Multimodal Conditioned Diffusion Model for Recommendation

Haokai Ma[1], Yimeng Yang[1], Lei Meng[1,2*], Ruobing Xie[3], Xiangxu Meng[1]

[1]Shandong University; [2]Shandong Research Institute of Industrial Technology; [3]WeChat, Tencent;

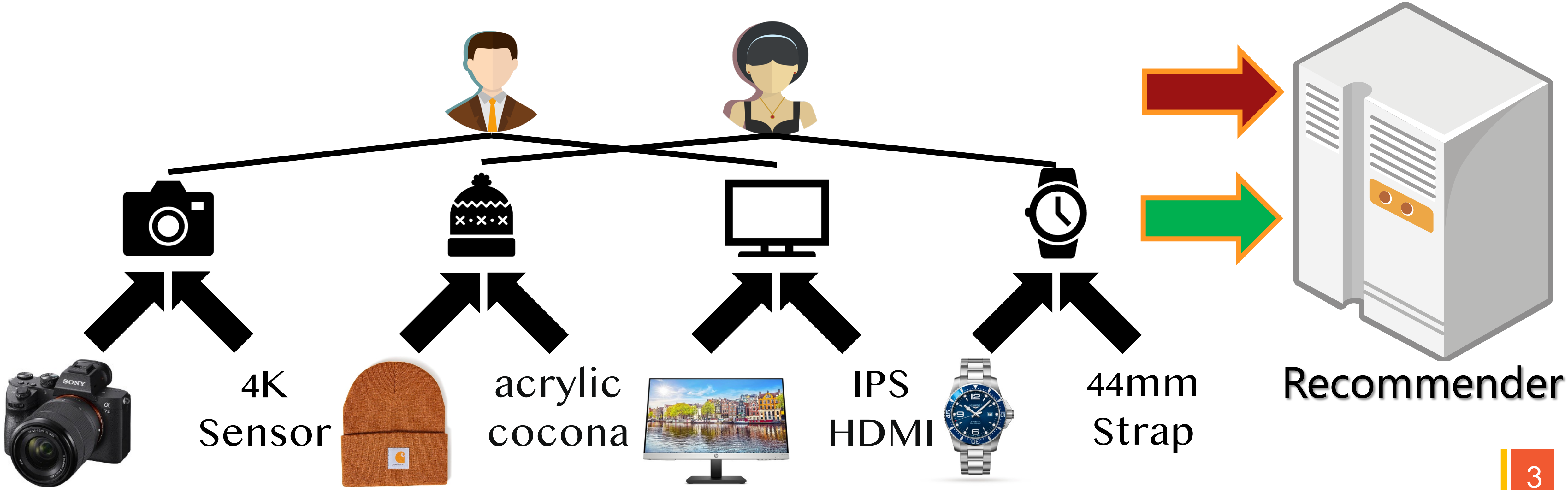
2024.05.13

❖Outline

1. Introduction
    - Multi-modal Recommendation
    - Diffusion Models in Recommendation
2. Motivation Analysis
3. Multimodal Conditioned Diffusion Model for Recommendation framework
4. Experiments
5. Conclusion

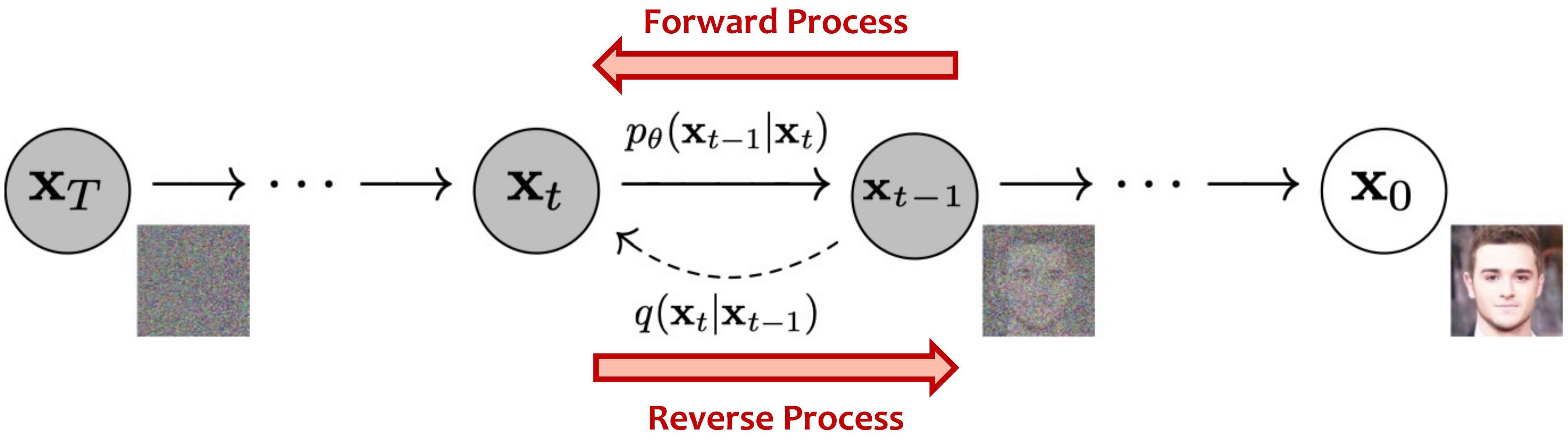
• 1.1 Multimodal Recommendation

- The ubiquity of the extensive corpus and the Matthew effect inevitably engenders the sparsity issue in real-world recommendation systems.
- Multimodal recommendation is proposed to use the additional multi-modal information to enhance the item representation modeling.



• 1.2 Diffusion Model

- Diffusion model is inspired by the non-equilibrium statistical physics and has demonstrated exceptional performance in CV and NLP.
- The classical DM methods typically consists of the Forward Process and the Reverse Process to inject the informative uncertainty into the representation.

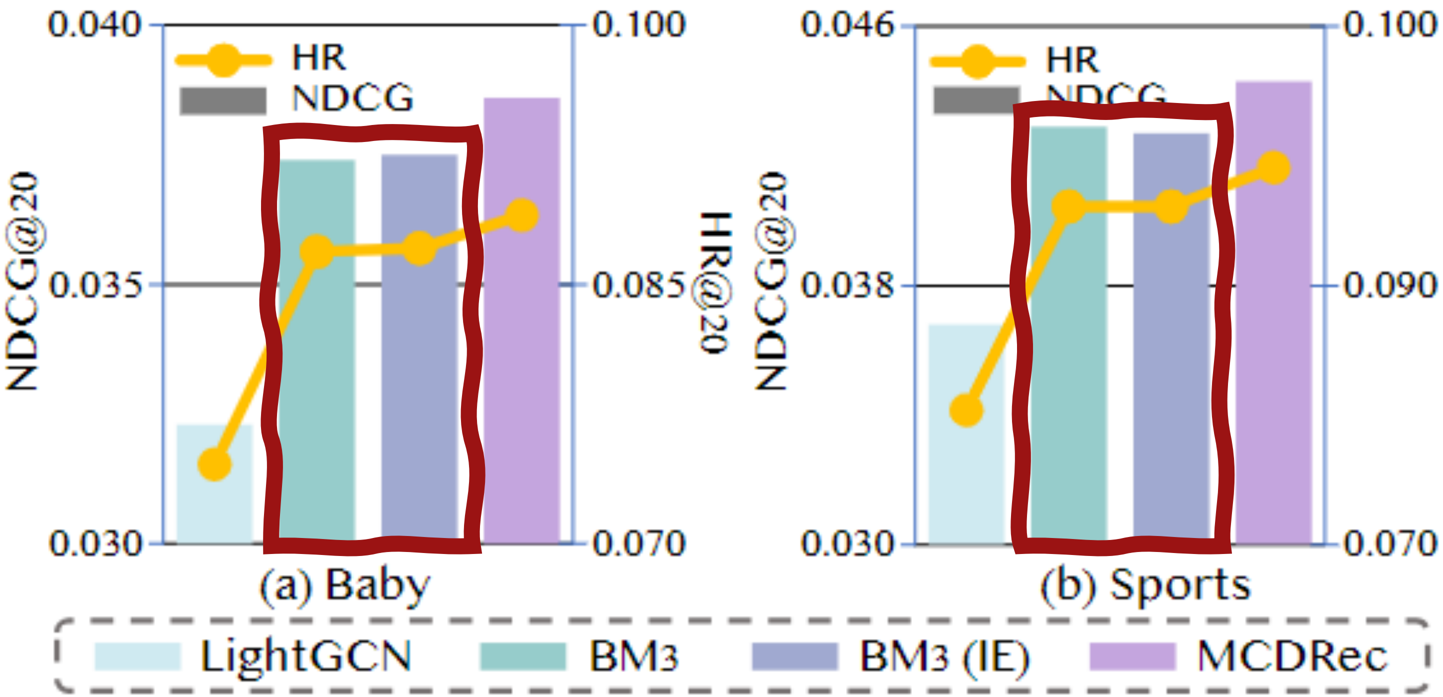


❖Outline

1. Introduction
2. **Motivation Analysis**
    - Challenges of Multi-modal recommenders
3. Multimodal Conditioned Diffusion Model for Recommendation framework
4. Experiments
5. Conclusion

2. Motivation Analysis

- **2.1 Challenges of Multi-modal recommenders**
  - **How to fully leverage the multi-modal knowledge from pretrained features?**

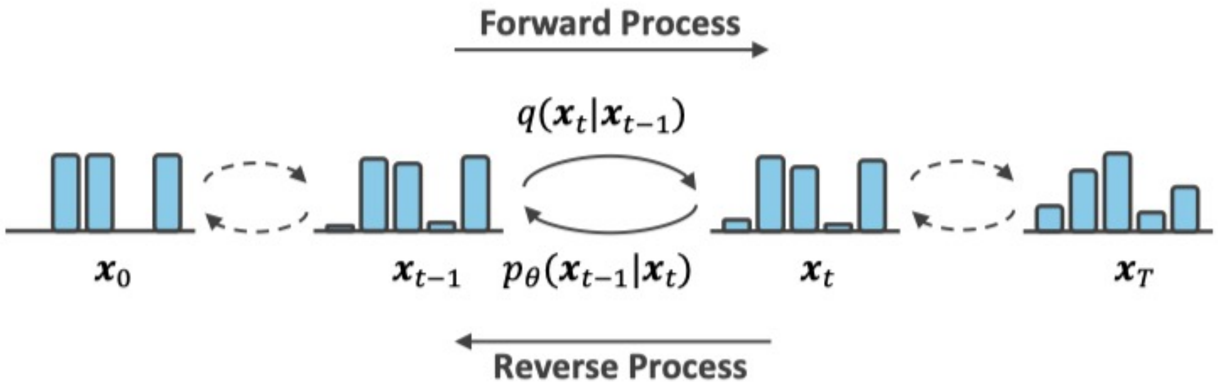


2. Motivation Analysis

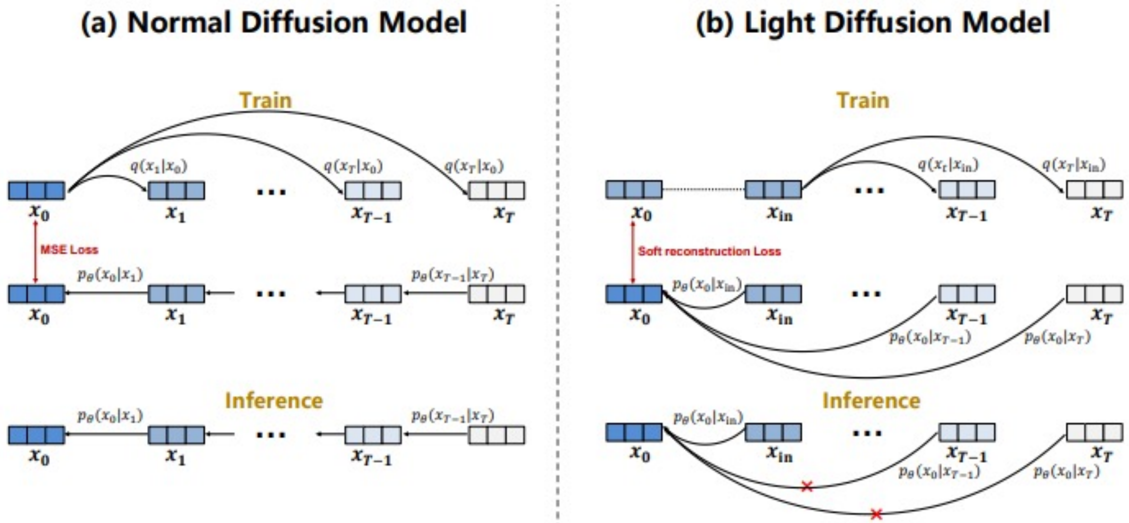
SHANDONG UNIVERSITY

• 2.2 Existing DM-based recommender paradigm

- DiffRec[1] : Reducing the added noises into the forward process to retain globally analogous yet personalized collaborative information in a denoising manner.
- LD4MRec [2] : Leveraging the continuous multi-modal representations for predicting discrete interaction probabilities.



(a) DiffRec[1]



(b) LD4MRec[2]

[1] Wenjie Wang, et al., Diffusion Recommender Model, SIGIR 2023
[2] Penghang Yu, et al., LD4MRec: Simplifying and Powering Diffusion Model for Multimedia Recommendation

❖Outline

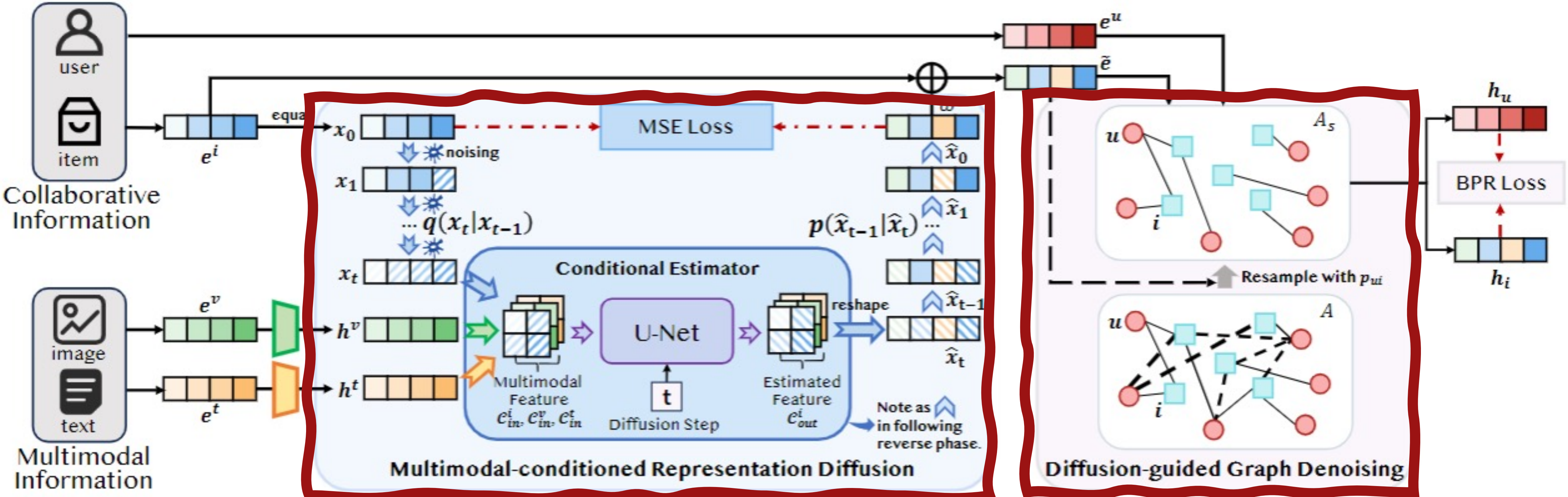
1. Introduction
2. Motivation Analysis
3. Multimodal Conditioned Diffusion Model for Recommendation framework
    - Overall Structure
4. Experiments
5. Conclusion

3. Multimodal Conditioned Diffusion Model for Recommendation framework

山东大学
SHANDONG UNIVERSITY

• 3.2 Overall Structure

- **MRD: Reducing the deviation between modality-aware features and the collaborative information and improve the modeling of item representation**
- **DGD : Denoising the user-item interaction graph accurately through the diffusion-aware item representations**



❖Outline

1. Introduction
2. Motivation Analysis
3. Multimodal Conditioned Diffusion Model for Recommendation framework
4. **Experiments**
    - Effectiveness of MCDRec
    - Ablation Study of MCDRec
    - Visualization of MCDRec
5. Conclusion

SHANDONG UNIVERSITY

- ## 4.1 Effectiveness of MCDRec

Version	Algorithms	Baby						Sports					
		R@5	R@10	R@20	N@5	N@10	N@20	R@5	R@10	R@20	N@5	N@10	N@20
CF-based recommenders	BPR-MF [18]	0.0199	0.0389	0.0442	0.0145	0.0167	0.0200	0.0275	0.0376	0.0435	0.0167	0.0198	0.0208
	LightGCN [8]	0.0282	0.0478	0.0746	0.0190	0.0250	0.0323	0.0342	0.0524	0.0804	0.0236	0.0296	0.0368
Multi-modal recommenders	MMGCN [28]	0.0253	0.0403	0.0646	0.0170	0.0219	0.0281	0.0231	0.0371	0.0631	0.0150	0.0196	0.0263
	SLMRec [23]	0.0438	0.0486	0.0741	0.0216	0.0271	0.0337	0.0418	0.0650	0.0967	0.0285	0.0361	0.0443
	DualGNN [26]	0.0324	0.0506	0.0799	0.0213	0.0274	0.0350	0.0348	0.0579	0.0892	0.0238	0.0316	0.0402
	LATTICE [32]	0.0349	0.0542	0.0845	0.0228	0.0292	0.037	0.0395	0.0625	0.0958	0.0262	0.0337	0.0423
	BM3 [36]	0.0326	0.0535	0.0869	0.0219	0.0288	0.0374	0.0401	0.0627	0.0961	0.0269	0.0343	0.0429
	MCDRec (BM3)	0.0355	0.0566	0.0890	0.0242	0.0306	0.0386	0.0419	0.0654	0.0991	0.0279	0.0355	0.0443
	Improvement	8.90%	5.79%	2.42%	10.50%	6.25%	3.21%.	4.49%	4.31%	3.12%	3.72%	3.50%	3.26%
	FREEDOM [35]	0.0376	0.0624	0.0985	0.0243	0.0324	0.0416	0.0455	0.0713	0.1075	0.0299	0.0384	0.0477
	MCDRec (FREEDOM)	**0.0397**	**0.0644**	**0.1013**	**0.0263**	**0.0343**	**0.0438**	**0.0466**	**0.0737**	**0.1100**	**0.0306**	**0.0392**	**0.0488**
	Improvement	5.59%	3.21%	2.84%	8.23%	5.86%	5.29%	2.42%	3.37%	2.33%	2.34%	2.08%	2.31%

MCDRec significantly outperforms all the baselines across all metrics and datasets.

The improvements are larger with smaller k

MCDRec achieves the most substantial improvement over BM3 and obtains peak results over FREEDOM.

4. Experiments

SHANDONG UNIVERSITY

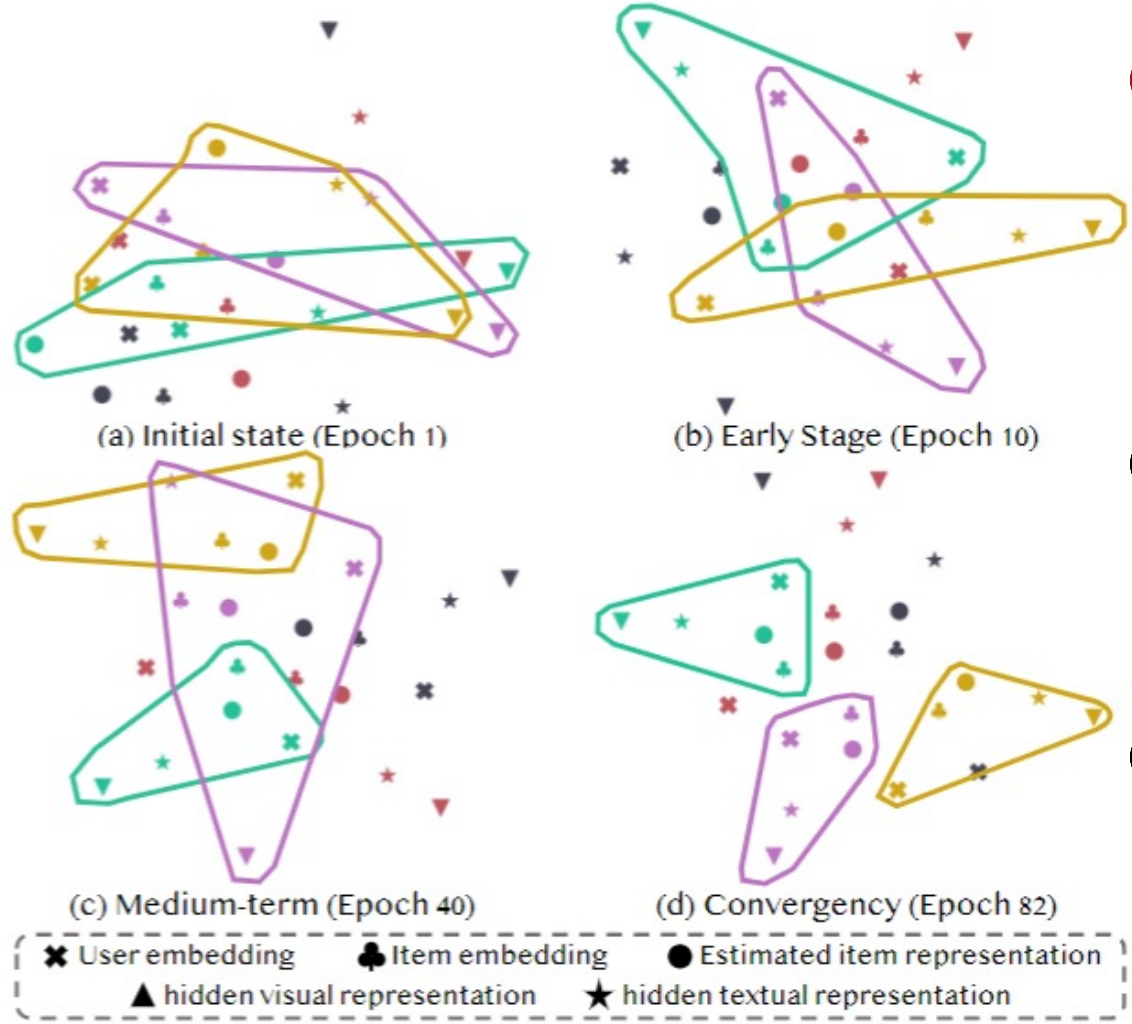
- **4.2 Ablation Study of MCDRec**

Datasets	Versions	R@5	R@10	R@20	N@5	N@10	N@20
Baby	BM3	0.0326	0.0535	0.0869	0.0219	0.0288	0.0374
	BM3+MK	0.0331	0.0541	0.0848	0.0220	0.0289	0.0368
	BM3+MRD	0.0348	0.0558	0.0886	0.0230	0.0297	0.0380
	BM3+DGD	0.0335	0.0547	0.0875	0.0226	0.0291	0.0375
	MCDRec (BM3)	**0.0355**	**0.0566**	**0.0890**	**0.0242**	**0.0306**	**0.0386**
Sport	BM3	0.0401	0.0627	0.0961	0.0269	0.0343	0.0429
	BM3+MK	0.0403	0.0620	0.0946	0.0268	0.0340	0.0424
	BM3+MRD	0.0411	0.0641	0.0988	0.0275	0.0350	0.0436
	BM3+DGD	0.0409	0.0644	0.0983	0.0276	0.0350	0.0435
	MCDRec (BM3)	**0.0419**	**0.0654**	**0.0991**	**0.0279**	**0.0355**	**0.0443**

[illegible]

[illegible]

• 4.3 Visualization of MCDRec



① **In the initial state**, the **intrinsic aggregation** of the pre-trained representations from **the same modality** poses a challenge for the subsequent recommender.

② With the training of MRD, we **progressively** achieve **consistent modeling** of multi-modal preferences from **the same user**.

③ Multimodal item representations from the same user exhibit **the significant clustering distributions** in the end.

❖Outline

1. Introduction
2. Motivation Analysis
3. Real Hard Negative Sampling framework
4. Experiments
5. Conclusion

5. Conclusion

SHANDONG UNIVERSITY

Conclusion

- We propose a novel Multimodal Conditioned Diffusion Model for Recommendation (MCDRec), which is able to co-model multi-modal guidance and diffusion guidance to enhance the performance of existing multi-modal recommenders.
- The proposed MRD and DGD in MCDRec are effective, model-agnostic and precisely capture users' modality-aware personalized preferences.
- MCDRec achieves significant and consistent improvements on different datasets and base multimodal recommenders.

Future Direction

- Exploit the fine-grained modeling of multi-modal representations in DM.
- Validate its effectiveness in more challenging scenarios such as multimodal sequential recommendation and cross-domain multimodal recommendation.

THANKS!

Yimeng Yang
y_yimeng@mail.sdu.edu.cn

山东大学软件学院
SCHOOL OF SOFTWARE, SHANDONG UNIVERSITY
— SINCE 2001 —

Tencent 腾讯