School of
Information Systems

SMU
SINGAPORE MANAGEMENT UNIVERSITY

Modeling Contemporaneous Basket Sequences with Twin Networks for Next-Item Recommendation

Duc-Trong Le, Hady W. Lauw, Yuan Fang

IJCAI 2018

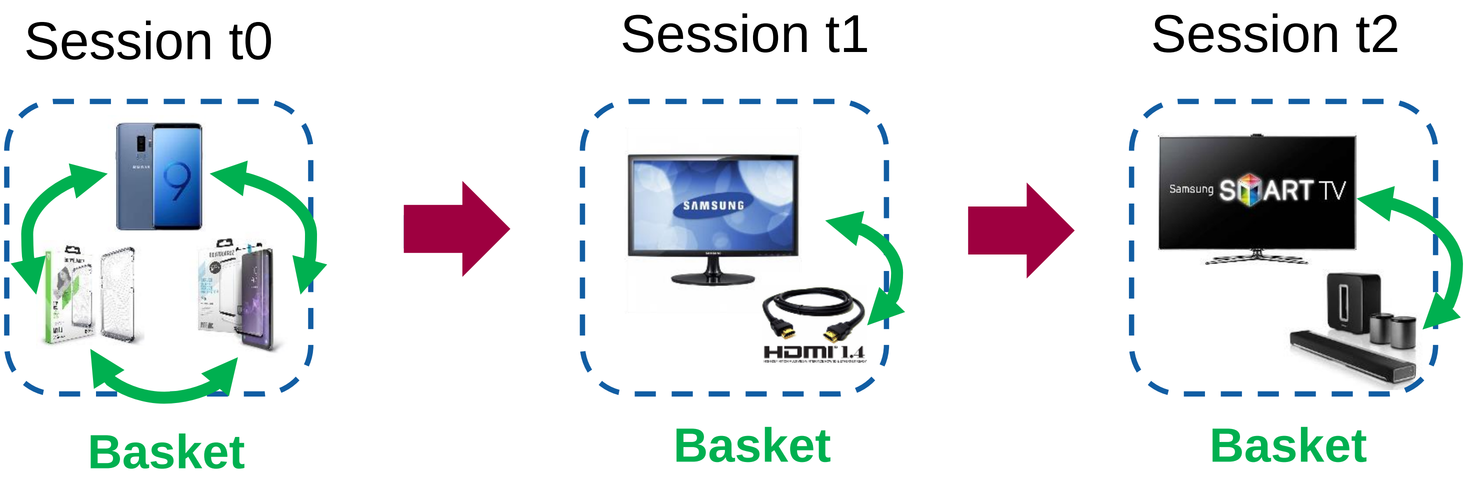
Stockholm, Sweden

Outline

- **Motivating examples**
- **Modeling Contemporaneous Basket Sequences (CBS)**
  - CBS with Siamese Networks (CBS-SN)
  - CBS with Concordant Fraternal Networks (CBS-CFN)
  - CBS with Discordant Fraternal Networks (CBS-DFN)
- **Experiments** on Alibaba, MovieLens-10M

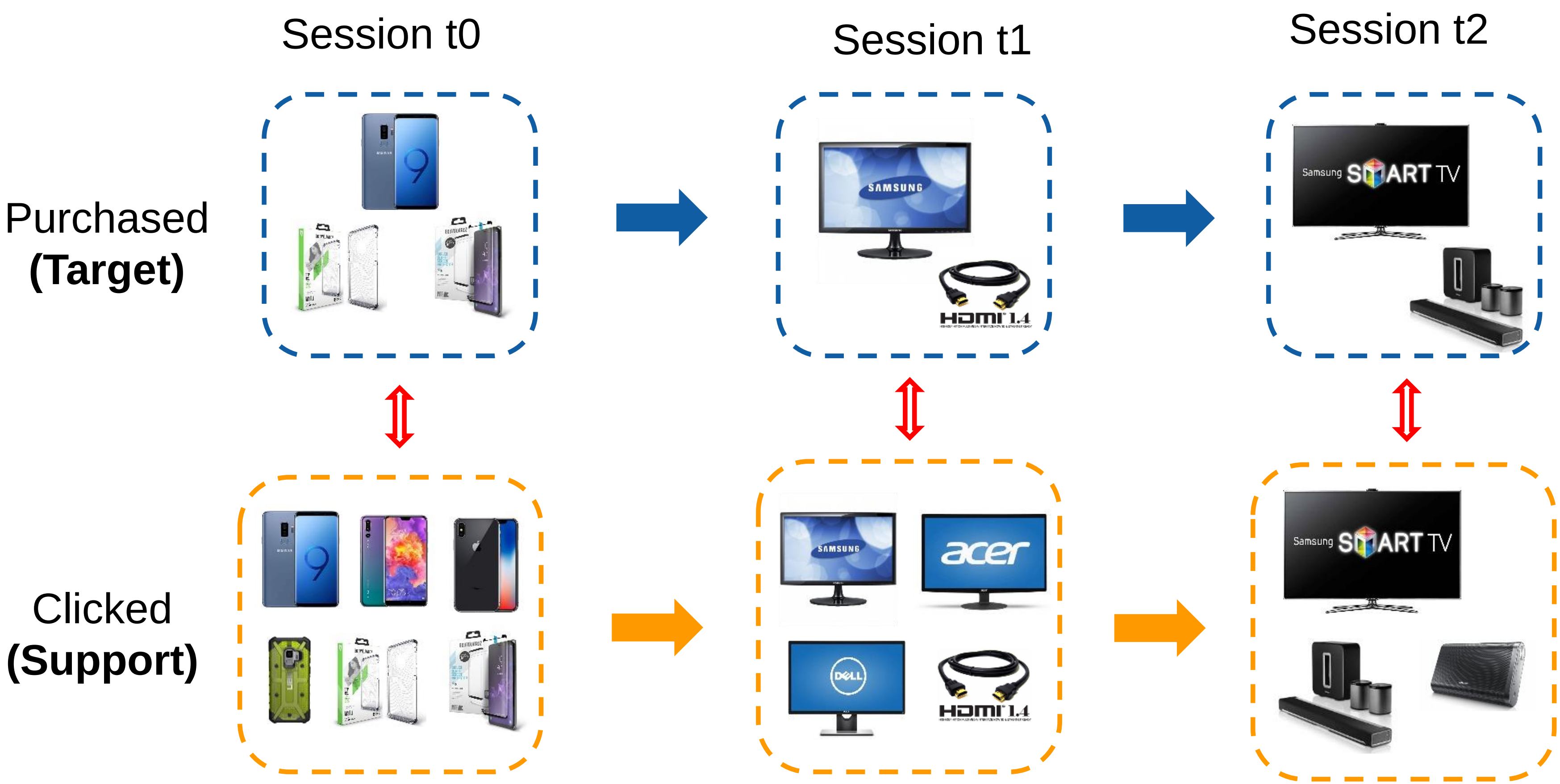
The Notion of Basket Sequence

Cross-session Purchased Products

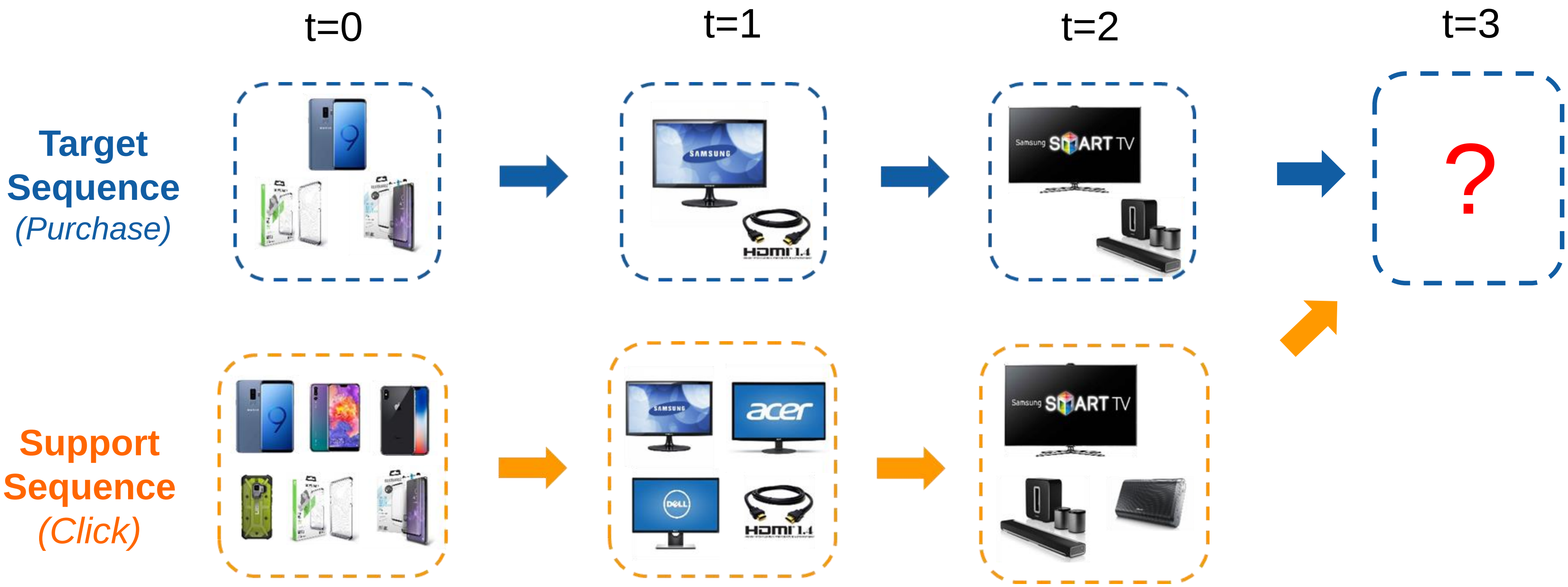


Correlative associations among items in a basket
Sequential associations across baskets in a sequence

The Notion of Contemporaneous Basket Sequences (CBS)



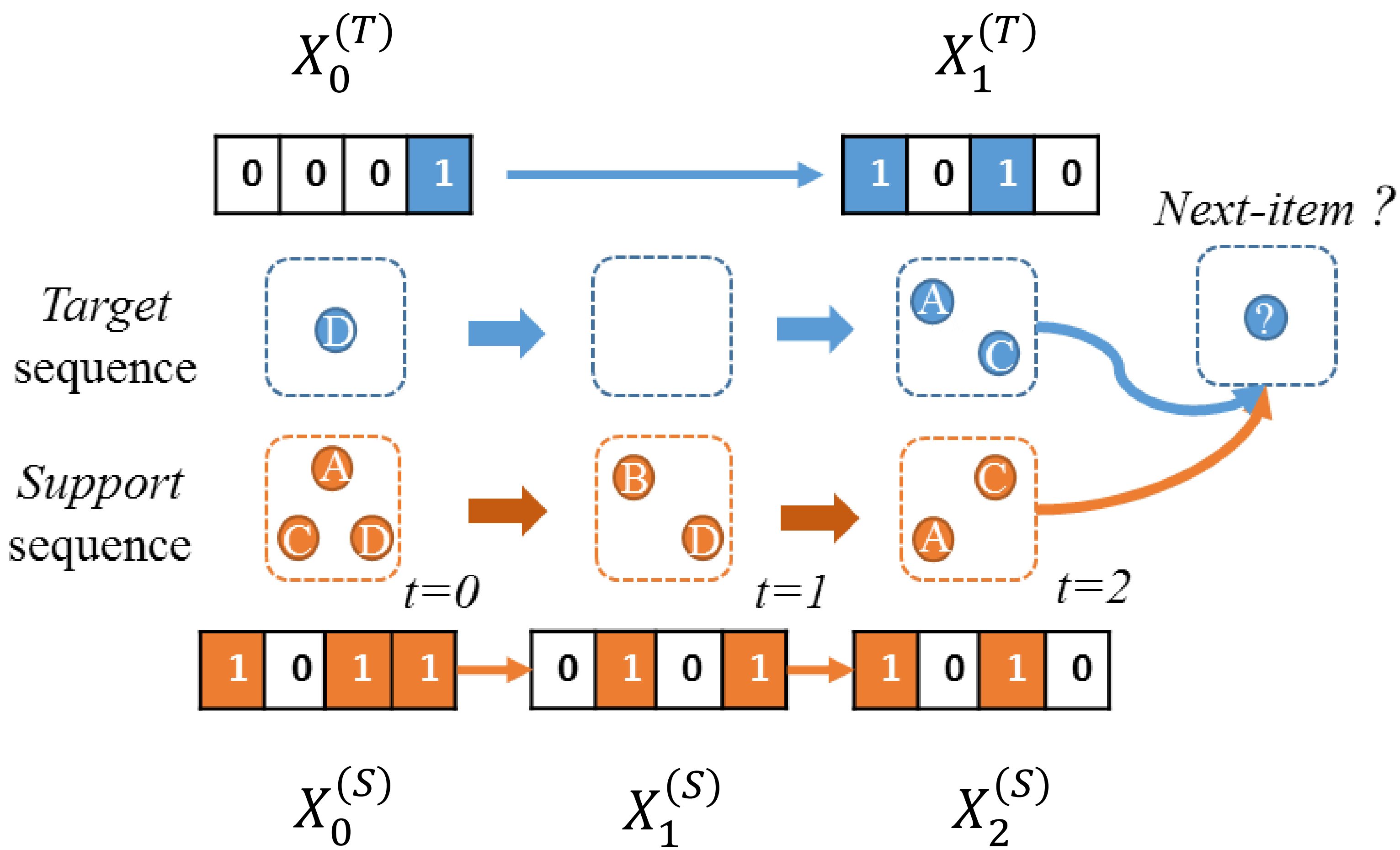
Next-Item Recommendation with CBS



TASK:

Modeling **correlative** & **sequential** associations in **CBS** concurrently to predict the **next "target" item**

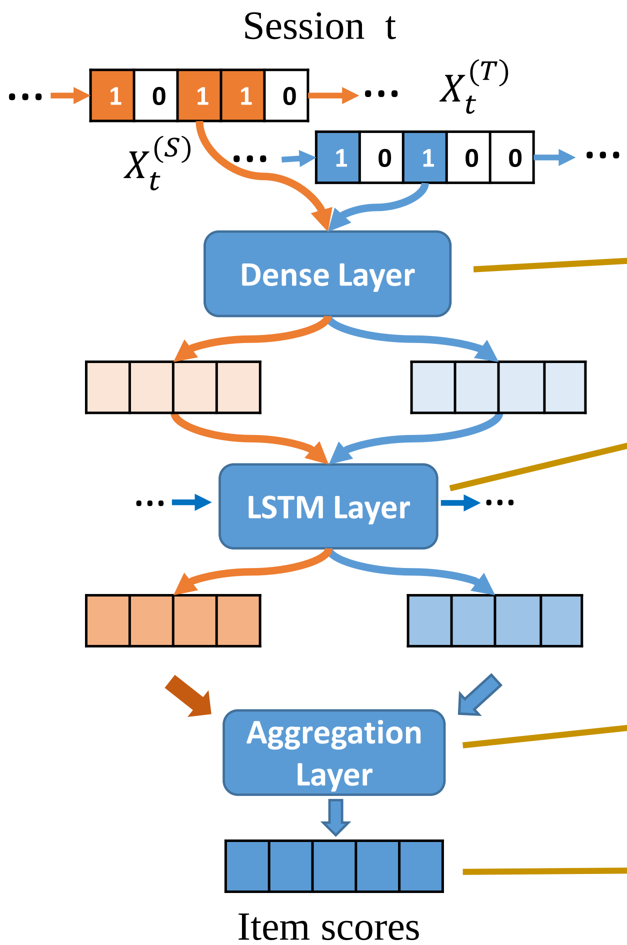
Next-Item Recommendation with CBS Binary Representation



Outline

- **Motivating examples**
- **Modeling Contemporaneous Basket Sequences (CBS)**
  - CBS with Siamese Networks (CBS-SN)
  - CBS with Concordant Fraternal Networks (CBS-CFN)
  - CBS with Discordant Fraternal Networks (CBS-DFN)
- **Experiments** on Alibaba, MovieLens-10M

CBS with Siamese Networks (CBS-SN)



- **Hypothesis**: CBS reflect the same underlying behavior
- **Formulation**: $X_t \in \{0,1\}^N$

$$b_t = f(\Theta_b X_t + \Omega_b)$$

$$\Theta_b \in \mathbb{R}^{L\times N}, \Omega_b \in \mathbb{R}^L$$

$$h_t = g(\Phi_b b_t + \Phi_h h_{t-1} + \Omega_h)$$

$$\Phi_b \in \mathbb{R}^{H\times L}, \Phi_h \in \mathbb{R}^{H\times H}, \Omega_h \in \mathbb{R}^H$$

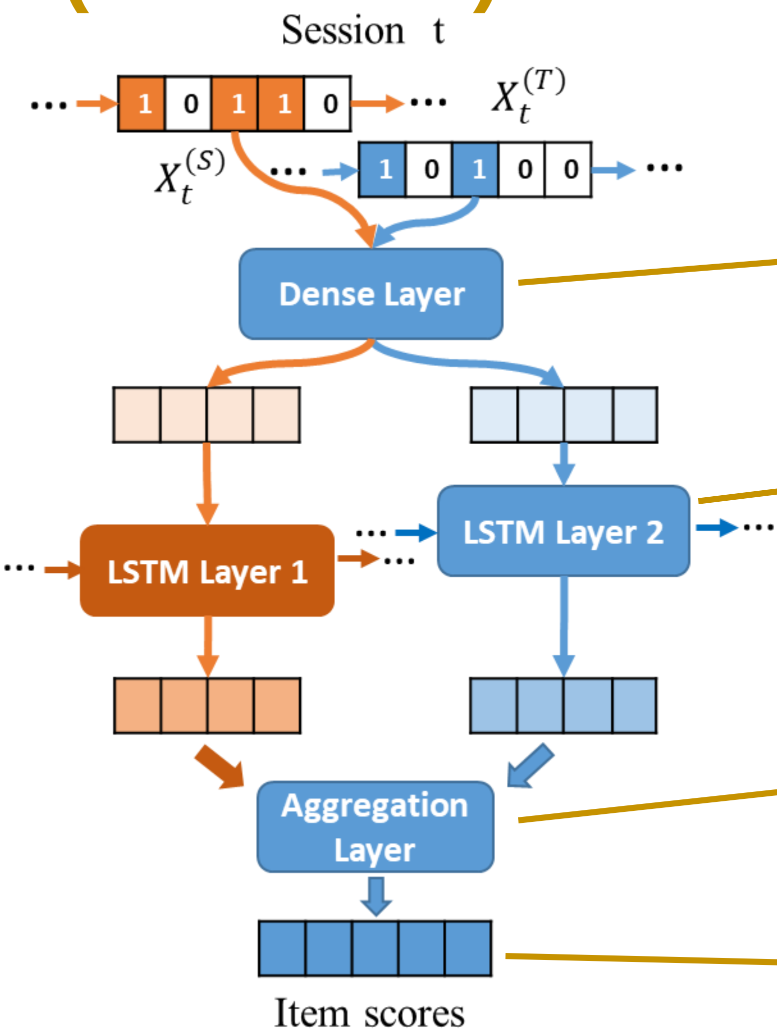
$$\hat{h} = concat(\hat{h}_T, \hat{h}_S)$$

$$r_{\text{agg}} = W.\hat{h} + \Omega_c$$

$$W \in \mathbb{R}^{N\times 2H}, \Omega_c \in \mathbb{R}^N$$

$$Y = softmax(r_{\text{agg}})$$

CBS with Concordant Fraternal Networks (CBS-CFN)



- **Hypothesis**: CBS reflect different long-term sequential dependencies
- **Formulation**: $X_t \in \{0,1\}^N$

$$b_t = f(\Theta_b X_t + \Omega_b)$$

$$\Theta_b \in \mathbb{R}^{L \times N}, \Omega_b \in \mathbb{R}^{L}$$

$$\boxed{\begin{aligned} h_t^{(S)} &= g(\Phi_b^{(S)} b_t^{(S)} + \Phi_h^{(S)} h_{t-1}^{(S)} + \Omega_h^{(S)}) \\ h_t^{(T)} &= g(\Phi_b^{(T)} b_t^{(T)} + \Phi_h^{(T)} h_{t-1}^{(T)} + \Omega_h^{(T)}) \end{aligned}}$$

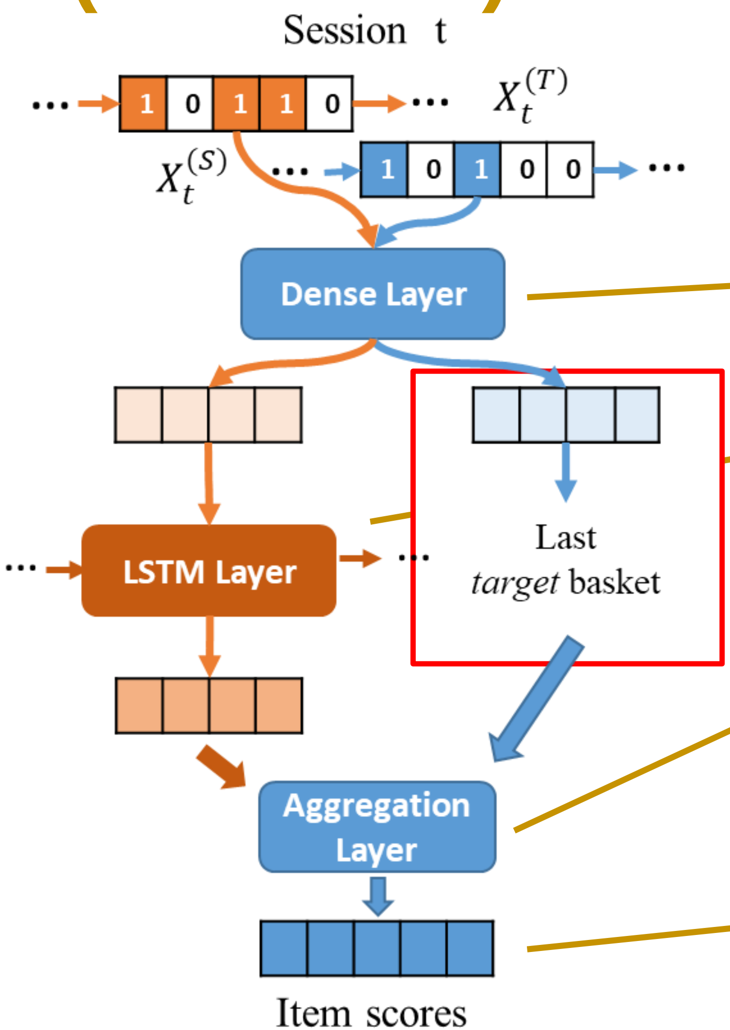
$$\Phi_b^{(*)} \in \mathbb{R}^{H \times L}, \Phi_h^{(*)} \in \mathbb{R}^{H \times H}, \Omega_h^{(*)} \in \mathbb{R}^{H}$$

$$\boxed{r_{\text{agg}} = W_1.\hat{h}_S + W_2.\hat{h}_T + \Omega_c}$$

$$W_1, W_2 \in \mathbb{R}^{N \times H}, \Omega_c \in \mathbb{R}^{N}$$

$$Y = softmax(r_{\text{agg}})$$

CBS with Discordant Fraternal Networks (CBS-DFN)



- **Hypothesis**: Short-term dependencies in "target" while long-term in "support"
- **Formulation**: $X_t \in \{0,1\}^N$

$$b_t = f(\Theta_b X_t + \Omega_b)$$

$$\Theta_b \in \mathbb{R}^{L \times N}, \Omega_b \in \mathbb{R}^L$$

$$h_t = g(\Phi_b b_t + \Phi_h h_{t-1} + \Omega_h)$$

$$\Phi_b \in \mathbb{R}^{H \times L}, \Phi_h \in \mathbb{R}^{H \times H}, \Omega_h \in \mathbb{R}^H$$

$$\boxed{r_{agg} = W_1 . \hat{h}_S + W_2 \hat{b}_T + \Omega_c}$$

$$W_1 \in \mathbb{R}^{N \times H}, W_2 \in \mathbb{R}^{N \times L}, \Omega_c \in \mathbb{R}^N$$

$$Y = softmax(r_{agg})$$

Outline

- **Motivating examples**
- **Modeling Contemporaneous Basket Sequences (CBS)**
  - CBS with Siamese Networks (CBS-SN)
  - CBS with Concordant Fraternal Networks (CBS-CFN)
  - CBS with Discordant Fraternal Networks (CBS-DFN)
- **Experiments** on Alibaba, Movielens-10M

Experimental Setup

- **Task**: Next item recommendation
  - For each testing sequence pair $< S, T >$
  - Hide the last target basket B to create $|B|$ testing instances.
  - Require each model to product a ranked list of candidates based on the testing instances with ground truth items.
- **Datasets**

Dataset		#Sequence	#Item	#Average Length	#Average Basket Size
Alibaba (E-commerce)	Support	23740	13498	11.2	5.5
	Target			5.3	1.8
MovieLens-10M (Movie Rating)	Support	189858	8202	34.5	2.5
	Target			16.6	1.8

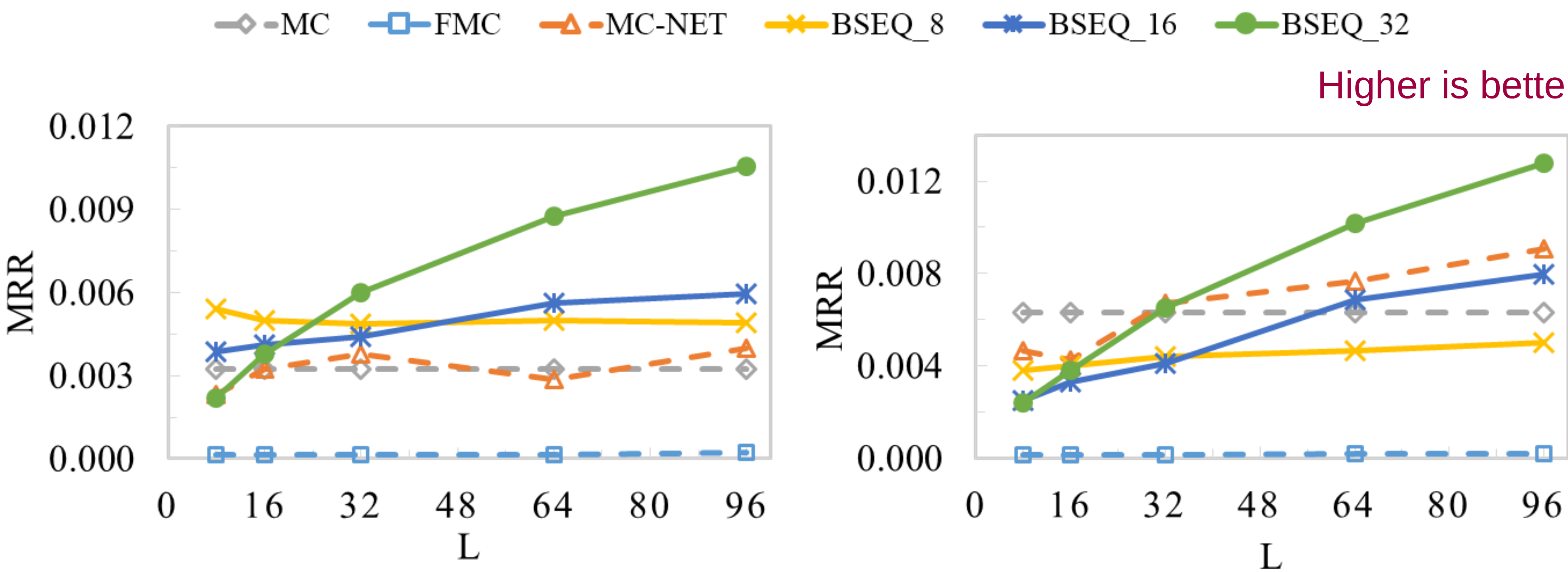
Pre-processing: Filter out too few items; sequences < 2 baskets.

- **Metric**: *Mean Reciprocal Rank*

$$MRR = \frac{\sum_B \sum_{v \in B} \frac{1}{\text{rank of } v \text{ for } (S, T \backslash \mathrm{B})}}{\#\text{total testing instances}}$$

RQ1: Is modelling sequential data useful?

Higher is better

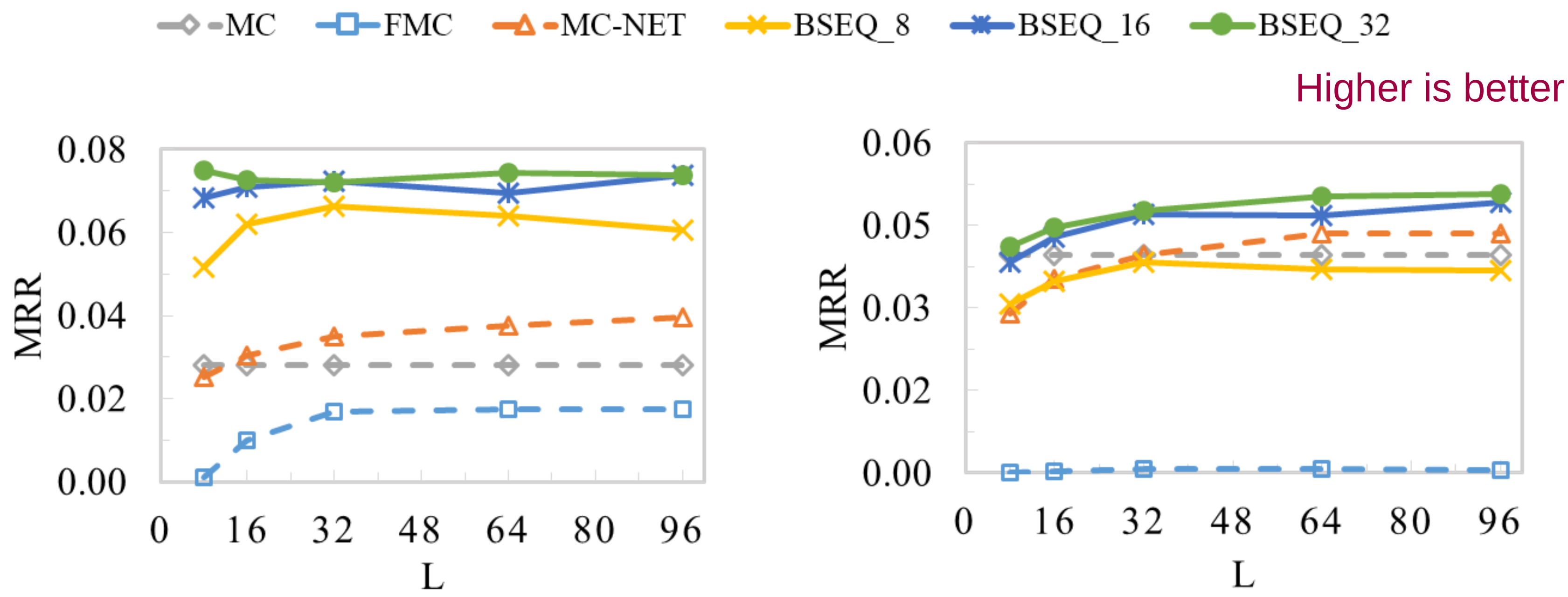


Alibaba –
Support Sequence Only

Alibaba –
Target Sequence Only

L is the latent dimension of the Dense Layer
H is the LSTM hidden state size ($BSEQ\_8 \rightarrow H = 8$)

RQ1: Is modelling sequential data useful?

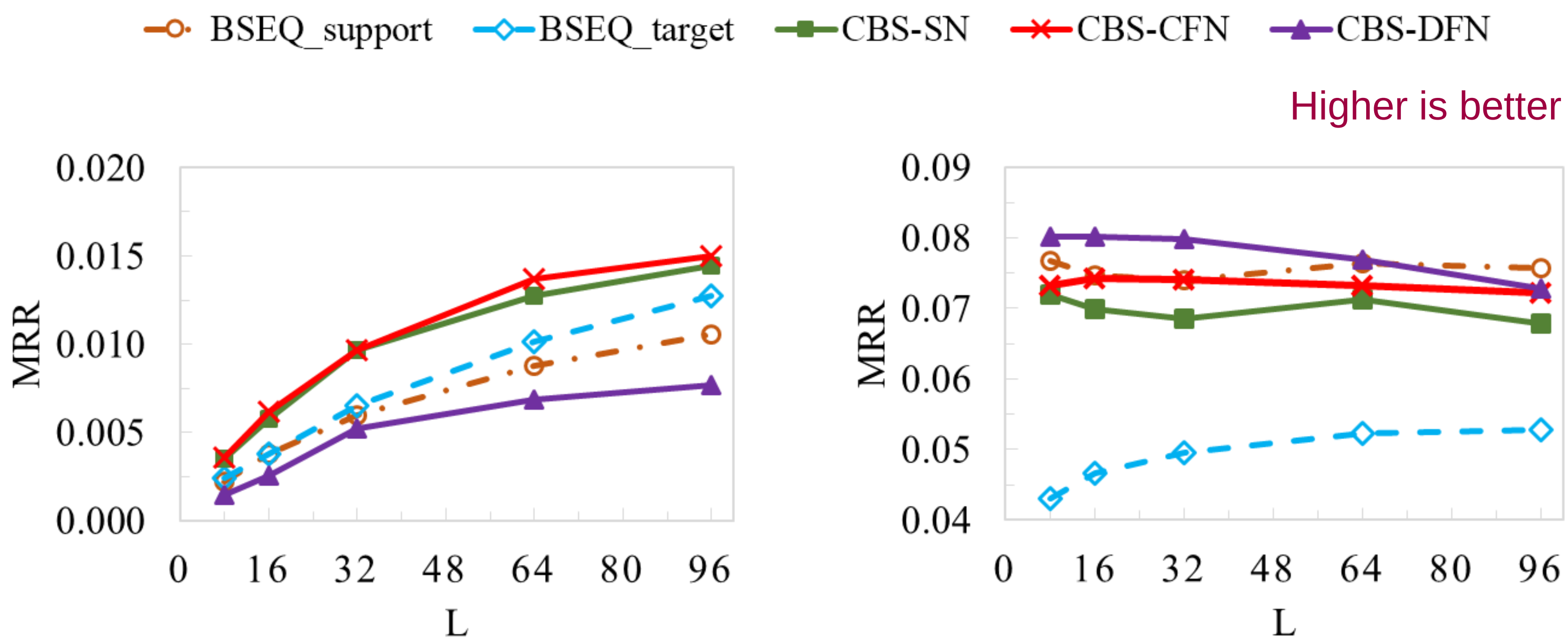


MovieLens – 10M
Support Sequence Only

MovieLens–10M
Target Sequence Only

L is the latent dimension of the Dense Layer
H is the LSTM hidden state size ($BSEQ\_8 \rightarrow H = 8$)

RQ2: Is modelling CBS useful?



Alibaba **MovieLens-10M**

L is the latent dimension of the Dense Layer
LSTM Hidden State Size H = 32

School of
Information Systems

RQ3: Do CBS Twin Networks outperform against other baselines?

LSTM Hidden State Size H = 32

Higher is better

Model	L	MRR
POP	-	0.004
$\text{DRM}_{support}$	64	0.011
DRM_{target}	32	0.004
$\text{BSEQ}_{support}$	96	0.011
BSEQ_{target}	96	0.013
CBS-SN	96	0.014
CBS-CFN	96	$\mathbf{0.015}^{\ddagger\S}$
CBS-DFN	96	0.008

Alibaba

Model	L	MRR
POP	-	0.006
$\text{DRM}_{support}$	96	0.002
DRM_{target}	16	0.001
$\text{BSEQ}_{support}$	8	0.075
BSEQ_{target}	64	0.050
CBS-SN	8	0.070
CBS-CFN	32	0.072
CBS-DFN	8	$\mathbf{0.078}^{\ddagger\S}$

MovieLens -10M

POP: The Popular-based recommendation method
DRM: The dynamic recurrent model with basket sequences

Conclusion

- Modeling **Contemporaneous Basket Sequences** in predicting the next-item adoption.

- Propose three Twin Network Structures (**CBS-SN, CBS-CFN** & **CBS-DFN**) contribute **statistically significant** improvements over:
  - **Single Basket Sequence Models (DRM, BSEQ)**
  - **Markov Chain-based Models (MC, FMC, MC-NET)**

  in term of top-K recommendations.

Thank you for listening!
Q&A