Multi-view Hypergraph-based Contrastive Learning Model for Cold-Start Micro-video Recommendation

Authors: Sisuo LYU, Xiuze ZHOU, Xuming HU
Affiliation: The Hong Kong University of Science and Technology (Guangzhou)

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

INTRODUCTION

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

RESEARCH BACKGROUND:

1. **Rapid Growth** of micro-video platforms (TikTok, Kwai) driving demand for personalized recommendations
2. **Rich Multi-modal Features** in micro-videos:
   - Textual metadata (titles, descriptions)
   - Visual elements (cover images)
   - Dynamic video content
3. **User Engagement** heavily influenced by these diverse features
4. **Unique Challenges** compared to traditional content recommendation

CHALLENGES:

1. **Over-smoothing Problem**
2. **Long-tail Distribution**
3. **Cold-start Scenarios**
4. **Information Underutilization**

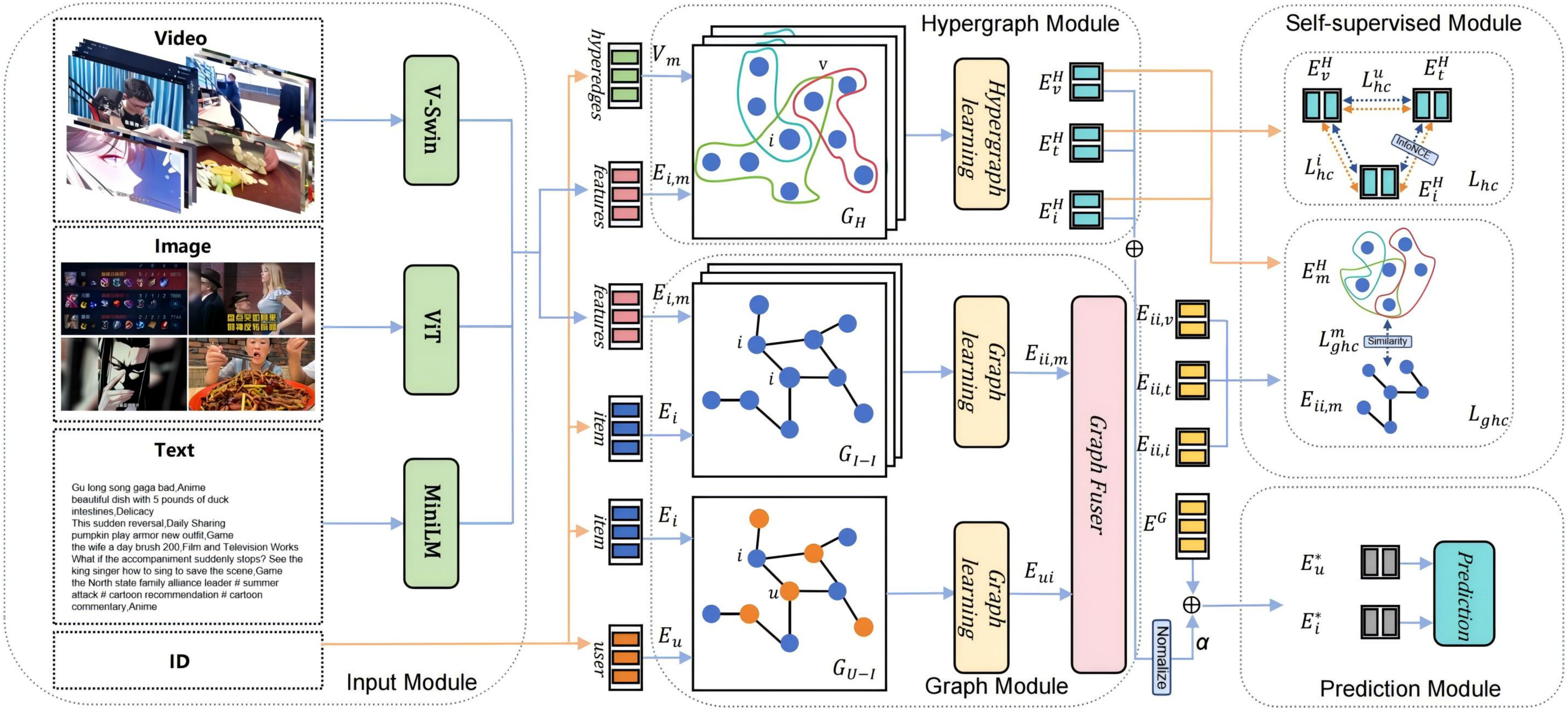
INTRODUCTION

OUR MOTIVATION:

1. **Beyond Traditional Graphs:**
   Need richer structures to capture higher-order relationships
2. **Multi-view Learning:**
   Different perspectives provide complementary information
3. **Self-supervised Signals:**
   Additional supervision can compensate for sparse interactions
4. **Cold-start Focus:**
   Improving recommendations for new/rarely viewed content benefits both creators and viewers

PROPOSED METHOD

MHCR:



PROPOSED METHOD

MULTI-VIEW FEATURE EXTRACTION:

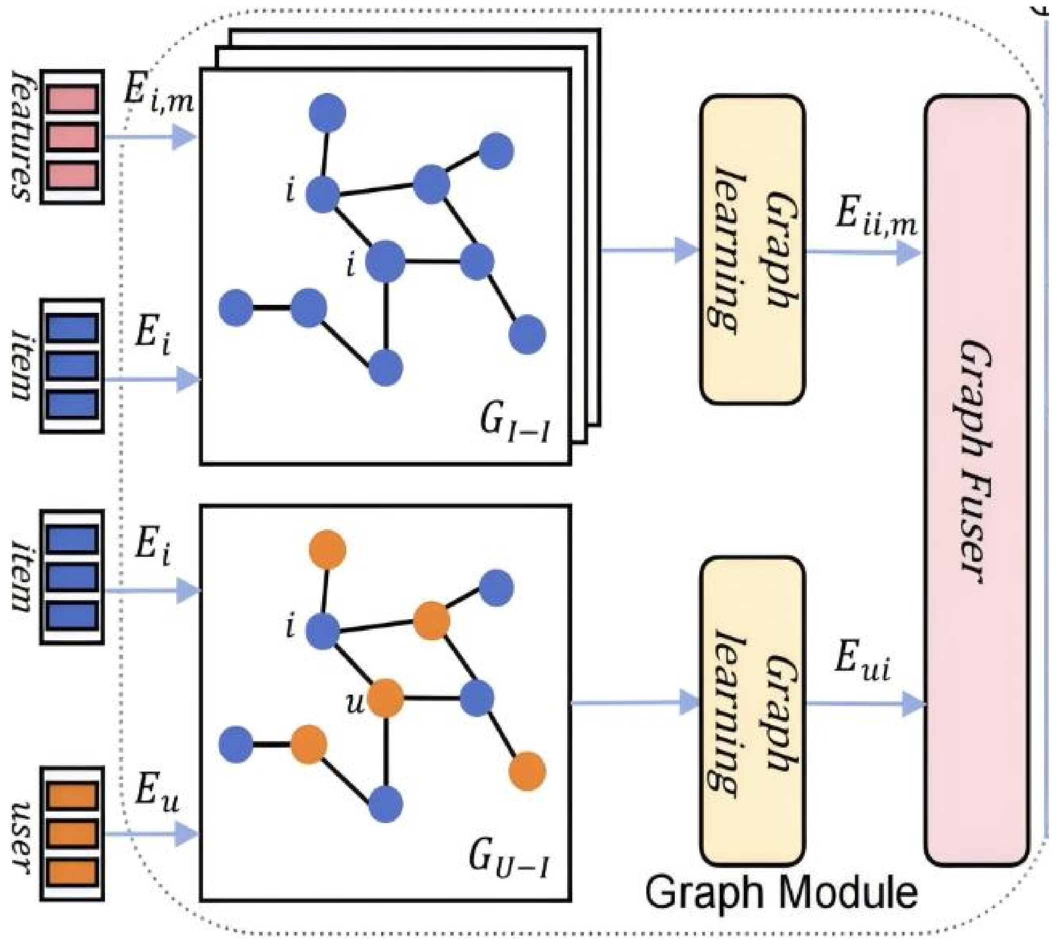
1. User-Item Graph:
   - Captures direct interaction patterns
   - Formula:

$$E_{ui}^{(l)} = \sum_{i \in N_u} \frac{1}{\sqrt{|N_u| \cdot |N_i|}} E_{ui}^{(l-1)},$$

   - Aggregates high-order collaborative signals
2. Item-Item Graph:
   - Models relationships between items
   - Modality-specific correlations
   - Affinity calculation:

$$s_{a,b}^m = \frac{(e_a^m)^{\mathrm{T}} \cdot e_b^m}{\|e_a^m\| \|e_b^m\|},$$

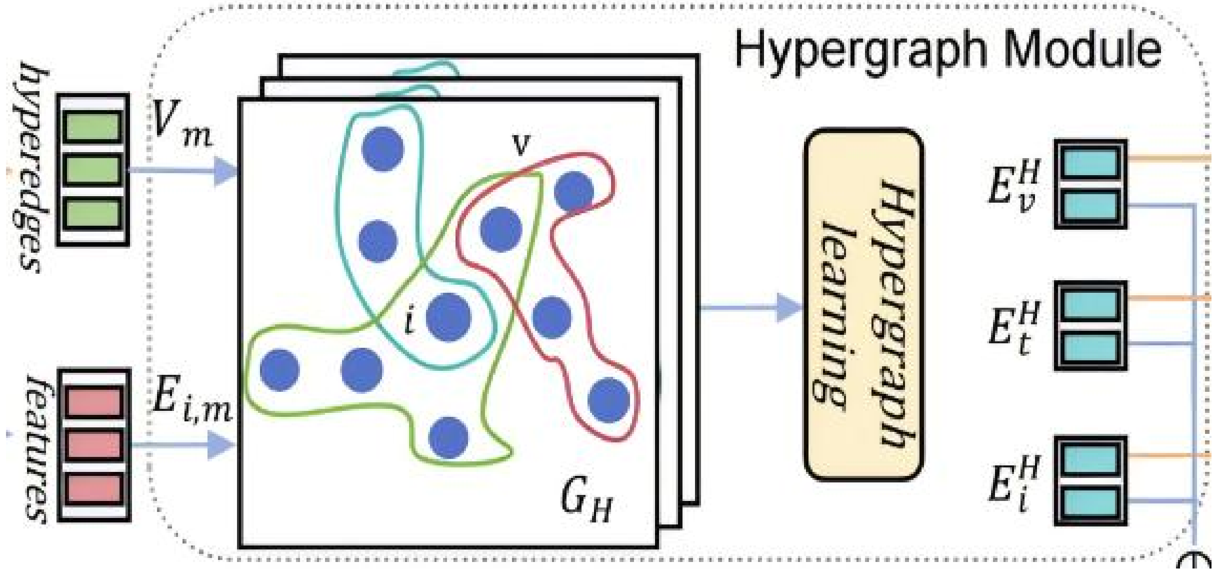
   - KNN sparsification for efficiency



PROPOSED METHOD

HYPERGRAPH MODULE:

1. Beyond Pairwise Relationships:
   - Hyperedges connect multiple nodes simultaneously
   - Captures complex interaction patterns
2. Learnable Hyperedge Embeddings:
   - $$H_i^m = E_i^m \cdot V_m^\top$$
   - $$H_u^m = X_u \cdot (H_i^m)^\top$$
3. Global Information Transfer:
   1. Message passing via hyperedges
   2. Facilitates information flow between disparate nodes



PROPOSED METHOD

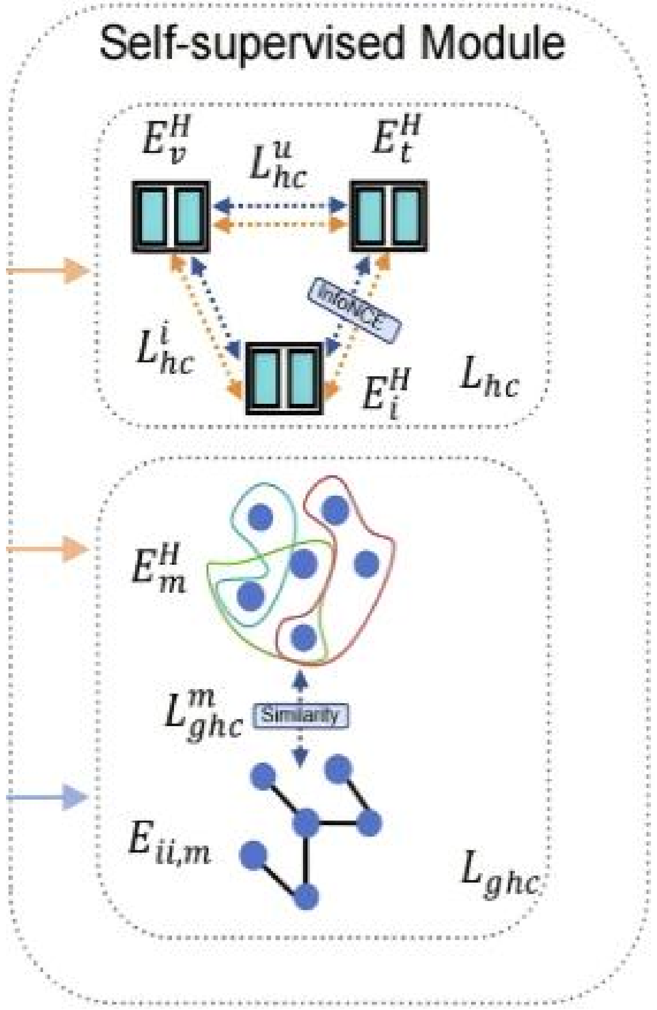
SELF-SUPERVISED LEARNING:

1. Cross-modal Hypergraph Contrastive Learning:
    1. Ensures effective fusion across modalities
    2. Pulls same entities closer in different modality spaces
    3. Loss function:

$$L_{hc} = \sum_{x \in U \cup I} -\log \frac{\sum_m \exp\left(\frac{s(E_x^{m,H}, E_x^{m',H})}{\tau}\right)}{\sum_{x' \in U \cup I} \sum_m \exp\left(\frac{s(E_x^{m,H}, E_x^{t,H})}{\tau}\right)},$$

2. Graph-Hypergraph Contrastive Learning:
    1. Aligns representations from different structural views
    2. Enhances consistency between graph and hypergraph embeddings
    3. Loss function:

$$L_{ghc} = \sum \left(-\log \frac{\exp\left(\frac{s(E^G, E^H)}{\tau}\right)}{\exp\left(\frac{s(E^G, E^H)}{\tau}\right) + \sum \exp\left(\frac{s(E^G, E^H)}{\tau}\right)} \right)$$



EXPERIMENTAL SETUP

DATASETS:

1. MicroLens-50K: 50,000 users, 19,220 items, 359,708 interactions
2. MicroLens-100K: 100,000 users, 19,738 items, 719,405 interactions

EVALUATION METRICS:

1. Recall@n ($n \in \{10, 20\}$)
2. NDCG@n ($n \in \{10, 20\}$)

BASELINES:

1. General: YouTube, VBPR
2. Graph-based: LightGCN, LayerGCN
3. Multimodal: MMGCN, GRCN, BM3, Freedom, MGCN

EXPERIMENTAL SETUP

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

OVERALL PERFORMANCE RESULTS:

TABLE I
PERFORMANCE METRICS OF DIFFERENT MODELS ON TWO MICROLENS DATASETS

Dataset	Metric	YouTube	VBPR	MMGCN	LightGCN	GRCN	LayerGCN	BM3	Freedom	MGCN	MHCR
MicroLens-50K	R@10	0.0375	0.0544	0.0403	0.0365	0.0631	0.0627	0.0565	0.0656	0.0708	**0.0736**
	R@20	0.0632	0.0888	0.067	0.0534	0.0982	0.0994	0.0918	0.1028	0.1089	**0.1102**
	N@10	0.0178	0.0273	0.0197	0.0284	0.0328	0.032	0.0281	0.0334	0.0363	**0.0383**
	N@20	0.0245	0.0361	0.0264	0.0345	0.0415	0.0414	0.0372	0.0429	0.0459	**0.0477**
MicroLens-100K	R@10	0.0392	0.0624	0.0405	0.0388	0.0682	0.0730	0.0601	0.0654	0.0717	**0.0798**
	R@20	0.0648	0.1002	0.0678	0.056	0.1057	0.1120	0.0975	0.1016	0.1096	**0.1187**
	N@10	0.0188	0.0314	0.0202	0.0306	0.0353	0.0382	0.0305	0.0337	0.0371	**0.042**
	N@20	0.0252	0.0410	0.0271	0.0367	0.0448	0.0480	0.0401	0.0431	0.0467	**0.0519**

MICROLENS-50K IMPROVEMENTS:

- Recall@10: 0.0736 (+3.96% vs MGCN)
- NDCG@10: 0.0383 (+5.51% vs MGCN)

MICROLENS-100K IMPROVEMENTS:

- Recall@10: 0.0798 (+11.30% vs MGCN)
- NDCG@10: 0.0420 (+13.19% vs MGCN)

EXPERIMENTAL SETUP

COLD-START PERFORMANCE:

TABLE II
PERFORMANCE COMPARISON OF DIFFERENT MODELS FOR COLD-START USERS ON TWO MICROLENS DATASETS

Dataset	Metric	GRCN	LayerGCN	BM3	Freedom	MGCN	MHCR
MicroLens-50K	R@10	0.0521	0.0527	0.0471	0.053	0.0588	**0.0616**
	R@20	0.081	0.0826	0.0789	0.0839	0.0912	**0.0937**
	N@10	0.0267	0.0267	0.0237	0.0267	0.0299	**0.0321**
	N@20	0.034	0.0341	0.0316	0.0344	0.038	**0.0396**
MicroLens-100K	R@10	0.0562	0.0588	0.0485	0.0544	0.0605	**0.0655**
	R@20	0.0874	0.092	0.0807	0.0882	0.0937	**0.0984**
	N@10	0.0288	0.0299	0.024	0.0275	0.0311	**0.0342**
	N@20	0.0366	0.0384	0.0321	0.0359	0.0396	**0.0423**

MICROLENS-50K IMPROVEMENTS:

- Recall@10: 0.0616 (+4.8% vs MGCN)
- NDCG@10: 0.0321 (+7.3% vs MGCN)

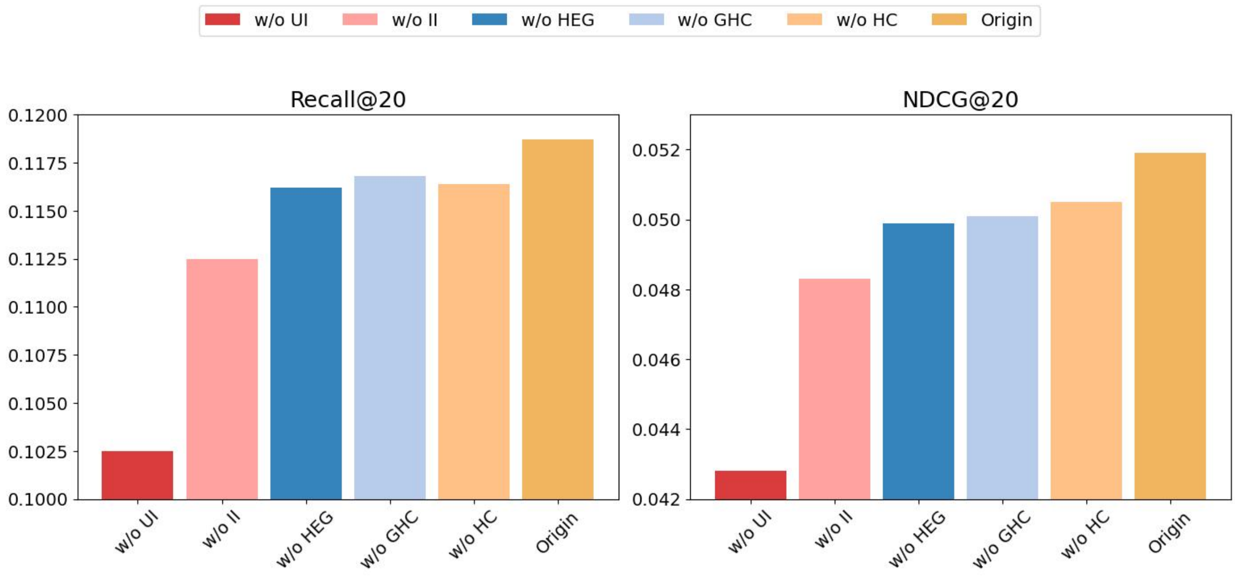
MICROLENS-100K IMPROVEMENTS:

- Recall@20: 0.0984 (+5.0% vs MGCN)
- NDCG@20: 0.0423 (+6.8% vs MGCN)

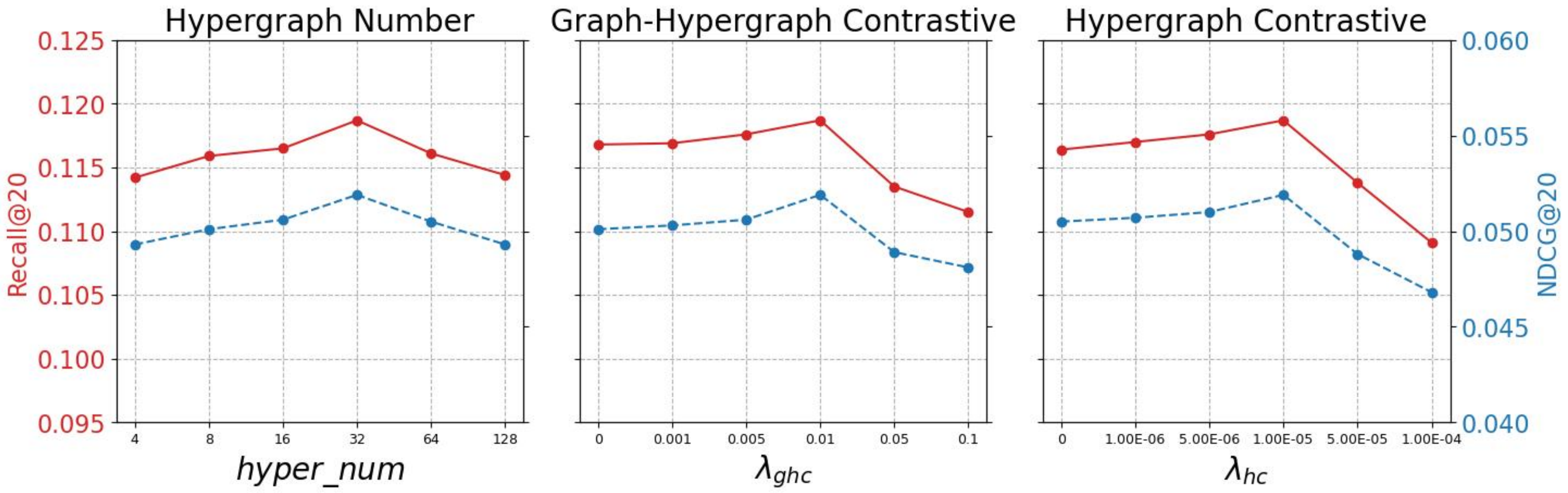
EXPERIMENTAL SETUP

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

COMPONENT CONTRIBUTIONS:



PARAMETER SENSITIVITY:



CONCLUSION

KEY CONTRIBUTIONS:

1. Novel integration of hypergraphs with contrastive learning
2. Multi-view feature extraction enhances representation quality
3. Significant performance gains, especially in cold-start scenarios

BROADER IMPACT:

1. Framework applicable to other recommendation domains
2. Helps address content discovery challenges for new creators
3. Enhances user experience with more diverse recommendations

FUTURE DIRECTIONS:

1. Dynamic hypergraph construction
2. Temporal modeling of evolving user interests

香港科技大学(广州)
THE HONG KONG
UNIVERSITY OF SCIENCE AND
TECHNOLOGY (GUANGZHOU)

THANK YOU!