Cold-start Bundle Recommendation via Popularity-based Coalescence and Curriculum Heating

Hyunsik Jeon[1], Jong-eun Lee[2], Jeongin Yun[2], and U Kang[2]

[1] University of California, San Diego
[2] Seoul National University

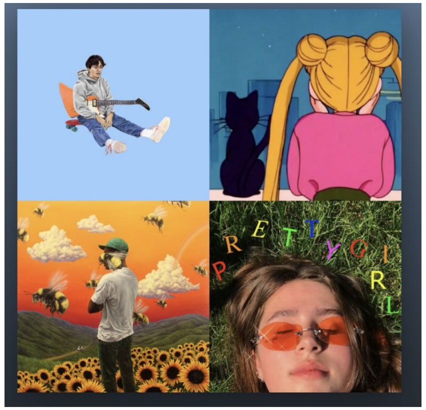
TheWebConf 2024

Outline

- **Introduction**
- Proposed Method
- Experiments
- Conclusion

Product Bundling

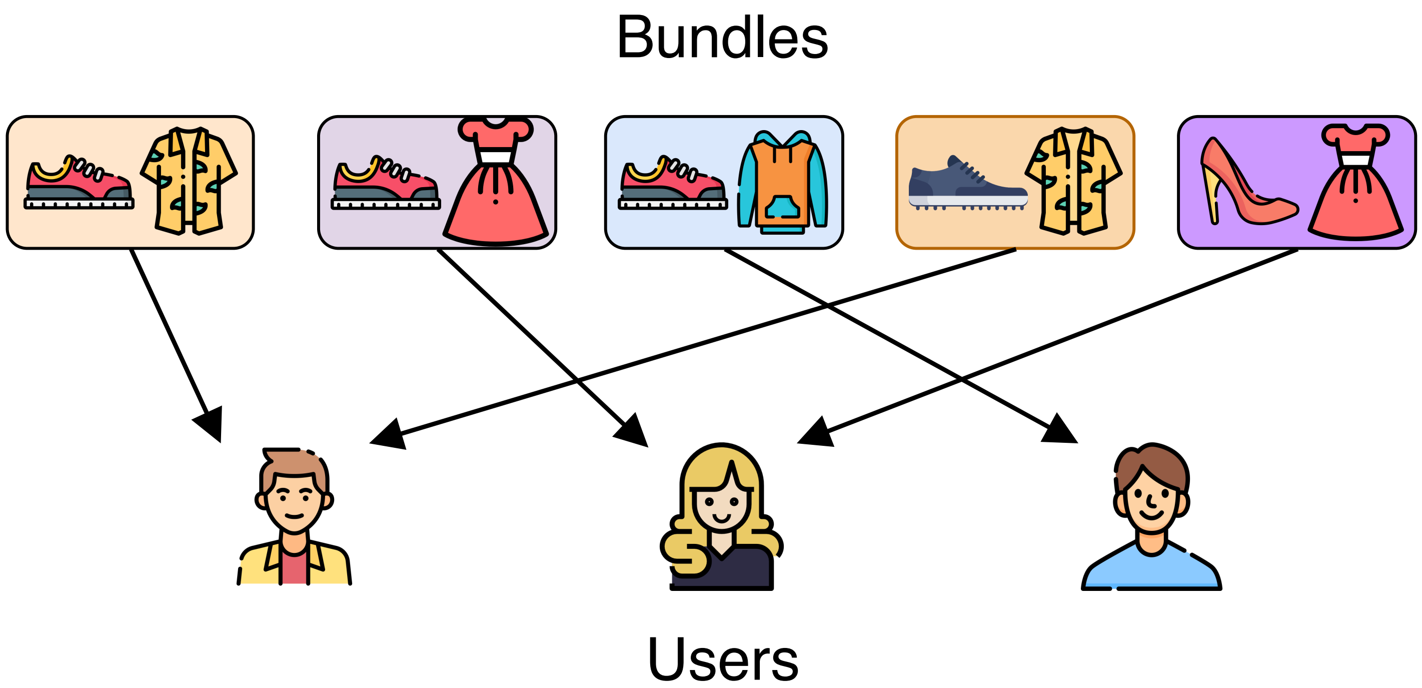
- **Prevalent strategy** in various platforms
  - One-stop convenience for customers
  - Increased exposure to lesser-known products
  - Cost-efficient offerings to customers



Examples of product bundling

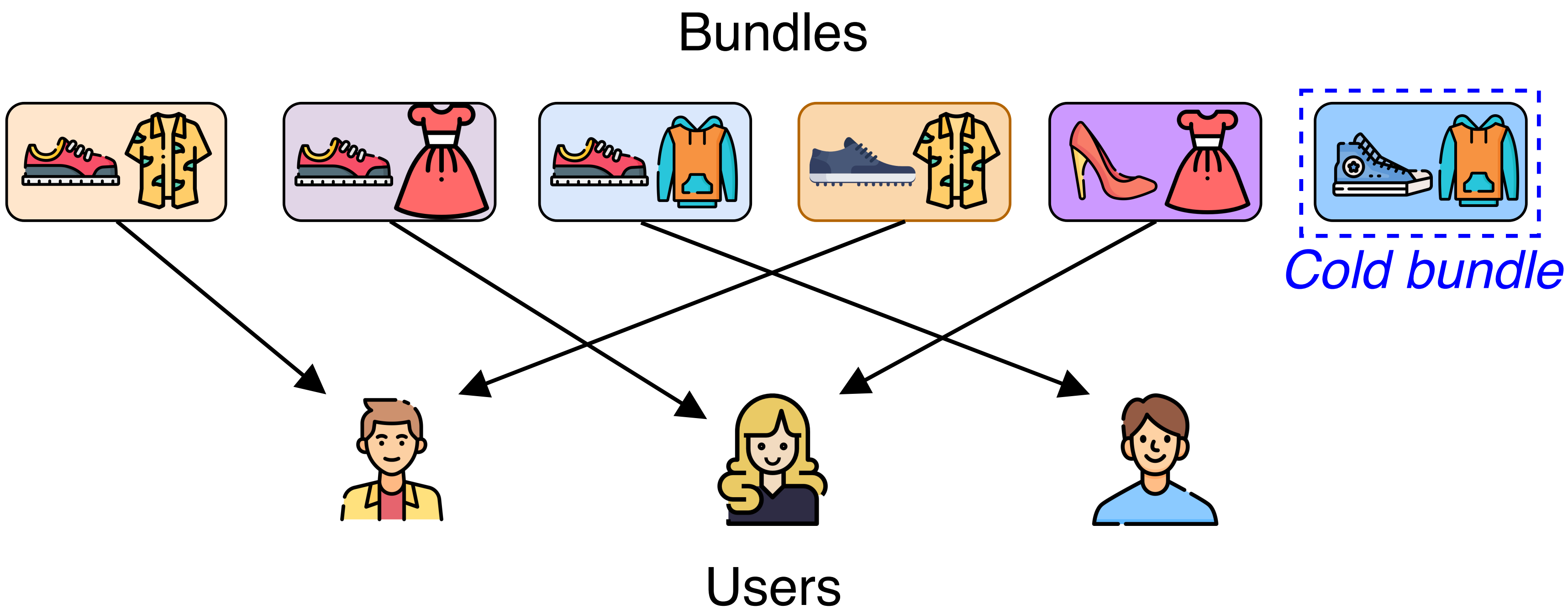
Bundle Recommendation

- Aims to recommend **bundles** instead of individual items
  - Became an important technique in industries



Cold-start Bundle Recommendation

- Aims to recommend **cold bundles** that have not been consumed by any users
  - New bundles are constantly created every day



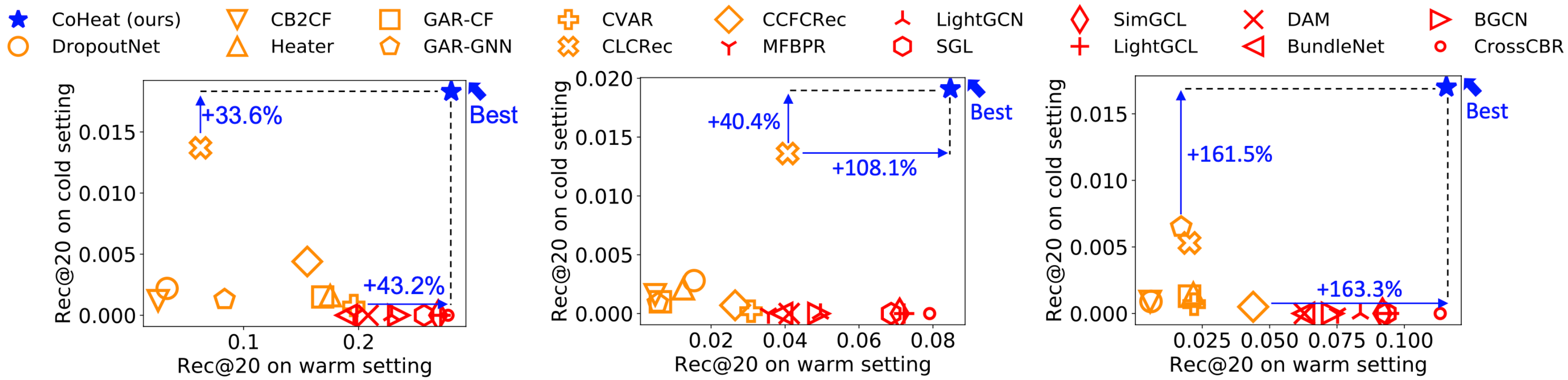
Problem Definition

Cold-start bundle recommendation

- **Given**
  - $\mathcal{U}$, $\mathcal{B}$, $\mathcal{I}$: sets of users, bundles, and items, resp.
  - A user-bundle interaction matrix $\mathbf{X} \in \mathbb{R}^{|\mathcal{U}| \times |\mathcal{B}_w|}$
    - $\mathcal{B}_w \subset \mathcal{B}$ refers to warm bundles
  - A user-item interaction matrix $\mathbf{Y} \in \mathbb{R}^{|\mathcal{U}| \times |\mathcal{I}|}$
  - A bundle-item affiliation matrix $\mathbf{Z} \in \mathbb{R}^{|\mathcal{B}| \times |\mathcal{I}|}$
- **Recommend**
  - k bundles from $\mathcal{B}$ to each user $u \in \mathcal{U}$
- **Challenge**
  - To estimate relationship between $u \in \mathcal{U}$ and $b \in \mathcal{B}_c$
    - $\mathcal{B}_c = \mathcal{B} \backslash \mathcal{B}_w$ refers to cold bundles

Our Contributions

- We propose ***CoHeat*** (Popularity-based **Co**alescence and Curriculum **Heat**ing)
  - **Problem**: tackling the cold-start problem in bundle recommendation, which has not been widely studied
  - **Method**: a novel learning framework
  - **Accurate**: outperforming ~17 competitors on real-world datasets



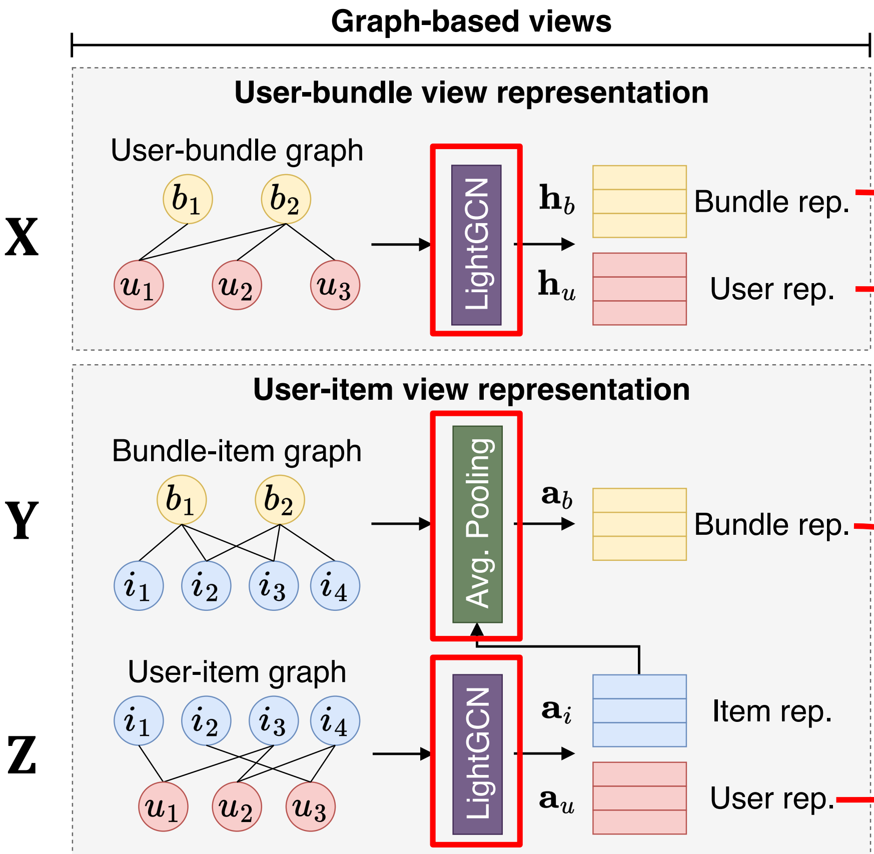
(a) Youshu (b) NetEase (c) iFashion

Outline

- Introduction
- **Proposed Method**
- Experiments
- Conclusion

Graph-based Views

- Form two graph-based views to fully utilize the given matrices $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$, as in *CrossCBR*



$$\mathbf{h}_u^{(k)} = \sum_{b \in \mathcal{N}_u} \frac{1}{\sqrt{|\mathcal{N}_u|}\sqrt{|\mathcal{N}_b|}} \mathbf{h}_b^{(k-1)}, \mathbf{h}_b^{(k)} = \sum_{u \in \mathcal{N}_b} \frac{1}{\sqrt{|\mathcal{N}_b|}\sqrt{|\mathcal{N}_u|}} \mathbf{h}_u^{(k-1)}$$

$$\mathbf{h}_u = \sum_{k=0}^{K} \frac{1}{k+1} \mathbf{h}_u^{(k)}, \mathbf{h}_b = \sum_{k=0}^{K} \frac{1}{k+1} \mathbf{h}_b^{(k)}$$

$$h_{ub} = \mathbf{h}_u^\top \mathbf{h}_b$$

$$\mathbf{a}_u^{(k)} = \sum_{i \in \mathcal{N}'_u} \frac{1}{\sqrt{|\mathcal{N}'_u|}\sqrt{|\mathcal{N}_i|}} \mathbf{a}_i^{(k-1)}, \mathbf{a}_i^{(k)} = \sum_{u \in \mathcal{N}_i} \frac{1}{\sqrt{|\mathcal{N}_i|}\sqrt{|\mathcal{N}'_u|}} \mathbf{a}_u^{(k-1)}$$

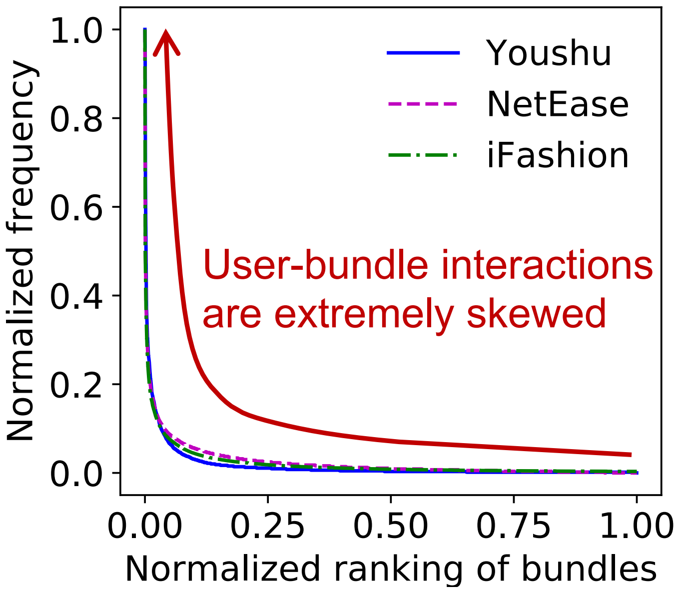
$$\mathbf{a}_u = \sum_{k=0}^{K} \frac{1}{k+1} \mathbf{a}_u^{(k)}, \mathbf{a}_i = \sum_{k=0}^{K} \frac{1}{k+1} \mathbf{a}_i^{(k)}$$

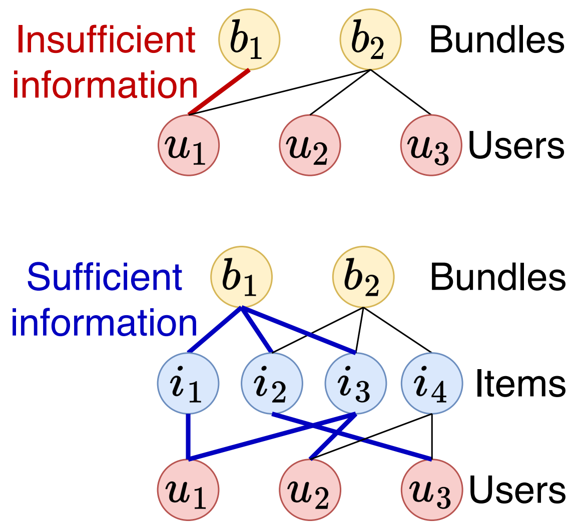
$$\mathbf{a}_b = \frac{1}{|\mathcal{N}'_b|} \sum_{i \in \mathcal{N}'_b} \mathbf{a}_i$$

$$a_{ub} = \mathbf{a}_u^\top \mathbf{a}_b$$

Popularity-based Coalescence (1/2)

- Real-world interactions are **<u>extremely skewed</u>**
  - Unpopular bundles are underrepresented in user-bundle (U-B) view
- User-item (U-I) view is more informative for them





Popularity-based Coalescence (2/2)

- Weighted strategy based on **bundle popularity** to estimate the relevance scores

$$\hat{y}_{ub} = \gamma_b h_{ub} + (1 - \gamma_b) a_{ub} \qquad \text{where } \gamma_b \in [0, 1]$$

$$\gamma_b > \gamma_{b'} \text{ if } n_b > n_{b'}$$

n_b: number of user interactions of bundle b

  - Mitigates the influence of U-B view for unpopular bundles
  - Instead, amplifies the influence of U-I view

$$\frac{\partial \hat{y}_{ub}}{\partial h_{ub}} < \frac{\partial \hat{y}_{ub'}}{\partial h_{ub'}} \text{ if } n_b < n_{b'} \qquad \frac{\partial \hat{y}_{ub}}{\partial a_{ub}} > \frac{\partial \hat{y}_{ub'}}{\partial a_{ub'}} \text{ if } n_b < n_{b'}$$

Curriculum Heating

- Cold bundles require well-represented embeddings from U-I view
  - U-I view is more difficult to learn than U-B view, as it entails understanding all affiliated items of a bundle
- Adopt a **curriculum learning** strategy
  - Focus initially on learning *easier* (U-B) view and gradually shift the focus to learning *harder* (U-I) view

$$\hat{y}_{ub}^{(t)} = \gamma_b^{(t)} h_{ub} + (1 - \gamma_b^{(t)}) a_{ub} \qquad \gamma_b^{(t)} = \tanh\left(\frac{n_b}{\psi^{(t)}}\right)$$

$$\psi^{(t)} = \epsilon^{t/T}, t : 0 \rightarrow T, \qquad \epsilon > 1$$

$$\frac{\partial \hat{y}_{ub}^{(t)}}{\partial h_{ub}} < \frac{\partial \hat{y}_{ub'}^{(t)}}{\partial h_{ub'}} \text{ if } n_b < n_{b'}$$

$$\frac{\partial \hat{y}_{ub}^{(t)}}{\partial a_{ub}} > \frac{\partial \hat{y}_{ub'}^{(t)}}{\partial a_{ub'}} \text{ if } n_b < n_{b'}$$

Alignment and Uniformity

- Aligning two views is essential, especially for cold bundles
- Adopt an **alignment and uniformity** (a sort of contrastive learning)

l_2-normalization

$$\tilde{\mathbf{h}}_u = \frac{\mathbf{h}_u}{\|\mathbf{h}_u\|_2}, \tilde{\mathbf{a}}_u = \frac{\mathbf{a}_u}{\|\mathbf{a}_u\|_2}, \tilde{\mathbf{h}}_b = \frac{\mathbf{h}_b}{\|\mathbf{h}_b\|_2}, \tilde{\mathbf{a}}_b = \frac{\mathbf{a}_b}{\|\mathbf{a}_b\|_2}$$

Alignment

$$l_{align} = \mathop{\mathbb{E}}_{u \sim p_{user}} \|\tilde{\mathbf{h}}_u - \tilde{\mathbf{a}}_u\|_2^2 + \mathop{\mathbb{E}}_{b \sim p_{bundle}} \|\tilde{\mathbf{h}}_b - \tilde{\mathbf{a}}_b\|_2^2$$

Uniformity

$$l_{uniform} = \log \mathop{\mathbb{E}}_{u,u' \sim p_{user}} e^{-2\|\tilde{\mathbf{h}}_u - \tilde{\mathbf{h}}_{u'}\|_2^2} + \log \mathop{\mathbb{E}}_{u,u' \sim p_{user}} e^{-2\|\tilde{\mathbf{a}}_u - \tilde{\mathbf{a}}_{u'}\|_2^2}$$
$$+ \log \mathop{\mathbb{E}}_{b,b' \sim p_{bundle}} e^{-2\|\tilde{\mathbf{h}}_b - \tilde{\mathbf{h}}_{b'}\|_2^2} + \log \mathop{\mathbb{E}}_{b,b' \sim p_{bundle}} e^{-2\|\tilde{\mathbf{a}}_b - \tilde{\mathbf{a}}_{b'}\|_2^2}$$

Contrastive loss

$$\mathcal{L}_{AU} = l_{align} + l_{uniform}$$

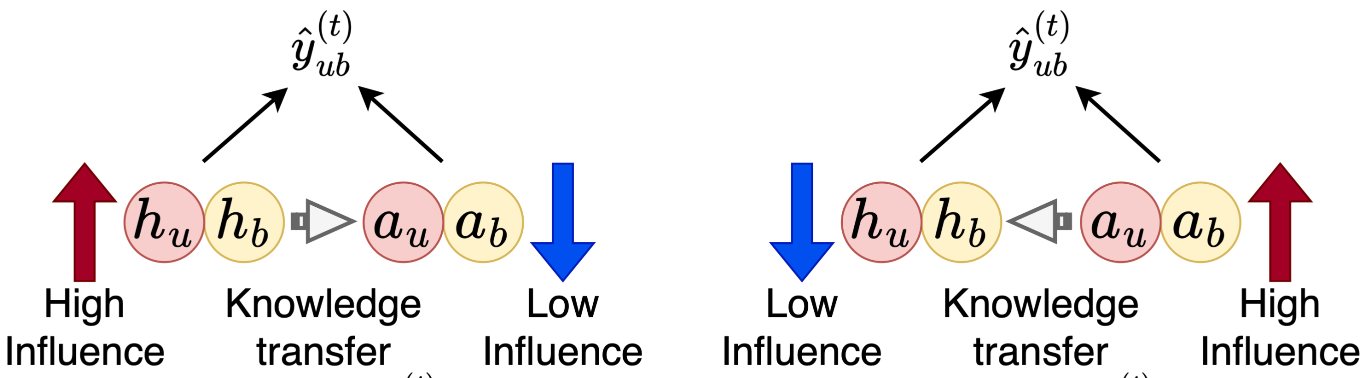
Objective and Dynamics

- Objective function of CoHeat

$$\mathcal{L}_{BPR}^{(t)} = \mathop{\mathbb{E}}_{(u,b^+,b^-)\sim p_{data}} -\ln\sigma(\hat{y}_{ub^+}^{(t)} - \hat{y}_{ub^-}^{(t)})$$

$$\mathcal{L}^{(t)} = \mathcal{L}_{BPR}^{(t)} + \lambda_1 \mathcal{L}_{AU} + \lambda_2 \|\Theta\|_2$$

- Dynamic mechanism of CoHeat



(a) Large $\gamma_b^{(t)}$ **(b) Small** $\gamma_b^{(t)}$

Outline

- Introduction
- Proposed Method
- **<u>Experiments</u>**
- Conclusion

Research Questions

- **Q1. Comparison w/ cold-start methods**
  - Does CoHeat show higher performance than other cold-start methods?
- **Q2. Comparison w/o warm-start methods**
  - Does CoHeat show comparable performance to other warm-start methods?
- **Q3. Comparison by cold bundle ratio**
  - How does the performance disparity change depending on the cold bundle ratio?

Datasets

- We use three real-world datasets

Dataset	Users	Bundles	Items	User-bundle (dens.)	User-item (dens.)	Bundle-item (dens.)	Avg. size of bundle
Youshu[1]	8,039	4,771	32,770	51,377 (0.13%)	138,515 (0.05%)	176,667 (0.11%)	37.03
NetEase[1]	18,528	22,864	123,628	302,303 (0.07%)	1,128,065 (0.05%)	1,778,838 (0.06%)	77.80
iFashion[1]	53,897	27,694	42,563	1,679,708 (0.11%)	2,290,645 (0.10%)	106,916 (0.01%)	3.86

  - *Youshu*: book review platform
  - *NetEase*: cloud music platform
  - *iFashion*: outfit sales platform

Baselines

- Cold-start methods
  - *DropoutNet*: robustness-based method
  - *CB2CF / Heater*: constraint-based method
  - *GAR-CF / GAR-GNN / CVAR*: generative method
  - *CLCRec / CCFCRec*: contrastive method
- Warm-start methods
  - *MFBPR / LightGCN / SGL / SimGCL / LightGCL*: item recommendation method
  - *DAM / BundleNet / BGCN / CrossCBR*: bundle recommendation method

Evaluation

- Three scenarios
  - *Cold-start*: bundles are randomly split
  - *Warm-start*: interactions are randomly split
  - *All-bundle*: half-n-half of warm/cold-start
- Metrics
  - *Recall@20*
  - *nDCG@20*

Q1. Comparison w/ Cold-start Methods

- CoHeat **outperforms** the competitors
  - Cold-start / warm-start / all-bundle scenarios

Model	Youshu						NetEase						iFashion					
	Recall@20			nDCG@20			Recall@20			nDCG@20			Recall@20			nDCG@20		
	Cold	*Warm*	*All*	*Cold*	*Warm*	*All*	*Cold*	*Warm*	*All*	*Cold*	*Warm*	*All*	*Cold*	*Warm*	*All*	*Cold*	*Warm*	*All*
DropoutNet [51]	.0022	.0336	.0148	.0007	.0153	.0055	.0028	.0154	.0046	.0015	.0078	.0024	.0009	.0060	.0039	.0008	.0045	.0027
CB2CF [2]	.0012	.0258	.0028	.0007	.0208	.0021	.0016	.0049	.0027	.0006	.0027	.0014	.0009	.0057	.0066	.0006	.0043	.0048
Heater [73]	.0016	.1753	.0541	.0007	.0826	.0286	.0021	.0125	.0102	.0010	.0064	.0054	.0015	.0217	.0123	.0010	.0151	.0083
GAR-CF [11]	.0015	.1688	.0529	.0011	.0726	.0317	.0010	.0063	.0014	.0005	.0035	.0008	.0013	.0203	.0090	.0013	.0143	.0055
GAR-GNN [11]	.0013	.0835	.0358	.0006	.0569	.0178	.0009	.0056	.0027	.0003	.0030	.0012	.0065	.0172	.0126	.0030	.0107	.0087
CVAR [70]	.0008	.1958	.0829	.0002	.1112	.0533	.0002	.0308	.0156	.0001	.0154	.0084	.0007	.0220	.0125	.0004	.0152	.0084
CLCRec [58]	.0137	.0626	.0367	.0087	.0317	.0194	.0136	.0407	.0259	.0075	.0215	.0138	.0053	.0203	.0126	.0043	.0135	.0085
CCFCRec [72]	.0044	.1554	.0702	.0022	.0798	.0425	.0007	.0265	.0130	.0004	.0128	.0068	.0005	.0439	.0252	.0003	.0304	.0172
CoHeat (ours)	**.0183**	**.2804**	**.1247**	**.0105**	**.1646**	**.0833**	**.0191**	**.0847**	**.0453**	**.0093**	**.0455**	**.0264**	**.0170**	**.1156**	**.0658**	**.0096**	**.0876**	**.0504**

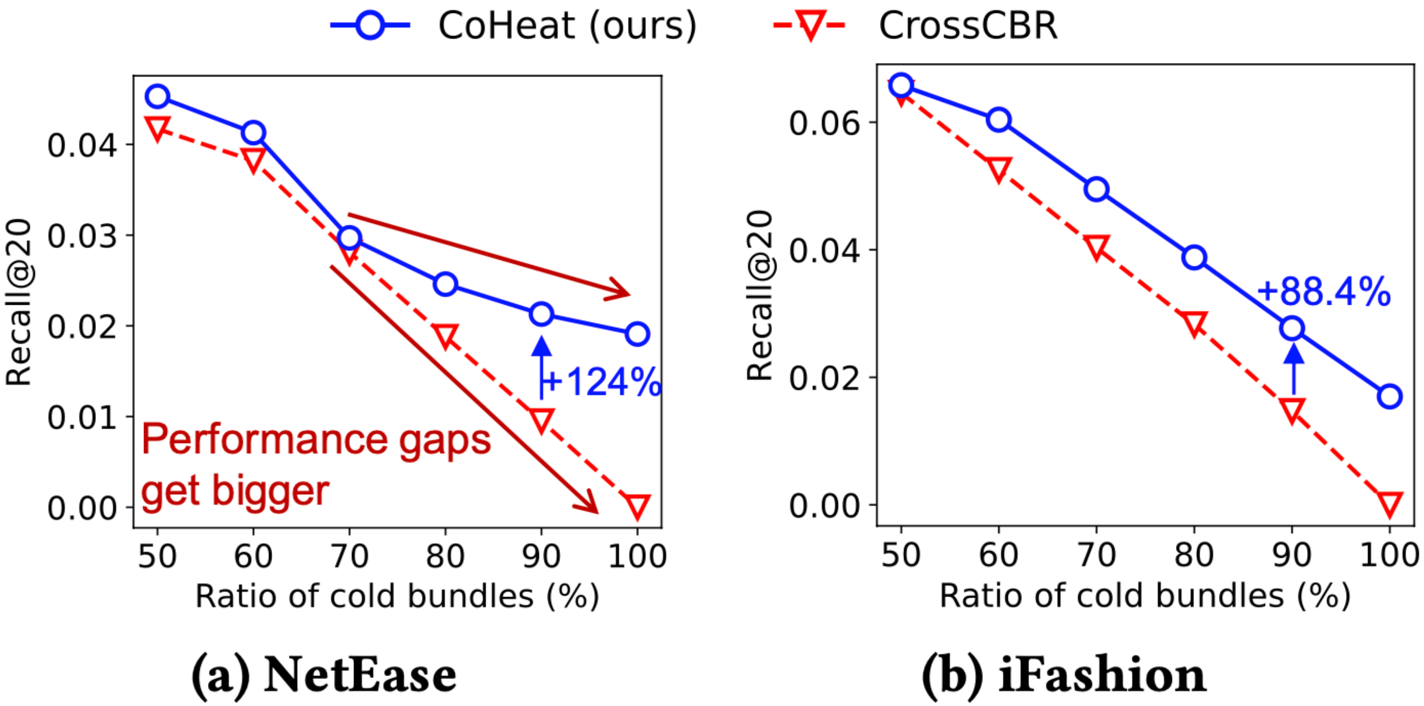
Q2. Comparison w/ Warm-start Methods

- CoHeat shows **comparable performance** to the baselines (even slightly higher perf.)
  - Warm-start scenario

Model	Youshu Recall @20	Youshu nDCG @20	NetEase Recall @20	NetEase nDCG @20	iFashion Recall @20	iFashion nDCG @20
MFBPR [42]	.1959	.1117	.0355	.0181	.0752	.0542
LightGCN [23]	.2286	.1344	.0496	.0254	.0837	.0612
SGL [59]	.2568	.1527	.0687	.0368	.0933	.0690
SimGCL [62]	.2691	.1593	.0710	.0377	.0919	.0677
LightGCL [5]	.2712	.1607	.0722	.0388	.0943	.0686
DAM [12]	.2082	.1198	.0411	.0210	.0629	.0450
BundleNet [17]	.1895	.1125	.0391	.0201	.0626	.0447
BGCN [8, 9]	.2347	.1345	.0491	.0258	.0733	.0531
CrossCBR [35]	.2776	.1641	.0791	.0433	.1133	.0875
CoHeat (ours)	**.2804**	**.1646**	**.0847**	**.0455**	**.1156**	**.0876**

Q3. Comparison by Cold Bundle Ratio

- **Performance disparity widens** as cold bundle ratios increase



Outline

- Introduction
- Proposed Method
- Experiments
- **Conclusion**

Conclusion

- Address the cold-start problem in bundle recommendation
- Propose a novel framework: **CoHeat**
  - Popularity-based coalescence
  - Curriculum heating
  - Representation alignment and uniformity
- Show the superiority of CoHeat through extensive experiments

Thank you!

Code: https://github.com/snudatalab/CoHeat

Personal website: https://jeon185.github.io

Appendix: Ablation Study

- All main components help improve the performance
  - *CoHeat-PC*: equal contribution of U-B and U-I views
  - *CoHeat-CH-Ant*: anti-curriculum learning
  - *CoHeat-CH-Fix*: temperature fixed as the maximum
  - *CoHeat-AU*: Omit contrastive loss

Model	**Youshu** Recall @20	**Youshu** nDCG @20	**NetEase** Recall @20	**NetEase** nDCG @20	**iFashion** Recall @20	**iFashion** nDCG @20
CoHeat-*PC*	.0000	.0000	.0000	.0000	.0000	.0000
CoHeat-*CH-Ant*	.0177	.0087	.0176	.0087	.0164	.0093
CoHeat-*CH-Fix*	.0180	.0092	.0182	.0090	.0164	.0092
CoHeat-*AU*	.0069	.0031	.0029	.0013	.0013	.0005
CoHeat (ours)	**.0183**	**.0105**	**.0191**	**.0093**	**.0170**	**.0096**