Bridging Items and Language: A Transition Paradigm for Large Language Model-Based Recommendation

Xinyu Lin[1], Wenjie Wang[1], Yongqi Li[2], Fuli Feng[3], See-Kiong Ng[1], and Tat-Seng Chua[1]
[1]National University of Singapore, [2]The Hong Kong Polytechnic University, [3]University of Science and Technology of China
{xylin1028,wenjiewang96,liyongqi0,fulifeng93}@gmail.com,{seekiong,dcscts}@nus.edu.sg

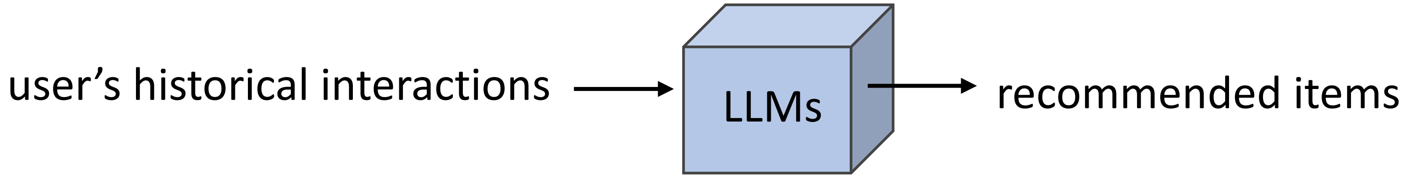
2024.05.13
Presenter: Xinyu Lin
2nd Workshop on Recommendation with Generative Models

Outline

- **Motivation**
- TransRec
- Experiments
- Future Work

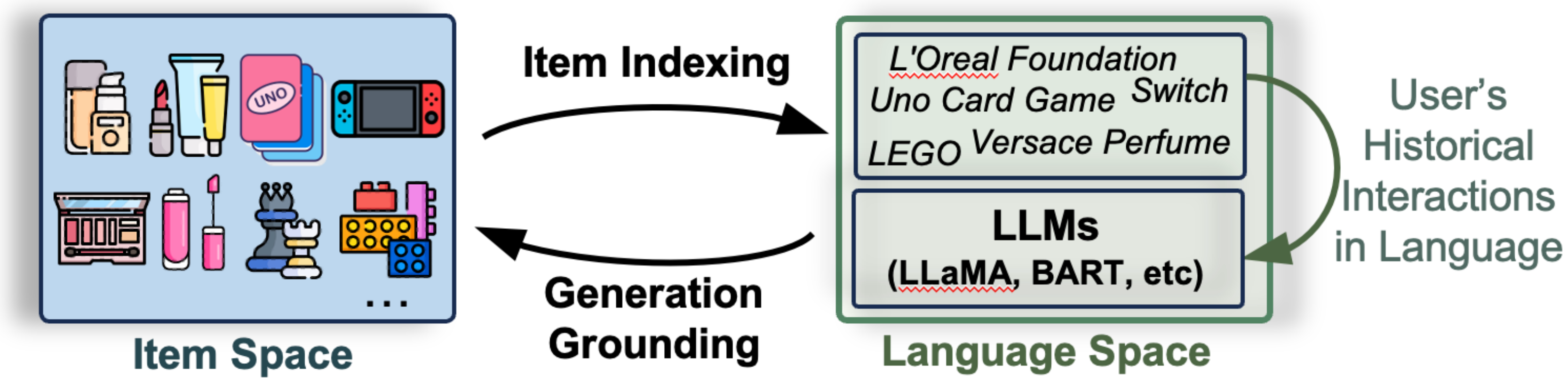
Motivation

- **LLM-based recommenders**



- **Gap between item space and language space**

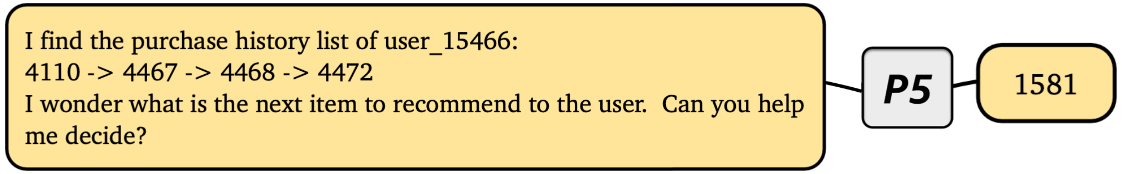
Two key steps



Motivation

❑ Existing work for item indexing

- **ID-based identifier**



- lack of semantics
- poor generalization ability

- **description-based identifier**

Instruction Input	
Instruction:	Given ten movies that the user watched recently, please recommend a new movie that the user likes to the user.
Input:	The user has watched the following movies before: “Traffic (2000)”, “Ocean’s Eleven (2001)”, ... “Fargo (1996)”
Instruction Output	
Output:	“Crouching Tiger, Hidden Dragon (Wu hu zang long) (2000)”

- inadequate distinctiveness
- Inconsistent with interactions

[1] Shijie Geng, et al. Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt & Predict Paradigm (P5). RecSys’22
[2] Bao Keqin et al. " A Bi-Step Grounding Paradigm for Large Language Models in Recommendation Systems“ arXiv 2023.

Motivation

- ❑ Existing work for generation grounding

autoregressive generation

+ exact matching — out-of-corpus recommendation

+ distance-based matching — computationally expensive

Motivation

Limitations

Item Indexing

- *ID-based identifier*: lack of semantics, poor generalization.
- *Description-based identifier*: inadequate distinctiveness

Generation Grounding

- out-of-corpus identifiers

Criteria

Identifier:

- distinctiveness
- semantics

Generation:

- constrained generation [1,2]

depend heavily on first token

- **position-free** constrained generation

[1] Zhixuan Chu, et al. Leveraging Large Language Models for Pre-trained Recommender Systems. 2023 Arxiv
[2] Kai Mei, et al. LightLM: A Lightweight Deep and Narrow Language Model for Generative Recommendation. 2023 Arxiv

Outline

- Motivation
- TransRec
- Experiments
- Future Work

TransRec

- A multi-facet **Transi**tion paradigm for LLM-based **Rec**ommendation
  - **Item Indexing: multi-facet identifier**
  - Instruction data construction
  - Generation Grounding

ID 15826 **(Distinctiveness)**

Title Wilson NBA Series Indoor and Outdoor Basketballs

Attribute Sports

(Semantics)

User's historical interactions in natural language in three facets

ID 15826; 8792; 513; 7382; 9014;

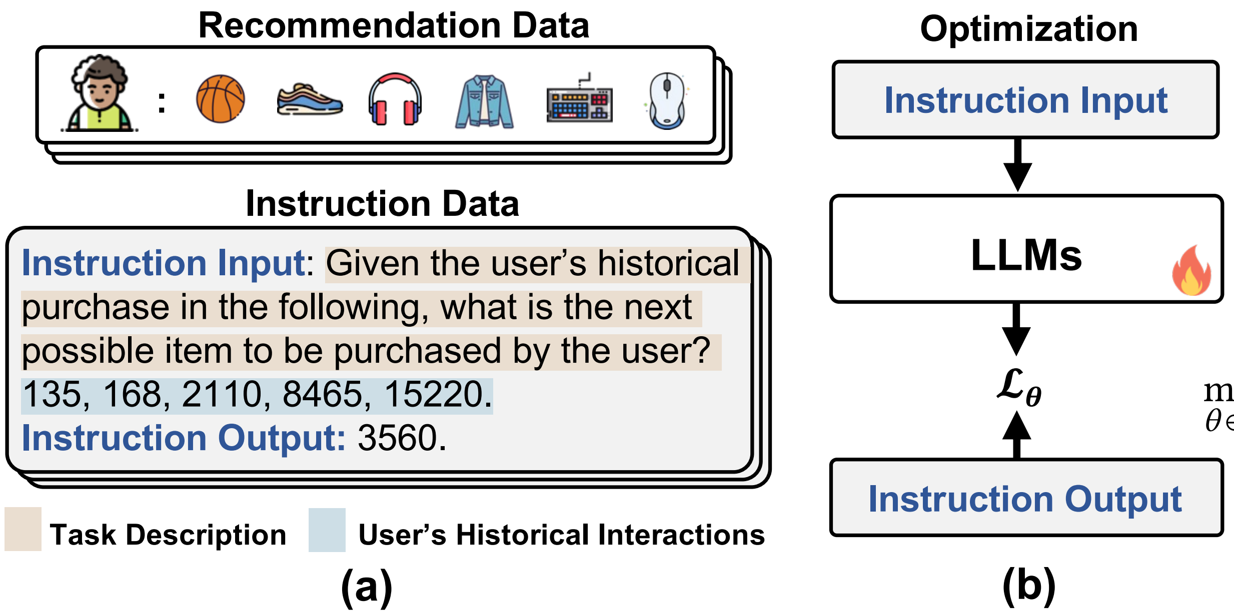
Title Wilson NBA Series Indoor and Outdoor Basketballs; LIT48 Advancourt Base Lifestyle Sneakers; …; Logitech 910-005604 Pebble M350 Wireless Keyboard;

Attribute Sports; Shoe; Headphone & Earphones; …; Electronics;

TransRec

❑ A multi-facet transition paradigm for LLM-based recommendation

- Item indexing
- Instruction data construction
- Generation grounding



$$\min_{\theta \in \Theta} \{\mathcal{L}_\theta = -\sum_{t=1}^{|y|} \log P_\theta(y_t | y_{<t}, x)\},$$

TransRec

❑ A multi-facet transition paradigm for LLM-based recommendation

- Item indexing
- **Instruction data construction**
- Generation grounding

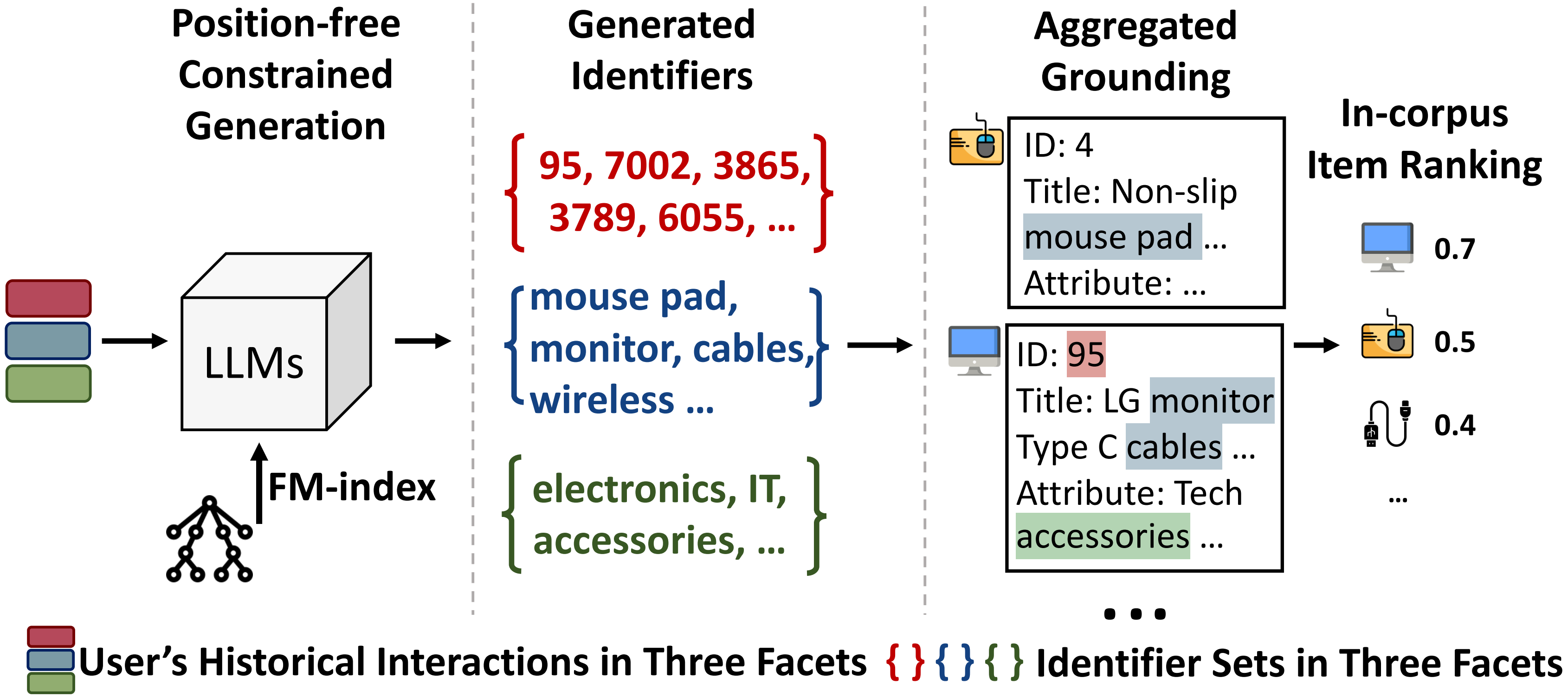
	Instruction Input	Instruction Output
ID	Given the following purchase history of a user, what is the next possible item to be purchased by the user? **15826; 8792; 513; 7382; 9014; || ID || +**	23
Title	Given the following purchase history of a user, what is the next possible item to be purchased by the user? **Wilson NBA Basketballs; Advancourt Sneakers; …; Logitech K270 Wireless Keyboard; || title || +**	Wireless Mouse
Attribute	Given the following attributes of purchase history of a user, what is the next possible attribute of item to be purchased by the user? **Sports; Shoe; Headphone & Earphones; …; Electronics; || attribute || +**	Electronics

substring of "Logitech 118-00735 M774 **Wireless Mouse**"

TransRec

A multi-facet transition paradigm for LLM-based recommendation

- Item indexing
- Instruction data construction
- Generation grounding



Outline

- Motivation
- TransRec
- **Experiments**
- Future Work

Experiments

- ❑ **RQ1:** How does our proposed TransRec perform compared to both traditional and LLM-based recommenders?
- ❑ Full training

	Model	Beauty				Toys				Yelp			
		R@5	R@10	N@5	N@10	R@5	R@10	N@5	N@10	R@5	R@10	N@5	N@10
traditional	MF	0.0294	0.0474	0.0145	0.0191	0.0236	0.0355	0.0153	0.0192	0.0220	0.0381	0.0138	0.0190
	LightGCN	0.0305	0.0511	0.0194	0.0260	0.0322	0.0508	0.0215	0.0275	0.0255	0.0427	0.0163	0.0218
	SASRec	0.0380	0.0588	0.0246	0.0313	0.0470	0.0659	0.0312	0.0373	0.0183	0.0296	0.0116	0.0152
	DCRec	0.0452	0.0635	0.0327	0.0385	0.0498	0.0674	0.0335	0.0406	0.0207	0.0328	0.0115	0.0154
	ACVAE	0.0503	0.0710	0.0356	0.0422	0.0488	0.0679	0.0350	0.0411	0.0211	0.0356	0.0127	0.0174
LLM-based	P5	0.0059	0.0107	0.0033	0.0048	0.0031	0.0069	0.0022	0.0034	0.0039	0.0062	0.0024	0.0031
	SID	0.0350	0.0494	0.0254	0.0301	0.0164	0.0218	0.0120	0.0139	0.0218	0.0332	0.0161	0.0187
	SemID+IID	0.0290	0.0429	0.0200	0.0245	0.0145	0.0260	0.0069	0.0123	0.0196	0.0304	0.0141	0.0160
	CID+IID	0.0484	0.0703	0.0337	0.0412	0.0169	0.0276	0.0104	0.0154	0.0265	0.0417	0.0184	0.0233
	TIGER	0.0377	0.0567	0.0249	0.0310	0.0278	0.0426	0.0176	0.0223	0.0183	0.0298	0.0119	0.0156
	TransRec-B	**0.0504**	**0.0735***	**0.0365***	**0.0450***	**0.0518***	**0.0764***	**0.0360***	**0.0420***	**0.0354***	**0.0457***	**0.0262***	**0.0306***

TransRec-B: TransRec instantiated on BART

- **Superior performance** compared to both **traditional** models and **LLM-based** models.
- **The superiority of TransRec** is attributed to 1) the utilization of multi-facet identifiers to simultaneously achieve semantics and distinctiveness. 2) the constrained and position-free generation for in-corpus item generation and mitigate the over-reliance on initial tokens.

Experiments

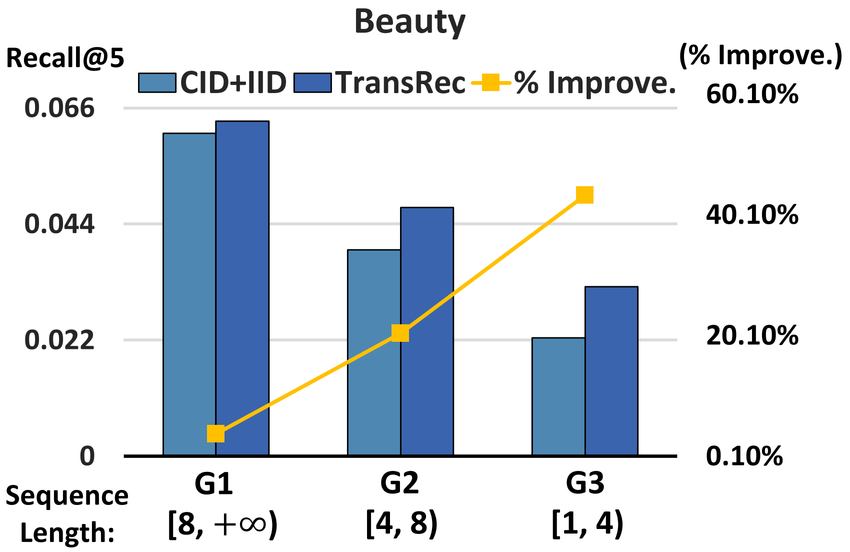
❑ Strong generalization ability

- **Few-shot training**
  - warm- and cold-start testing

N-shot	Model	Warm R@5	Warm N@5	Cold R@5	Cold N@5
1024	LightGCN	0.0205	0.0125	0.0005	0.0003
	ACVAE	0.0098	0.0057	0.0047	0.0026
	CID+IID	0.0100	0.0066	0.0085	0.0071
	TransRec-B	0.0042	0.0028	0.0029	0.0021
	TransRec-L	**0.0141**	**0.0070**	**0.0159**	**0.0097**
2048	LightGCN	0.0186	0.0117	0.0005	0.0004
	ACVAE	0.0229	0.0136	0.0074	0.0044
	CID+IID	0.0150	0.0101	0.0078	0.0062
	TransRec-B	0.0057	0.0031	0.0045	0.0026
	TransRec-L	**0.0194**	**0.0112**	**0.0198**	**0.0124**

* The bold results highlight the superior performance compared to the best LLM-based recommender baseline.

- **User group analysis**
  - from dense users to sparse users



- **Remarkable generalization ability** of LLMs with vase knowledge base, especially on cold-start recommendation under limited data.
- On user side, TransRec significantly **improves the performance of sparse users** with fewer interactions.

Outline

- Motivation
- TransRec
- Experiments
- **Future Work**

Future Work

Following this direction, many promising ideas deserve further exploration:

- although incorporating ID, title, and attribute is effective, it is worthwhile to **automatically construct multi-facet identifiers** to reduce the noises in natural descriptions;
- it is meaningful to devise better strategies for grounding modules, to **effectively combine the ranking scores from different facets**, such as using neural models in an end-to-end learning manner.

Thanks for Your Listening!

🙌 Contact: xylin1028@gmail.com