User-Item Matching and Auto-encoders for Cold-start Next-Item Recommendation

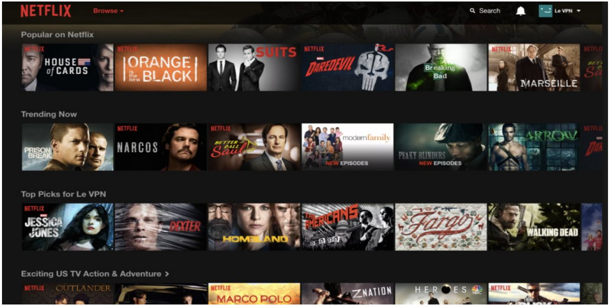
Michael K. Ng
Department of Mathematics, The University of Hong Kong
Email: michael.ng@hku.hk

The 2021 IEEE International Conference on e-Business Engineering (ICEBE)

Research supported in part by HKRGC GRF grants and HKU-TCL Joint Research Center for AI

Recommendation System

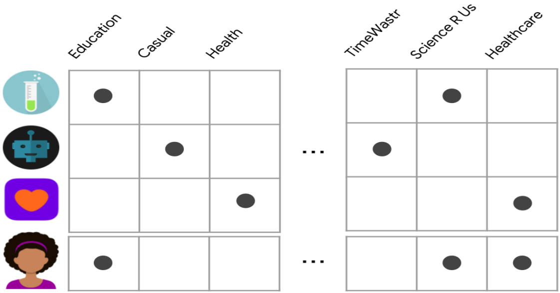
- Discover relevant information of interest for users
- Apply to plenty of scenarios such as online shopping, movie recommendation, etc.



citation from https://vwo.com/

Content-based Filtering

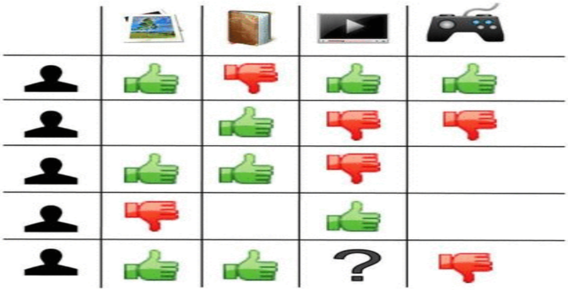
Content-based filtering uses item features to recommend other items similar to what the user likes, based on their previous actions or explicit feedback.



citation from http://developers.google.com/

Collaborative Filtering

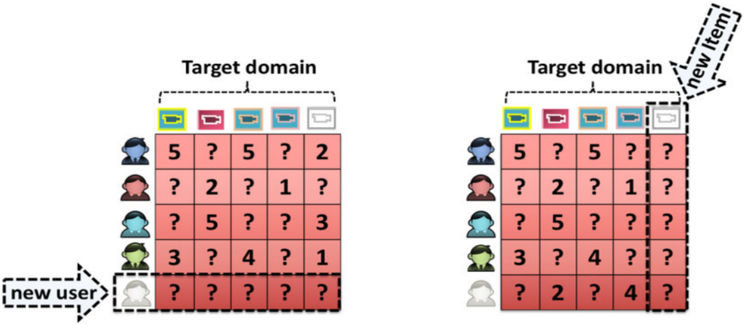
- Current recommendation system approaches are mainly based on collaborative filtering.
- Collaborative filtering uses similarities between users and items simultaneously to provide recommendations.



citation from https://en.wikipedia.org/

Key Issue

- One of the challenging issues in recommendation systems is the cold-start item problem.
- Recommendations are needed for items that no one has yet interacted with.



citation from DOI:10.13140/RG.2.2.27407.02725

Cold-Start Item Recommendation

1. Y. Meng, X. Yan, W. Liu, H. Wu, and J. Cheng, "Wasserstein collaborative filtering for item cold-start recommendation," in Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization, 2020, pp. 318-322.
2. M. Saveski and A. Mantrach, "Item cold-start recommendations: learning local collective embeddings," in Proceedings of the 8th ACM Conference on Recommender systems, 2014, pp. 89-96.
3. O. Barkan, A. Caciularu, I. Rejwan, O. Katz, J. Weill, I. Malkiel, and N. Koenigstein, "Cold item recommendations via hierarchical item2vec," in IEEE International Conference on Data Mining (ICDM). IEEE, 2020, pp. 912-917.
4. S. A. P. Parambath and S. Chawla, "Simple and effective neural-free soft-cluster embeddings for item cold-start recommendations," Data Mining and Knowledge Discovery, vol. 34, no. 5, pp. 1560-1588, 2020.
5. J. Misztal-Radecka, B. Indurkhya, and A. Smywinski-Pohl, "Meta-user2vec model for addressing the user and item cold-start problem in recommender systems," User Modeling and User-Adapted Interaction, pp. 1-26, 2020.

Cold-Start Item Recommendation

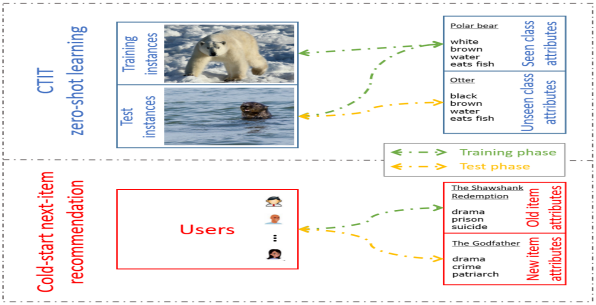
- In [1], Meng et al. proposed to evaluate the similarity between the warm items and the cold-start ones, then a user's preference on cold-start items can be achieved by minimizing the Wasserstein distance between the distributions over these two types of items.
- In [2], Saveski et al. proposed a matrix factorization method named Local Collective Embeddings, which exploits item attributes and user historical preferences while enforcing the manifold structure exhibited by the collective embeddings.
- In [5], Misztal-Radecka et al. proposed to extend the Doc2Vec model by building embeddings of users' metadata labels along with item representations.

Limitation

- Current methods mainly aim to discover a latent space to bridge the users and cold-start items.
- Current methods barely consider preserving either the interaction relationships between users and their historical selected items or the information of original input data.
- Important information may lose during the model training.

Zero-shot Learning

Cold-start recommendation is naturally similar to the zero-shot learning, which predicts labels for instances whose classes are unseen during training.



Zero-shot Learning

- Given the attributes of training classes and training instances with their labels, the goal is to recognize the classes that have no overlap with the training classes of the test instances.
- In the cold-start item recommendation, given training items based on users' preferences, we aim to recommend new items to users who are different from training items.
- In the cold-start user recommendation, given training users with their preferences on items, we aim to recommend items to new users who are different from training users.

Cold-start Next-item Recommendation

How to recommend a user his/her next preferred item, which has not yet been interacted with by any users within the system ?

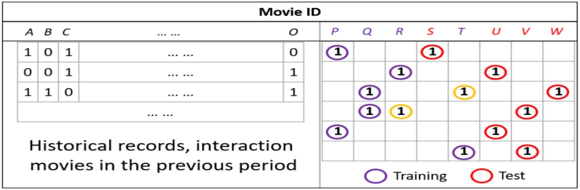


Fig. 2. Split of the training and testing data for movie recommendation. Movies A, B, . . ., and O are the historical movies. The movies circled by purple color, *i.e.*, movies P, Q, R, and T, are training movies (*i.e.*, seen classes in ZSL), and the movies surrounded by red circles *i.e.*, movies S, U, V, and W, are test cold-start movies (*i.e.*, unseen classes in ZSL).

The Proposed Method

- To exploit a hypergraph to learn latent user embedding based on the interaction relationships between users and the historical items
- To preserve the interaction relationships between users and the historical items and the information of original items
- To capture the relationships between users and their training items

Cold-start Next-item Recommendation

N_U	number of users	$\mathbf{U}$	user features
N_M	number of items	$\mathbf{M}$	item features
N_M^{hist}	number of historical items	$\mathbf{M}^{hist}$	historical item features
N_M^{train}	number of training items	$\mathbf{M}^{train}$	training item features
N_M^{test}	number of test items	$\mathbf{M}^{test}$	test item features
$\mathbf{H}$	historical interactions	$\mathbf{Y}$	training labels

Definition 2 (Cold-start Next-item Recommendation). In this paper, $\mathcal{M}^{train}$ and $\mathcal{M}^{test}$ are the sets of training and test items, respectively, and $\mathcal{M}^{train} \cap \mathcal{M}^{test} = \emptyset$. Given user data $\{\mathbf{U}, \mathbf{Y}\}$ interacting with the training items $\mathcal{M}^{train}$, cold-start next-item recommendation aim-s to learn a model that can recommend the next items belonging to the test items $\mathcal{M}^{test}$ to users. Specifically, user data $\{\mathbf{U}, \mathbf{Y}\}$, item attributes $\mathbf{M}$ are used in model learning.

Cold-start Next-item Recommendation

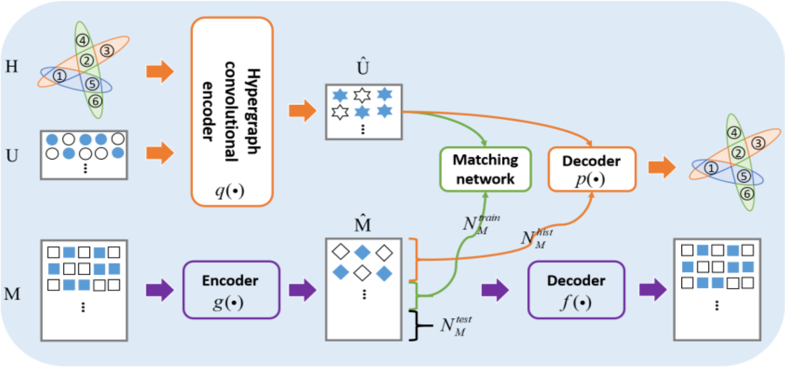
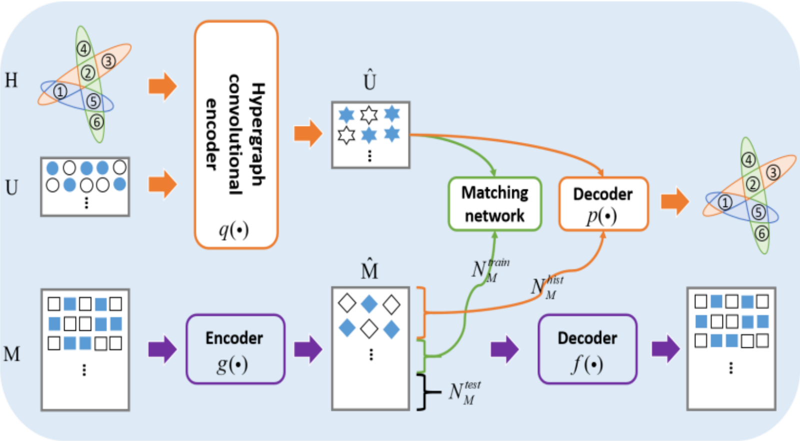


Fig. 4. An illustration of the three components plotting by different colors of the proposed UIMA. The progress in purple is the auto-encoder for learning item embedding, the one in orange presents the auto-encoder for learning user embedding, the one in green is the matching network.

Cold-start Next-item Recommendation

To learn the latent informative embeddings for both users and items from the sight of zero-shot learning, so as to evaluate the distances between users and cold-start items for recommendations.

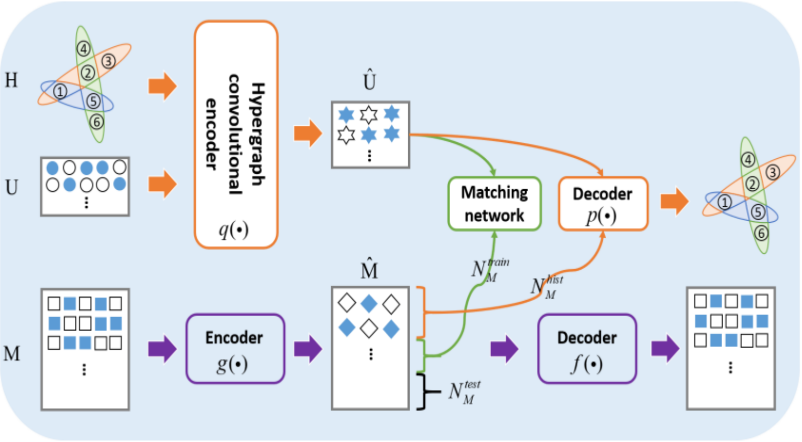
Cold-start Next-item Recommendation



An auto-encoder network to achieve latent item embedding:

$$Loss_{g,f}(M, f(g(M)))$$

Cold-start Next-item Recommendation



An auto-encoder network to preserve the interaction relationships between users and the historical items so as to obtain latent user embedding

Hypergraph

	e1	e2	e3
v1	1	0	0
v2	1	0	1
v3	0	1	1
v4	1	1	1
v5	0	1	0

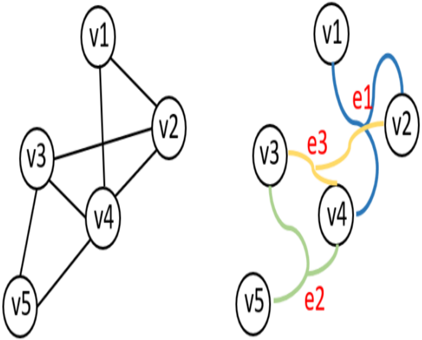


Fig. 1. A comparison between a simple graph and a hypergraph. Left: a set of movies {v1, v2, v3, v4, v5} and a set of actors {e1, e2, e3}. Middle: each edge in the simple graph can only connect two vertices. Right: each hyperedge in the hypergraph can connect more than two vertices.

Hypergraph

- The historical user interactions H is indeed an incidence matrix whose values equal to 1 or 0, which represent the relationship between users and historical item.
- H can be regarded as a hypergraph whose hyperedges are item-set.
- H and the user representation U can form inputs of a graph network, where we develop hypergraph convolutional operations to learn the informative user embedding.
- The propagation effective information between nodes can be proceeded by fully exploiting the complex relationships and local clustering structures.

Hypergraph

- Hypergraph convolution operation:

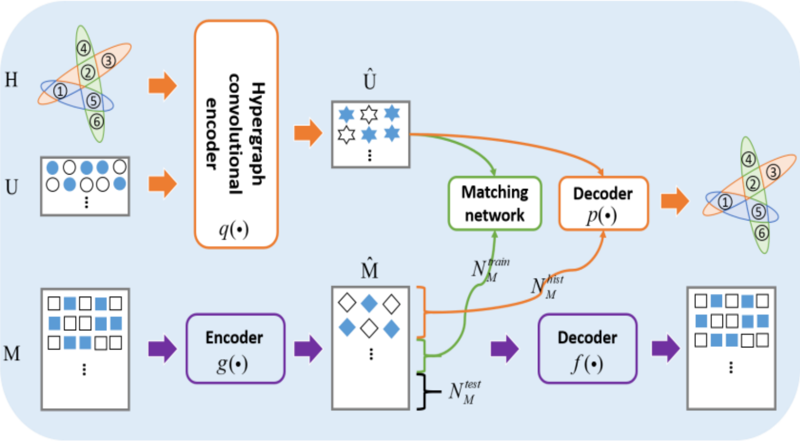
$$U^{(l+1)} = \sigma(L^{-1/2}H\Omega H^T L^{-1/2}U^{(l)}W)$$

  L is the diagonal vertex degree matrix and B is the diagonal hyperedge degree matrix.

- Together with historical item embedding, we can decode the hypergraph indicating the relationships between users and historical items:

$$Loss_{q,p}(H, p(\hat{H}|q(\hat{U} \mid U, H), \hat{M}^{hist}))$$

Cold-start Next-item Recommendation



A matching network to explore the relationships between users and their training labels

Matching Network

The training labels Y indicate the next interacted items for users. A user-item matching network to explore the relationships between users and their training items:

$$Loss_{q,g}(q(\hat{U} \mid U, H), Yg(M^{train})),$$

where $Loss$ can be the mean squared error.

Experimental Results

TABLE 2
Statistical information on the used data sets.

	Duration	History	N_U	N_M^{hist}	N_M^{train}	N_M^{test}
MovieLens	1	2000-2002	385	4830	146	149
	2	2001-2003	452	5838	141	166
	3	2002-2004	493	6939	168	157
	4	2003-2005	614	7441	197	131
LastFM	1	2005-2007	154	4280	118	140
	2	2006-2008	191	6177	163	184
	3	2007-2009	186	6758	157	175
	4	2008-2010	101	6500	99	99
Delicious	1	1-3, 2010	316	4997	305	314
	2	2-4, 2010	326	5521	305	320
	3	3-5, 2010	365	6137	341	359
	4	4-6, 2010	434	7087	399	425

Evaluation Metrics:

$$\mathtt{MRR} = \frac{1}{N_U} \sum_{i=1}^{N_u} \frac{1}{rank_i} \qquad \mathtt{Top@KError} = \frac{|e_K|}{N_U}$$

Experimental Results

TABLE 3
Results in terms of MRR (↑) and Top@K Error (↓) of different methods on the MovieLens data set.

Method	Duration 1					Duration 2				
	MRR	Top@5	Top@10	Top@15	Top@20	MRR	Top@5	Top@10	Top@15	Top@20
KNN	0.0352	0.9688	0.9558	0.9221	0.8779	0.0427	0.9602	0.9093	0.8739	0.8208
PCA	0.0426	0.9610	0.9169	0.8779	0.8442	0.0354	0.9735	0.9425	0.9071	0.8783
LCE	0.1413	0.8312	0.8182	0.8000	0.7792	0.0772	0.9270	0.9004	0.8805	0.8319
LLAE*	0.0682	0.9247	0.8182	0.7714	0.7247	0.0532	0.9513	0.8850	0.8429	0.7987
UIMA-I	0.1485	0.8234	**0.7922**	0.7714	**0.7169**	**0.0781**	**0.8916**	**0.8473**	**0.8031**	**0.7412**
UIMA-II	**0.1486**	**0.8130**	0.7974	**0.7506**	0.7221	0.0607	0.9226	0.9004	0.8473	0.7677

Method	Duration 3					Duration 4				
	MRR	Top@5	Top@10	Top@15	Top@20	MRR	Top@5	Top@10	Top@15	Top@20
KNN	0.0448	0.9615	0.9331	0.8884	0.8519	0.0543	0.9446	0.8941	0.8453	0.7899
PCA	0.0416	0.9594	0.9229	0.8763	0.8256	0.0450	0.9577	0.9202	0.8795	0.8208
LCE	0.0722	0.9391	0.8966	0.8418	0.7890	0.0731	0.9169	0.8388	0.8192	0.7899
LLAE*	0.0484	0.9452	0.9047	0.8580	0.8154	0.0643	0.9186	0.8616	0.8208	0.7801
UIMA-I	0.1045	**0.8864**	0.8377	0.7830	0.7302	0.1460	**0.8274**	**0.7785**	**0.7541**	**0.7199**
UIMA-II	**0.1075**	0.9026	**0.8195**	**0.7566**	**0.7181**	**0.1466**	0.8453	0.8029	0.7720	0.7443

Experimental Results

TABLE 4
Results in terms of MRR (↑) and Top@K Error (↓) of different methods on the LastFM data set.

Method	Duration 1					Duration 2				
	MRR	Top@5	Top@10	Top@15	Top@20	MRR	Top@5	Top@10	Top@15	Top@20
KNN	0.0515	0.9416	0.9156	0.8766	0.8442	0.0476	0.9581	0.9058	0.8796	0.8534
PCA	0.0374	0.9740	0.9286	0.9026	0.8571	0.0329	0.9738	0.9476	0.9215	0.8953
LCE	0.0652	0.9545	0.9026	0.8636	0.8377	0.0484	0.9476	0.9162	0.8901	0.8482
LLAE*	0.0468	0.9416	0.9026	0.8571	0.8377	0.0721	0.9267	0.8848	0.8586	0.8325
UIMA-I	0.0734	**0.9026**	**0.8571**	0.8312	**0.7792**	**0.1047**	**0.8691**	0.8325	0.7958	**0.7539**
UIMA-II	**0.0763**	0.9156	0.8701	**0.8247**	0.7987	0.0970	0.8901	**0.8272**	**0.7906**	**0.7539**

Method	Duration 3					Duration 4				
	MRR	Top@5	Top@10	Top@15	Top@20	MRR	Top@5	Top@10	Top@15	Top@20
KNN	0.0405	0.9677	0.9194	0.8817	0.8602	0.0730	0.9307	0.8515	0.8020	0.7723
PCA	0.0409	0.9624	0.9570	0.9247	0.8871	0.0434	0.9604	0.9604	0.9010	0.8911
LCE	0.0465	0.9624	0.9247	0.9032	0.8710	0.0912	0.8911	0.8317	0.7624	0.7030
LLAE*	0.0574	0.9355	0.9032	0.8763	0.8548	0.0922	0.9208	0.8119	0.7822	0.7327
UIMA-I	**0.1002**	0.8925	**0.8387**	**0.8011**	**0.7742**	**0.1597**	0.8218	**0.7327**	**0.6733**	**0.6040**
UIMA-II	0.0952	**0.8871**	0.8656	0.8441	0.8226	0.1471	**0.8119**	0.7723	0.7129	0.6634

Experimental Results

TABLE 5
Results in terms of MRR (↑) and Top@K Error (↓) of different methods on the Delicious data set.

Method	Duration 1					Duration 2				
	MRR	Top@5	Top@10	Top@15	Top@20	MRR	Top@5	Top@10	Top@15	Top@20
KNN	0.0253	0.9778	0.9494	0.9430	0.9335	0.0264	0.9785	0.9540	0.9325	0.9202
PCA	0.0232	0.9810	0.9715	0.9589	0.9241	0.0216	0.9877	0.9663	0.9479	0.9325
LCE	0.0281	0.9778	0.9494	0.9241	0.9114	0.0326	0.9693	0.9448	0.9264	0.9018
LLAE*	0.0453	0.9589	0.9335	0.9082	0.9019	0.0399	0.9540	0.9356	0.9080	0.8834
UIMA-I	**0.0831**	**0.9082**	**0.8797**	**0.8449**	**0.8133**	**0.0621**	**0.9294**	0.9018	**0.8681**	0.8589
UIMA-II	0.0685	0.9272	0.9019	0.8576	0.8418	0.0567	0.9356	**0.8957**	**0.8681**	**0.8466**

Method	Duration 3					Duration 4				
	MRR	Top@5	Top@10	Top@15	Top@20	MRR	Top@5	Top@10	Top@15	Top@20
KNN	0.0264	0.9753	0.9562	0.9452	0.9233	0.0294	0.9700	0.9539	0.9286	0.9101
PCA	0.0258	0.9726	0.9534	0.9452	0.9233	0.0254	0.9724	0.9585	0.9309	0.9194
LCE	0.0268	0.9753	0.9589	0.9397	0.9288	0.0250	0.9793	0.9700	0.9516	0.9378
LLAE*	0.0519	0.9425	0.9151	0.8959	0.8658	0.0557	0.9332	0.9194	0.9101	0.8986
UIMA-I	0.0712	0.9233	0.9014	0.8795	**0.8548**	**0.0685**	**0.9286**	**0.8940**	**0.8756**	**0.8571**
UIMA-II	**0.0724**	**0.9178**	**0.8877**	**0.8767**	0.8712	0.0565	0.9424	0.9194	0.8986	0.8848

Experimental Results

TABLE 6

Ablation study in terms of MRR (↑) and Top@K Error (↓) of UIMA-II on the MovieLens data set.

	$\mathcal{L}_{\mathrm{U}}$	$\mathcal{L}_{\mathrm{M}}$	$\mathcal{L}_{\mathrm{UM}}$	MRR	Top@5	Top@10	Top@15	Top@20
1	√			0.0555	0.9558	0.9093	0.8518	0.8009
2		√		0.0465	0.9535	0.9093	0.8850	0.8496
3			√	0.0541	0.9425	0.9159	0.8673	0.8562
4	√	√		0.0563	0.9558	0.8938	0.8496	0.8009
5	√		√	0.0564	0.9580	0.9071	**0.8385**	0.7965
6		√	√	0.0316	0.9735	0.9358	0.9071	0.8717
7	√	√	√	**0.0607**	**0.9226**	**0.9004**	0.8473	**0.7677**

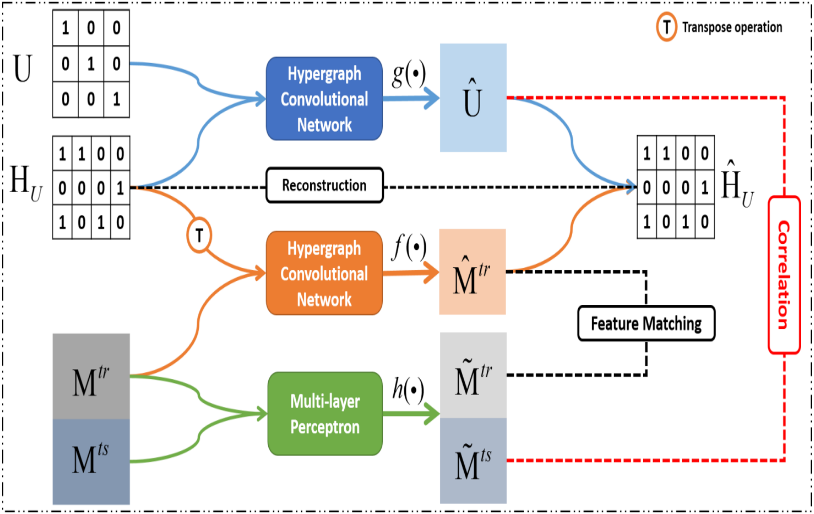
Experimental Results

TABLE 7
Results in terms of MRR (↑) and Top@K Error (↓) of using different graphs for UIMA on the LastFM data set.

Method	Duration 1					Duration 2				
	MRR	Top@5	Top@10	Top@15	Top@20	MRR	Top@5	Top@10	Top@15	Top@20
UIMA-I (simple graph)	0.0699	0.9221	0.8701	**0.8182**	0.7922	0.0444	0.9476	0.9110	0.8691	0.8168
UIMA-I (hypergraph)	0.0734	**0.9026**	**0.8571**	0.8312	**0.7792**	**0.1047**	**0.8691**	0.8325	0.7958	**0.7539**
UIMA-II (simple graph)	0.0475	0.9416	0.9091	0.8701	0.8182	0.0443	0.9372	0.8848	0.8743	0.8220
UIMA-II (hypergraph)	**0.0763**	0.9156	0.8701	0.8247	0.7987	0.0970	0.8901	**0.8272**	**0.7906**	**0.7539**

Method	Duration 3					Duration 4				
	MRR	Top@5	Top@10	Top@15	Top@20	MRR	Top@5	Top@10	Top@15	Top@20
UIMA-I (simple graph)	0.0409	0.9624	0.9355	0.8925	0.8710	0.0776	0.9109	0.8317	0.8119	0.7624
UIMA-I (hypergraph)	**0.1002**	0.8925	**0.8387**	**0.8011**	**0.7742**	**0.1597**	0.8218	**0.7327**	**0.6733**	**0.6040**
UIMA-II (simple graph)	0.0422	0.9516	0.9355	0.9032	0.8602	0.0916	0.9010	0.8119	0.7723	0.7327
UIMA-II (hypergraph)	0.0952	**0.8871**	0.8656	0.8441	0.8226	0.1471	**0.8119**	0.7723	0.7129	0.6634

Feature Matching Machine



Feature Matching Machine

TABLE I

STATISTICAL INFORMATION ON THE USED DATASETS. FOR COLD-START ITEM RECOMMENDATION, ON THE MOVIELENS DATASET, WE DENOTE MOVIELENS 1/2/3/4/5 AS THE FIRST, SECOND, THIRD, FOURTH, AND FIFTH DURATIONS, RESPECTIVELY. SIMILARLY, ON THE LASTFM DATASET, WE DENOTE LASTFM 1/2/3 AS THE FIRST, SECOND, AND THIRD DURATIONS, RESPECTIVELY.

Task	Dataset	Training				Test		
		#user	#item	#record	density	#user	#item	#record
Cold-start item recommendation	MovieLens 1	296	3,719	120,048	10.91%	296	637	7,035
	MovieLens 2	388	4,352	162,238	9.61%	388	411	5,939
	MovieLens 3	576	4,802	240,577	8.70%	576	340	6,912
	MovieLens 4	657	5,137	278,349	8.25%	657	313	7,654
	MovieLens 5	705	5,411	282,878	7.42%	705	379	8,703
	LastFM 1	207	4,617	10,735	1.12%	207	1,602	2,007
	LastFM 2	275	6,569	17,273	0.96%	275	1,508	1,845
	LastFM 3	291	7,402	18,751	0.87%	291	1,230	1,548
Cold-start user recommendation	BlogCatalog	9,280	39	-	-	1,032	39	-
	Ciao	1,937	8,000	-	-	216	8,000	-
	Epinions 1	1,791	8,000	-	-	200	8,000	-
	Epinions 2	5,328	200	-	-	592	200	-

Metrics

- Precision is the proportion of correctly recommended items, i.e., preferred items existing in the recommendation list
- Recall is the proportion of correctly recommended items to the total number of items which should be recommended
- `HR@K` is defined as the total number of correctly recommended items divided by the total number of all users' preferred items
- NDCG: Cumulative Gain is the sum of all the relevance scores in a recommendation set

Feature Matching Machine (Cold-Start Item)

TABLE III
RESULTS ON THE MOVIELENS AND LASTFM DATASETS FOR COLD-START ITEM RECOMMENDATION. THE BEST RESULTS ARE HIGHLIGHTED IN BOLD.

Dataset	Method	Precision			Recall			NDCG			HR		
		@20	@50	@100	@20	@50	@100	@20	@50	@100	@20	@50	@100
MovieLens 1	PCA	0.0764	0.0636	0.0545	0.0430	0.0935	0.1669	0.0913	0.0956	0.1261	0.0643	0.1338	0.2294
	LCE	0.1576	0.0923	0.0735	0.1801	0.2506	0.3738	0.2278	0.2265	0.2745	0.1326	0.1942	0.3092
	Heater	0.0662	0.0457	0.0431	0.0631	0.1002	0.1861	0.0652	0.0725	0.1069	0.0557	0.0962	0.1812
	FMM	**0.2064**	**0.1369**	**0.1005**	**0.2317**	**0.3681**	**0.5004**	**0.3037**	**0.3224**	**0.3745**	**0.1737**	**0.2880**	**0.4229**
MovieLens 2	PCA	0.0713	0.0573	0.0454	0.0908	0.1705	0.2630	0.0962	0.1248	0.1618	0.0931	0.1872	0.2965
	LCE	0.1693	0.1057	0.0852	0.2638	0.3969	0.6072	0.2348	0.2708	0.3502	0.2212	0.3453	0.5563
	Heater	0.1847	0.1112	0.0840	0.2833	0.4238	0.6008	0.2572	0.2939	0.3611	0.2413	0.3634	0.5487
	FMM	**0.1995**	**0.1339**	**0.0907**	**0.3239**	**0.4995**	**0.6389**	**0.2930**	**0.3488**	**0.4025**	**0.2606**	**0.4373**	**0.5925**
MovieLens 3	PCA	0.0589	0.0518	0.0432	0.0822	0.1736	0.2903	0.0768	0.1099	0.1520	0.0982	0.2157	0.3597
	LCE	0.1161	0.0741	0.0571	0.2436	0.3436	0.5020	0.1957	0.2257	0.2796	0.1936	0.3086	0.4761
	Heater	0.1345	0.1019	0.0661	0.2706	0.4804	0.5972	0.2037	0.2720	0.3114	0.2241	0.4248	0.5512
	FMM	**0.1359**	**0.1068**	**0.0739**	**0.2839**	**0.5080**	**0.6755**	**0.2453**	**0.3225**	**0.3780**	**0.2266**	**0.4452**	**0.6155**
MovieLens 4	PCA	0.0686	0.0540	0.0425	0.1016	0.1875	0.2929	0.0898	0.1197	0.1560	0.1178	0.2319	0.3645
	LCE	0.1209	0.0828	0.0600	0.2500	0.4029	0.5513	0.1993	0.2474	0.2954	0.2076	0.3555	0.5146
	Heater	0.1390	0.0961	0.0639	0.2593	0.4657	0.5786	0.2350	0.2970	0.3354	0.2387	0.4126	0.5485
	FMM	**0.1656**	**0.1104**	**0.0761**	**0.3503**	**0.5298**	**0.6969**	**0.2404**	**0.3002**	**0.3538**	**0.2843**	**0.4740**	**0.6530**
MovieLens 5	PCA	0.0466	0.0390	0.0373	0.0572	0.1245	0.2410	0.0628	0.0841	0.1252	0.0755	0.1581	0.3025
	LCE	0.1457	0.1002	0.0645	0.2960	0.4803	0.5788	0.1967	0.2565	0.2897	0.2361	0.4057	0.5223
	Heater	0.1407	0.0808	0.0589	0.2792	0.3711	0.5266	0.3042	0.3225	0.3709	0.2280	0.3274	0.4772
	FMM	**0.1938**	**0.1193**	**0.0733**	**0.3837**	**0.5523**	**0.6560**	**0.3646**	**0.4095**	**0.4417**	**0.3139**	**0.4831**	**0.5937**
LastFM 1	PCA	0.0089	0.0079	0.0072	0.0240	0.0396	0.0772	0.0133	0.0185	0.0283	0.0184	0.0409	0.0747
	LCE	0.0157	0.0101	0.0090	0.0381	0.0560	0.1142	0.0343	0.0380	0.0527	0.0324	0.0523	0.0932
	Heater	0.0196	0.0135	0.0098	**0.0454**	**0.0736**	**0.1231**	0.0358	**0.0432**	**0.0550**	0.0404	0.0698	0.1006
	FMM	**0.0205**	**0.0138**	**0.0110**	0.0404	0.0695	0.1046	**0.0369**	0.0431	0.0543	**0.0424**	**0.0713**	**0.1136**
LastFM 2	PCA	0.0027	0.0041	0.0039	0.0037	0.0288	0.0547	0.0046	0.0119	0.0183	0.0081	0.0309	0.0585
	LCE	0.0036	0.0028	0.0032	0.0121	0.0302	0.0609	0.0067	0.0106	0.0174	0.0108	0.0211	0.0482
	Heater	**0.0135**	**0.0089**	0.0071	**0.0640**	**0.1046**	0.1514	**0.0309**	**0.0410**	0.0516	**0.0401**	0.0661	0.1057
	FMM	0.0122	**0.0089**	**0.0077**	0.0624	0.1026	**0.1611**	0.0301	0.0409	**0.0550**	0.0363	**0.0667**	**0.1149**
LastFM 3	PCA	0.0055	0.0038	0.0040	0.0197	0.0375	0.0857	0.0120	0.0164	0.0269	0.0207	0.0362	0.0749
	LCE	0.0015	0.0023	0.0022	0.0068	0.0267	0.0392	0.0037	0.0091	0.0128	0.0058	0.0213	0.0420
	Heater	0.0040	0.0045	0.0044	0.0189	0.0617	0.1057	0.0088	0.0193	0.0296	0.0149	0.0420	0.0827
	FMM	**0.0072**	**0.0076**	**0.0065**	**0.0416**	**0.0869**	**0.1372**	**0.0197**	**0.0323**	**0.0441**	**0.0271**	**0.0717**	**0.1221**

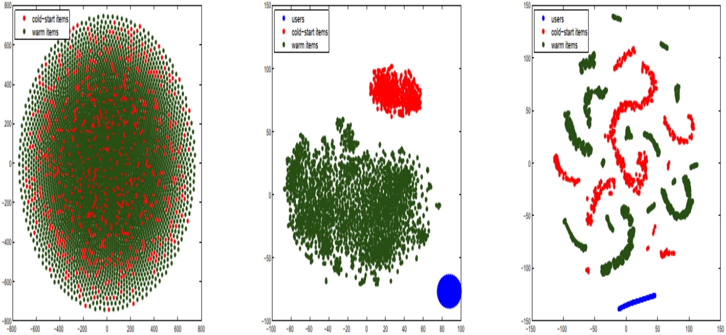
Feature Matching Machine (Cold-Start User)

TABLE IV

RESULTS ON THE BLOGCATALOG, CIAO, AND EPINIONS DATASETS FOR COLD-START USER RECOMMENDATION. THE BEST RESULTS ARE HIGHLIGHTED IN BOLD.

Dataset	Method	Precision		Recall		HR	
		@20	@30	@20	@30	@20	@30
BlogCatalog	PCA	0.0142±0.0007	0.0213±0.0004	0.1991±0.0106	0.4421±0.0122	0.2028±0.0107	0.4567±0.0081
	CMF	0.0411±0.0012	0.0368±0.0009	0.6128±0.0210	0.8067±0.0132	0.5874±0.0155	0.7892±0.0136
	LoCo	0.0586±0.0009	0.0432±0.0007	0.8493±0.0113	0.9321±0.0090	0.8376±0.0093	0.9259±0.0088
	Heater	0.0389±0.0027	0.0322±0.0027	0.5678±0.0409	0.7003±0.0528	0.5550±0.0355	0.6891±0.0535
	FMM	**0.0592±0.0010**	**0.0450±0.0006**	**0.8529±0.0098**	**0.9668±0.0059**	**0.8453±0.0085**	**0.9644±0.0059**
Ciao	PCA	0.0030±0.0007	0.0030±0.0007	0.0032±0.0024	0.0049±0.0023	0.0024±0.0006	0.0035±0.0008
	CMF	0.0738±0.0117	0.0597±0.0105	0.0595±0.0096	0.0720±0.0118	0.0576±0.0092	0.0699±0.0123
	LoCo	0.0522±0.0053	0.0431±0.0043	0.0430±0.0042	0.0537±0.0051	0.0408±0.0051	0.0507±0.0065
	Heater	0.0104±0.0029	0.0102±0.0017	0.0065±0.0027	0.0094±0.0027	0.0082±0.0025	0.0119±0.0022
	FMM	**0.0839±0.0099**	**0.0706±0.0075**	**0.0669±0.0069**	**0.0857±0.0093**	**0.0654±0.0079**	**0.0828±0.0099**
Epinions 1	PCA	0.0084±0.0030	0.0079±0.0029	0.0028±0.0009	0.0042±0.0015	0.0027±0.0008	0.0039±0.0011
	CMF	0.0363±0.0058	0.0320±0.0046	0.0129±0.0031	0.0173±0.0048	0.0121±0.0023	0.0160±0.0028
	LoCo	0.0551±0.0037	0.0492±0.0039	0.0188±0.0019	0.0251±0.0023	0.0183±0.0012	0.0245±0.0019
	Heater	0.0261±0.0078	0.0248±0.0056	0.0105±0.0025	0.0146±0.0045	0.0087±0.0027	0.0124±0.0030
	FMM	**0.0773±0.0055**	**0.0673±0.0066**	**0.0273±0.0022**	**0.0353±0.0034**	**0.0257±0.0028**	**0.0335±0.0041**
Epinions 2	PCA	0.0227±0.0021	0.0227±0.0017	0.0821±0.0062	0.1248±0.0061	0.0750±0.0057	0.1123±0.0066
	CMF	0.0550±0.0037	0.0489±0.0032	0.1883±0.0196	0.2474±0.0214	0.1821±0.0128	0.2429±0.0165
	LoCo	0.0530±0.0020	0.0475±0.0018	0.1765±0.0065	0.2346±0.0098	0.1753±0.0075	0.2357±0.0081
	Heater	0.0374±0.0057	0.0350±0.0037	0.1289±0.0298	0.1805±0.0254	0.1240±0.0205	0.1738±0.0191
	FMM	**0.0602±0.0027**	**0.0532±0.0021**	**0.2033±0.0086**	**0.2769±0.0095**	**0.1992±0.0088**	**0.2639±0.0064**

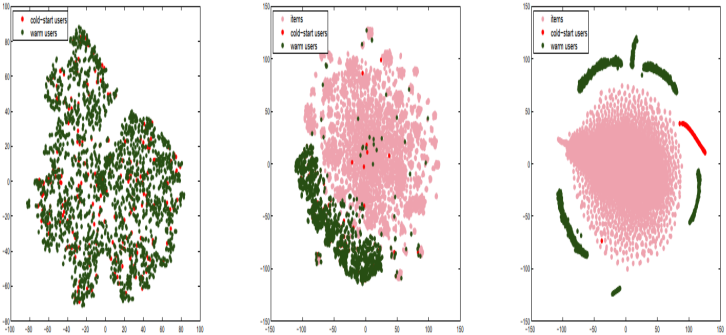
Feature Matching Machine (Cold-Start Item)



(a) Original (b) LCE (c) FMM

Fig. 3. Visualization of item distribution on the MovieLens 1 dataset. The ideal distribution is the users, warm items, and cold-start items are clustered and the warm and cold-start items are mixed from the view of users, as the distribution of FMM.

Feature Matching Machine (Cold-Start User)



(a) Original (b) LoCo (c) FMM

Fig. 4. Visualization of user distribution on the Ciao dataset. The ideal distribution is the items, warm users, and cold-start users are clustered and the warm and cold-start users are mixed from the view of users, as the distribution of FMM.

Revisit of Hypergraph Convolution

- Hypergraph convolution operation:

$$U^{(l+1)} = \sigma(L^{-1/2} H \Omega H^T L^{-1/2} U^{(l)} W)$$

  Perform convolution operation on the vertices

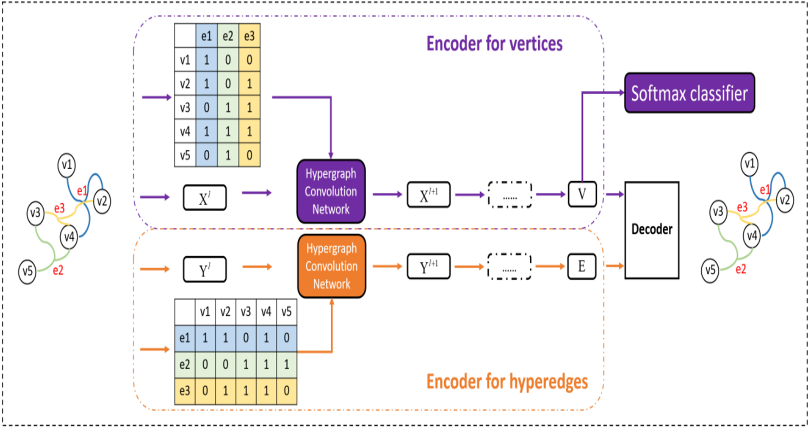
- "Compressed" hyperedges information are incorporated in the matrix multiplication:

$$H \Omega H^T$$

  Can it be well-explored ?

- Define the convolution operations on the hyperedges

Hypergraph Convolution on Vertices-Hyperedges Network



$$X^{(l+1)} = \sigma(L^{-1/2} H \Omega H^T L^{-1/2} X^{(l)} W)$$

$$Y^{(l+1)} = \sigma(D^{-1/2} H^T \Psi H D^{-1/2} Y^{(l)} U)$$

Experimental Results

Table 2. Data set statistics on Setting I collected from [23].

Data	Vertices	Edges	Classes	Features	Label rate
Citeseer	3327	4732	6	3703	0.036
Cora	2708	5429	7	1433	0.052
Pubmed	19717	44338	3	500	0.003

Table 3. Data set statistics on Setting II collected from [12].

Data	Vertices	Edges	Classes	Features	Label rate
Citeseer co-citation	1498	1107	6	3703	0.150
Pubmed co-citation	3840	7963	3	500	0.020
Cora co-authorship	16313	7389	10	1000	0.052

Experimental Results

Table 4. Experimental results (%) of different semi-supervised vertex classification methods on Setting I.

Method	Citeseer	Pubmed	Cora
DeepWalk [40]	43.2	65.3	67.2
ICA [41]	69.1	73.9	75.1
Planetoid [62]	64.7	77.2	75.7
Chebyshev [11]	67.4 ± 1.0	76.7 ± 0.1	78.6 ± 0.6
GCN [23]	67.9 ± 0.5	78.9 ± 0.7	80.1 ± 0.5
MoNet [35]	-	78.8 ± 0.4	81.7 ± 0.5
Variance Reduction [7]	70.9 ± 0.2	79.0 ± 0.4	82.0 ± 0.8
FeaStNet [48]	69.3 ± 1.1	76.6 ± 0.6	80.4 ± 0.7
GAT [47]	72.5 ± 0.7	79.0 ± 0.3	83.0 ± 0.7
GWNN [57]	71.7	79.1	82.8
CayleyNet [29]	67.1 ± 2.4	75.6 ± 3.6	81.2 ± 1.2
HighOrder [36]	64.2 ± 1.0	75.0 ± 2.6	76.6 ± 1.2
ARMA [4]	72.5 ± 0.4	78.9 ± 0.3	83.4 ± 0.6
HGNN [14]	-	80.1	81.6
HCHA [2]	71.2 ± 0.4	78.4 ± 0.3	82.7 ± 0.3
HCVH	**74.7 ± 0.5**	**80.5 ± 0.6**	**83.5 ± 0.6**

Experimental Results

Table 5. Experimental results (%) of different semi-supervised vertex classification methods on Setting II.

Method	co-citation		co-authorship
	Citeseer	Pubmed	Cora
FastHyperGCN [58]	56.1 ± 11.2	54.4 ± 10.0	45.2 ± 12.9
HyperGCN [58]	54.7 ± 9.8	60.0 ± 10.7	55.0 ± 0.9
HGNN [14]	61.1 ± 2.2	63.3 ± 2.2	58.2 ± 0.3
HNHN [12]	64.8 ± 1.6	75.9 ± 1.5	63.9 ± 0.8
HCVH	**71.4 ± 1.2**	**77.1 ± 3.6**	**65.9 ± 0.5**

Experimental Results

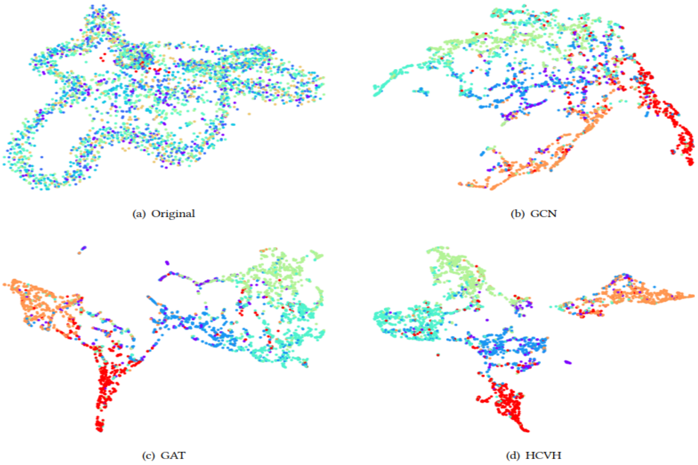
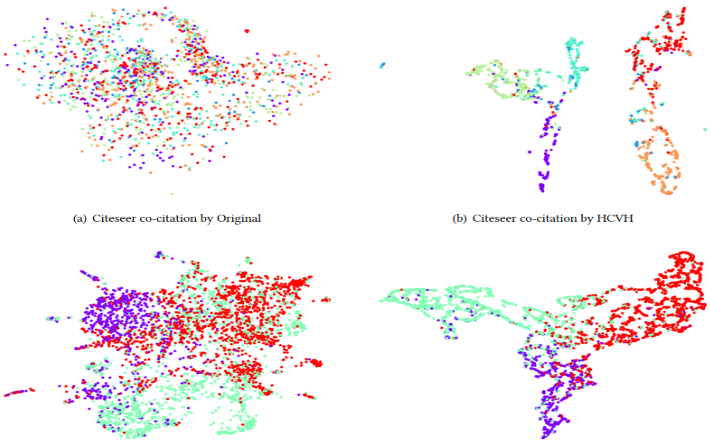


Fig. 6. Visualization of samples represented by original features and different graph learning features on the Citeseer data set on Setting I. The color indicates the class label in data set.

Experimental Results



(a) Citeseer co-citation by Original

(b) Citeseer co-citation by HCVH

(c) Pubmed co-citation by Original

(d) Pubmed co-citation by HCVH

Fig. 7. Visualization of samples represented by original features and HCVH features on the Citeseer and Pubmed data sets on Setting II. The color indicates the class label in data set.

Experimental Results

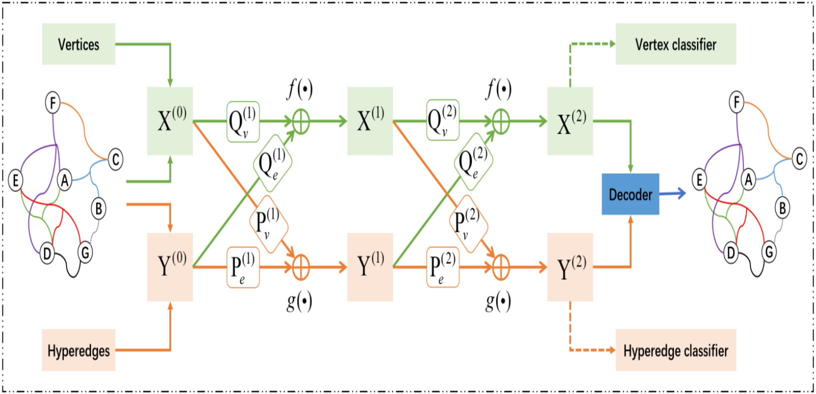


Fig. 2. An overview of the proposed HCoN model. The superscripts "(0)", "(1)", and "(2)" denote different layers. The latent vertex representations are learned by aggregating the information from previous vertices and hyperedges via network parameters $\mathbf{Q}_v$ and $\mathbf{Q}_e$. Similarly, the latent hyperedge representations are obtained by aggregating the information from previous vertices and hyperedges via network parameters $\mathbf{P}_v$ and $\mathbf{P}_e$. The notations $\mathbf{X}^{(2)}$ and $\mathbf{Y}^{(2)}$ are output representations of vertices and hyperedges, which are also represented by $\mathbf{V}$ and $\mathbf{E}$ in Eq. (11), respectively.

Summary

- Solve hypergraph representation learning problem
- Solve cold-start next-item recommendation problem: preserves the relationships between users and the historical items and the information of original item attributes, and captures the relationships between users and their training items
- The experimental results show that the recommendation performance of the proposed method is better than that of the baseline methods
- Research work for graphs with multiple relations
- Analysis of generalization error for the proposed approach

Thank you very much !