Chongqing University of Technology

Enhancing Information Diffusion Prediction with Self-Supervised Disentangled User and Cascade Representations

Zhangtao Cheng*
Wenxue Ye*
Leyuan Liu†
zhangtao.cheng@outlook.com
wenxue.ye@outlook.com
leyuanliu@uestc.edu.cn
University of Electronic Science and Technology of China
Chengdu, Sichuan, China

Wenxin Tai†
Fan Zhou
wxtai@std.uestc.edu.cn
fan.zhou@uestc.edu.cn
University of Electronic Science and Technology of China
Chengdu, Sichuan, China
Kash Institute of Electronics and Information Industry
Kashgar, Xinjiang, China

https://github.com/CZ-TAO12/DisenIDP

—— CIKM 2023

Reported by Yuyang Lai

Introduction

(1) User intents behind re-sharing behaviors are entangled.

interest- and dependency-wise hypergraphs

(2) Cascade temporal influence is dynamic drifting.

two separate attention-based encoders

To supervise the disentanglement process and prevent information loss, we design a self-supervised auxiliary task to guide disentanglement and learn fine-grained user and cascade representations.

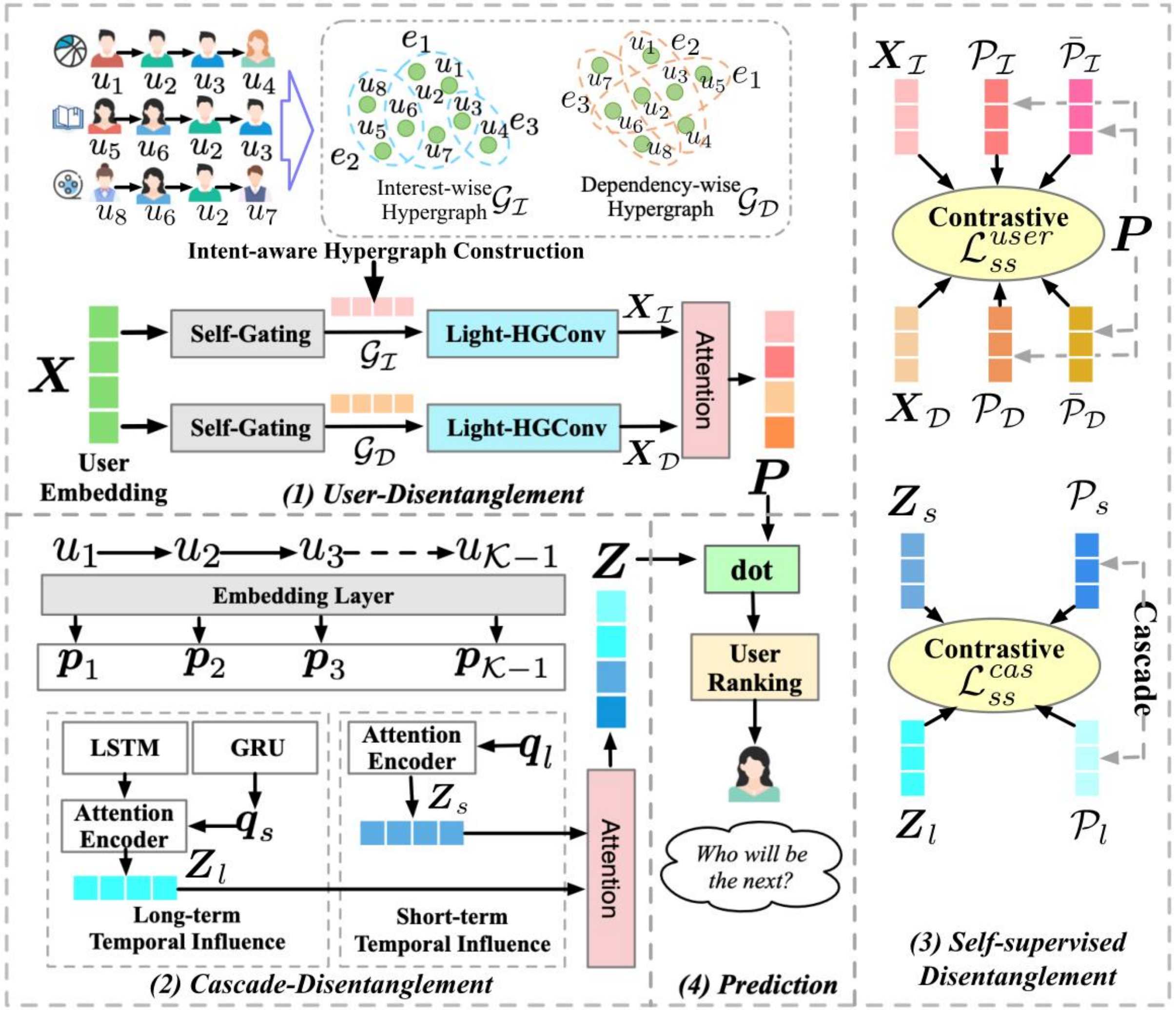

Figure 1: Overall framework of DisenIDP.

Method

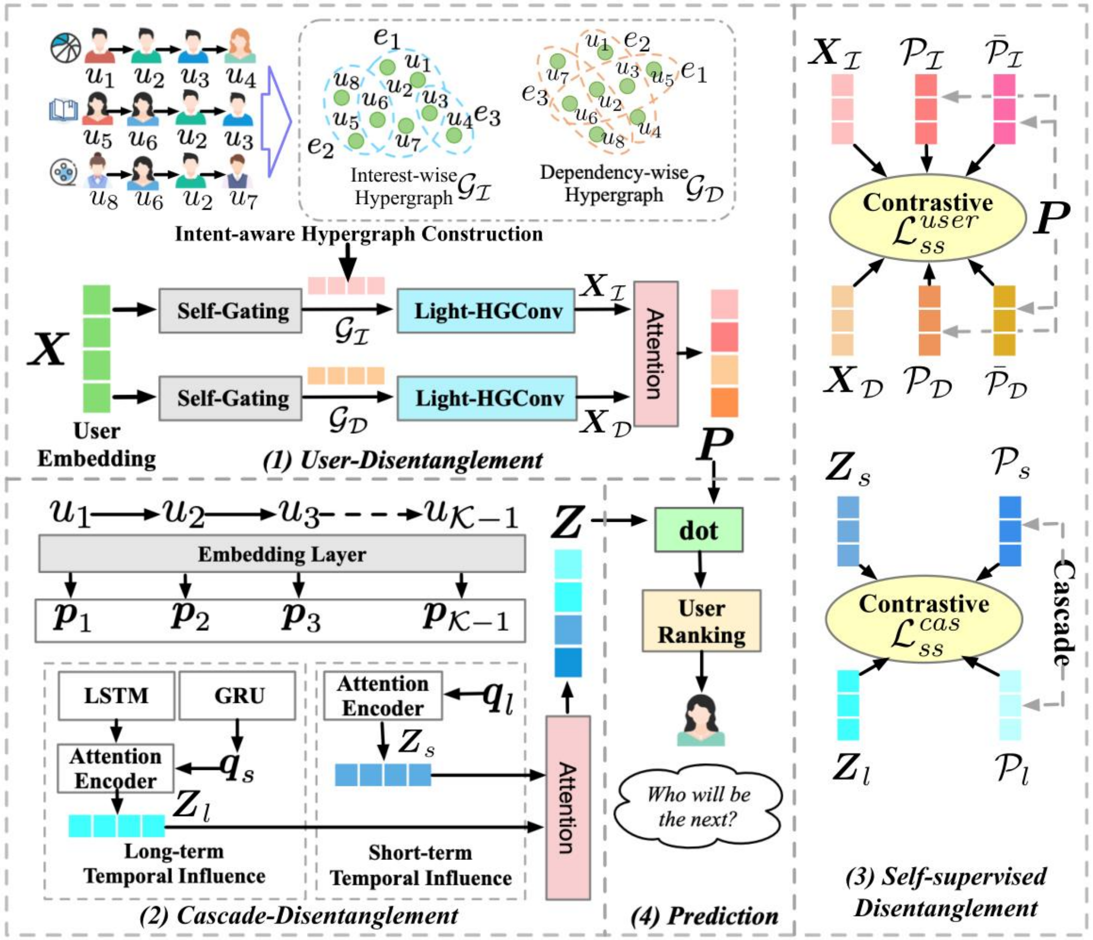

Figure 1: Overall framework of DisenIDP.

$$X_s = f^s_{gate}(X) = X \odot \text{sigmoid}(X W_s + b_s) \qquad s \in \{\mathcal{I}, \mathcal{D}\}$$

$$X_s^{(l+1)} = D_s^{-1} H_s B_s^{-1} H_s^{\mathrm{T}} X_s^{(l)}$$

$$\bar{X}_s = \frac{1}{L+1} \sum_{l=0}^{L} X_s^l$$

$$\alpha_s = \frac{\exp\left(a^\top \cdot W_a \bar{X}_s\right)}{\sum_{s' \in \{\mathcal{D}, \mathcal{I}\}} \exp\left(a^\top \cdot W_a \bar{X}_{s'}\right)}, \qquad (1)$$

$$P = \sum_{s \in \{\mathcal{D}, \mathcal{I}\}} \alpha_s \bar{X}_s$$

Method

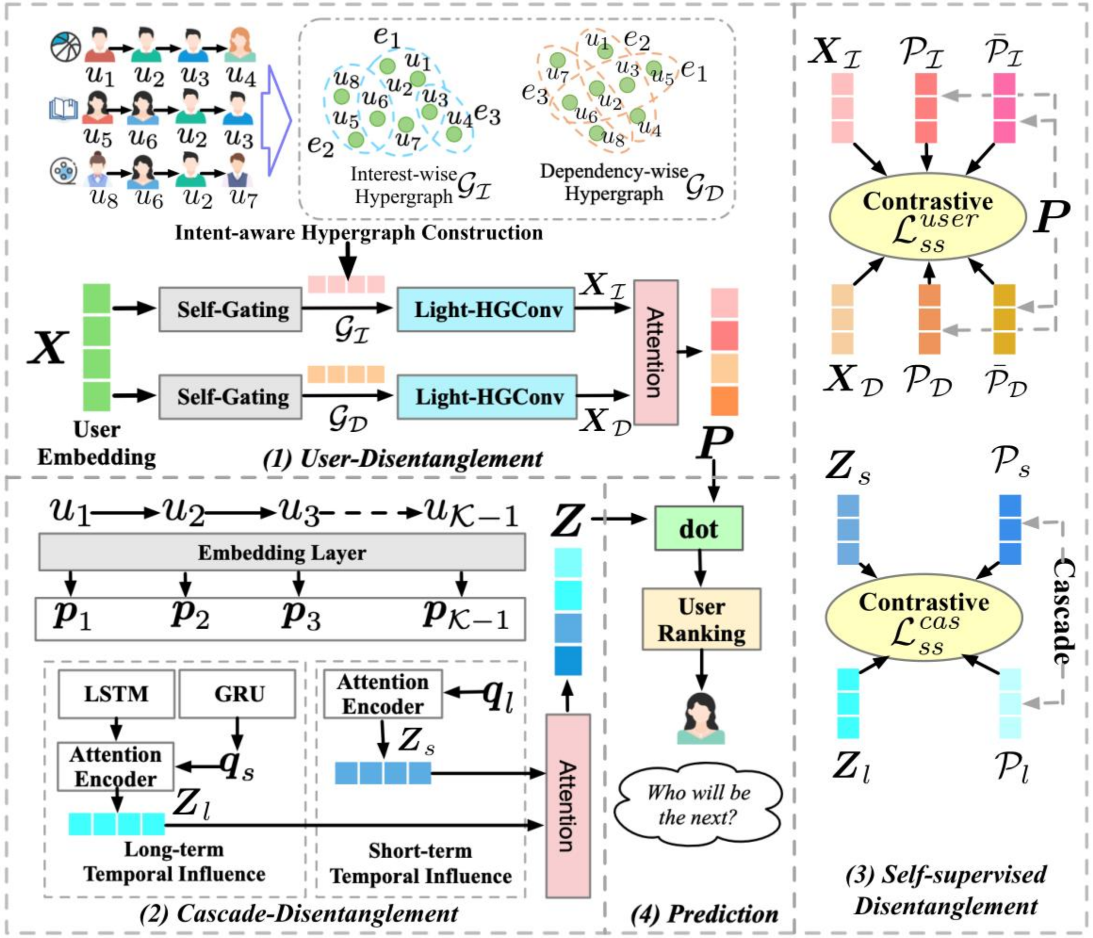

Figure 1: Overall framework of DisenIDP.

$$\alpha_j = \frac{\exp(\langle W_l^q p_{u_1}, W_l^k p_{u_j}\rangle)}{\sum_{j=1}^{\mathcal{K}-1}\exp(\langle W_l^q p_{u_1}, W_l^k p_{u_j}\rangle)}. \tag{2}$$

$$z_l^{\mathcal{K}} = \sum_{j=1}^{\mathcal{K}-1} \alpha_j W_l^v p_{u_j} \qquad z_s^{\mathcal{K}}$$

$$\begin{aligned} z_{k,q} &= \mathrm{ATTENTION}(z_k^* W_q^{\mathrm{query}}, z_k^* W_q^{\mathrm{key}}, z_k^* W_q^{\mathrm{value}}), \\ z_k &= (z_{k,1} || z_{k,2} || \dots || z_{k,q} || \dots || z_{k,Q}) W^O, \end{aligned} \tag{3}$$

Method

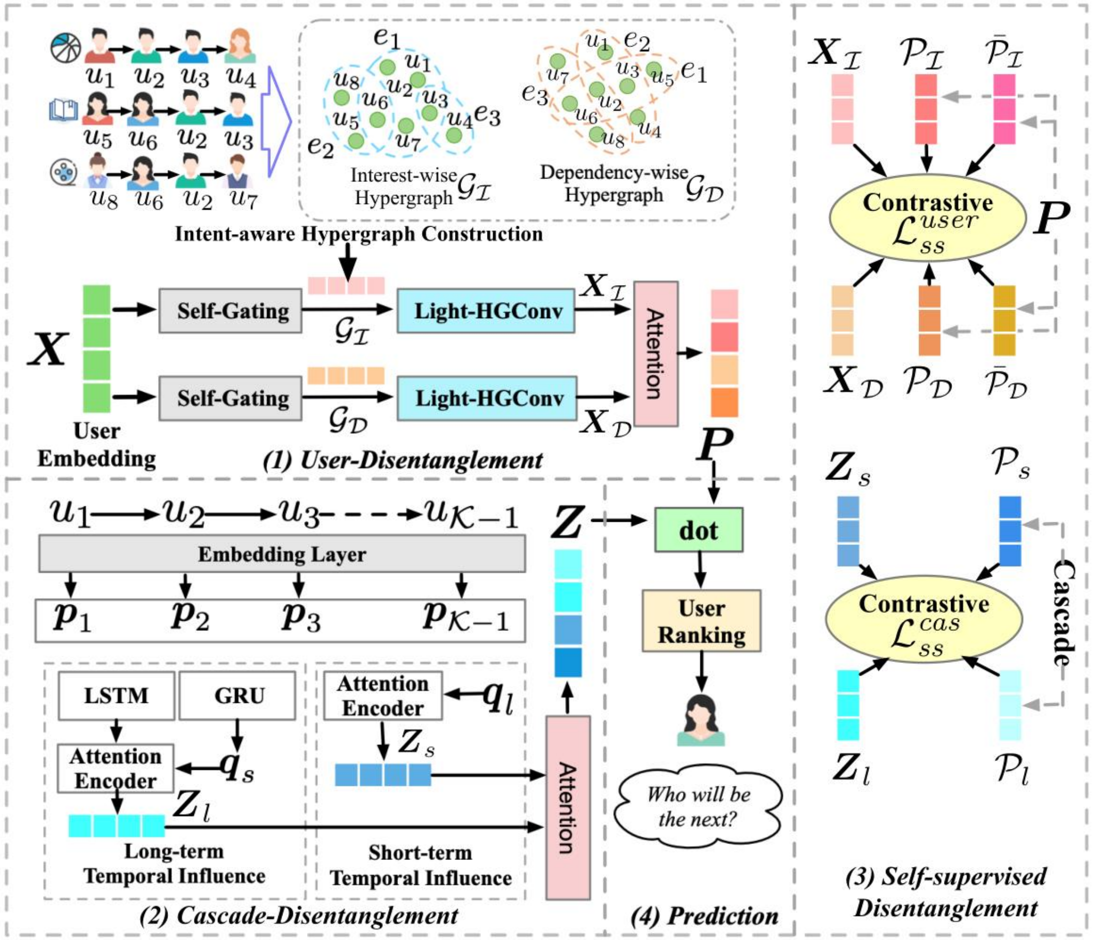

Figure 1: Overall framework of DisenIDP.

$$\mathcal{L}_{ss}^{user} = -\sum_{s\in\{\mathcal{D},\mathcal{I}\}} \log\sigma\left(f_D\left(X_s,\mathcal{P}_s\right) - f_D(X_s,\tilde{\mathcal{P}}_s)\right), \quad (4)$$

$$\mathcal{L} = \mathcal{L}_{cross} + \beta\mathcal{L}_{ss}^{user} + \lambda\mathcal{L}_{ss}^{cas},$$

Experiments

Table 1: Statistics of the datasets.

Dataset	#Users	#Cascades	#Train	#Val	#Test	Avg.Length
Twitter	12,627	3,454	2,763	345	346	38.22
Weibo	26,537	35,070	28,056	3,507	3,507	29.41

Experiments

Table 2: Performance comparisons on two datasets.

Model	Twitter				Weibo			
	H@10	H@100	M@10	M@100	H@10	H@100	M@10	M@100
DeepDiffuse	5.72	21.61	5.93	6.99	0.74	5.73	0.23	0.36
Topo-LSTM	10.45	25.42	9.51	14.68	1.86	12.89	0.60	0.90
NDM	22.45	35.12	15.59	16.03	9.85	39.31	4.05	4.95
SNIDSA	25.67	43.59	16.34	18.89	10.73	39.51	4.75	5.52
FOREST	30.28	50.12	21.45	22.36	15.59	52.55	7.55	8.63
Inf-VAE	14.93	46.42	19.83	21.82	10.37	38.05	5.90	6.58
DyHGCN	32.78	58.53	21.57	22.45	14.65	51.65	7.13	8.27
MS-HGAT	29.12	56.68	16.44	17.37	12.67	40.05	6.50	7.38
DisenIDP	**34.01**	**60.39**	**23.04**	**23.94**	**17.81**	**57.40**	**8.23**	**9.55**
% Improv.	3.75	3.17	6.81	6.63	14.23	9.22	9.00	10.66

Table 3: Ablation study of DisenIDP.

Model		Twitter		Weibo	
		H@100	M@100	H@100	M@100
DisenIDP	**All**	**60.39**	**23.94**	**57.40**	**9.55**
Multi-type Hypergraph	w/o $\mathcal{G}_{\mathcal{D}}$	59.52	23.26	56.43	8.89
	w/o $\mathcal{G}_{\mathcal{I}}$	59.19	23.04	56.37	8.77
User-Disentanglement	HyperGAT	58.93	22.98	55.55	8.93
	HyperGCN	58.89	22.88	54.21	8.76
Cascade-Disentanglement	w/o LTIE	55.20	17.62	44.98	7.34
	w/o STIE	57.59	19.95	45.35	7.44
Self-supervised-Disentanglement	w/o UD	59.63	23.35	56.63	8.81
	w/o CD	60.02	23.75	56.42	8.79

Thank you!