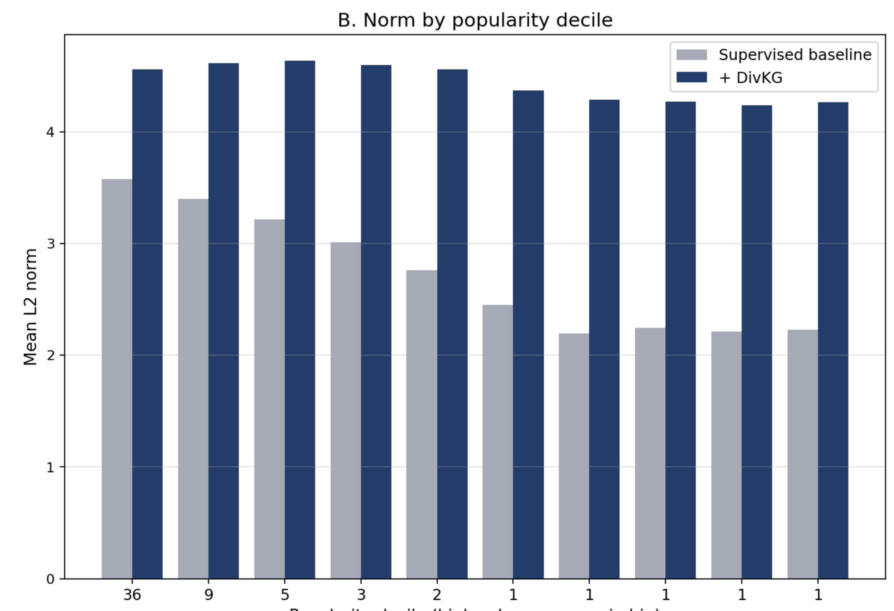
Test results on ReDial (%). SRD rows are Method 1; DivKG rows are Method 2. Cov@10 jumps over the matching baseline: CR-Walker +10%, KBRD +78%, KGSF +81%, TREA +75%.

COVERAGE CURVE



80% of recommendations come from **175** items under baseline; DivKG needs **445** items for the same share — recommendation mass spreads across a far wider catalog slice.

EMBEDDING POPULARITY GAP



Norm-popularity correlation flattens. Baseline assigns systematically smaller norms to long-tail items; DivKG closes the gap so tail items compete in the inner-product score.

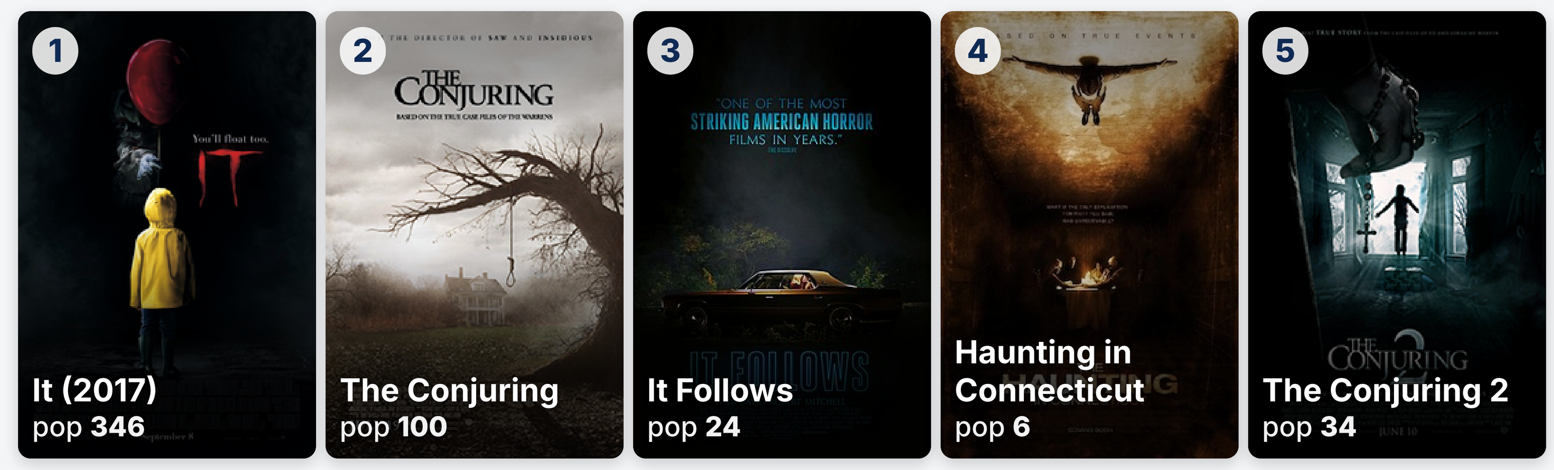
Qualitative — top-5 recommendations

ReDial test index 3803. User mentions horror seeds *A Perfect Getaway*, *Orphan*, *The Purge*; gold is the **long-tail** *Children of the Corn* (training pop 6). The baseline buries the gold under the head-of-distribution horror cluster; DivKG **recovers the gold** AND surfaces a seed movie.

KGSF baseline

avg pop **82.1**

All five slots are filled by head-of-distribution horror titles — *It (2017)* and *The Conjuring* dominate. The training-rare gold ***Children of the Corn*** (training pop 6) is **buried at rank 9**, and none of the user's seed mentions (*A Perfect Getaway*, *Orphan*, *The Purge*) surface in top-5. Cross-entropy alone collapses onto the popular cluster.



KGSF + DivKG

avg pop **53.2**

DivKG **recovers the gold *Children of the Corn*** at **rank 4**, lifts two long-tail titles (*Orphan* and *The Fourth Kind*, training pop 4 each) into the top-5, and keeps a popular safety net (*The Conjuring 2*). Average popularity drops **82.1 → 53.2** while the user's seed intent is preserved.

REFERENCES **KBRD** Chen et al., EMNLP 2019 · **KGSF** Zhou et al., KDD 2020 · **CR-Walker** Ma et al., EMNLP 2021 · **TREA** Li et al., ACL 2023
REINFORCE Williams, Mach. Learn. 1992 · **RLHF** Ouyang et al., NeurIPS 2022 · **ReDial** Li et al., NeurIPS 2018 · **TG-ReDial** Zhou et al., COLING 2020