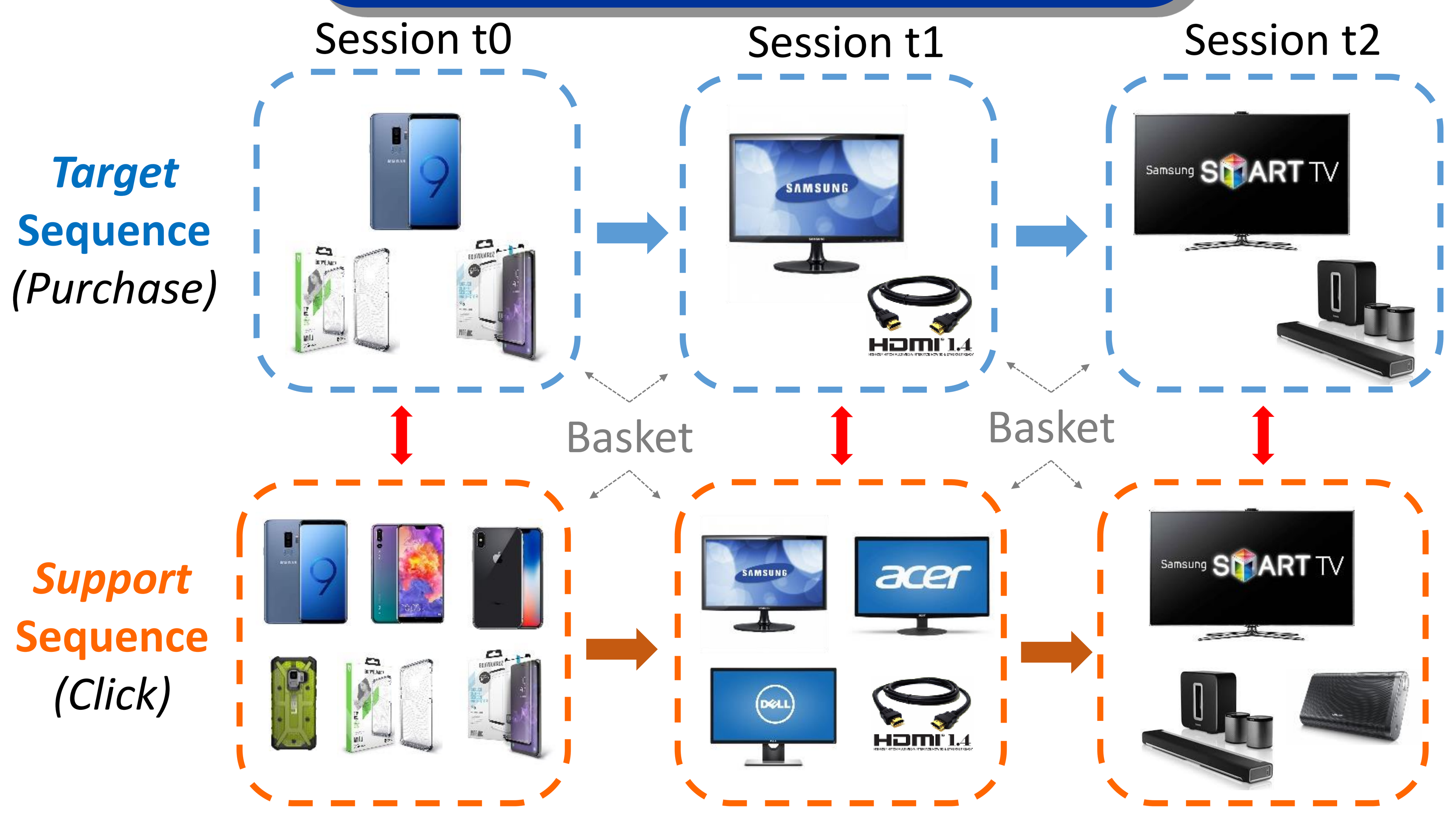
Modeling Contemporaneous Basket Sequences with Twin Networks for Next-Item Recommendation

SMU SINGAPORE MANAGEMENT UNIVERSITY
School of Information Systems

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

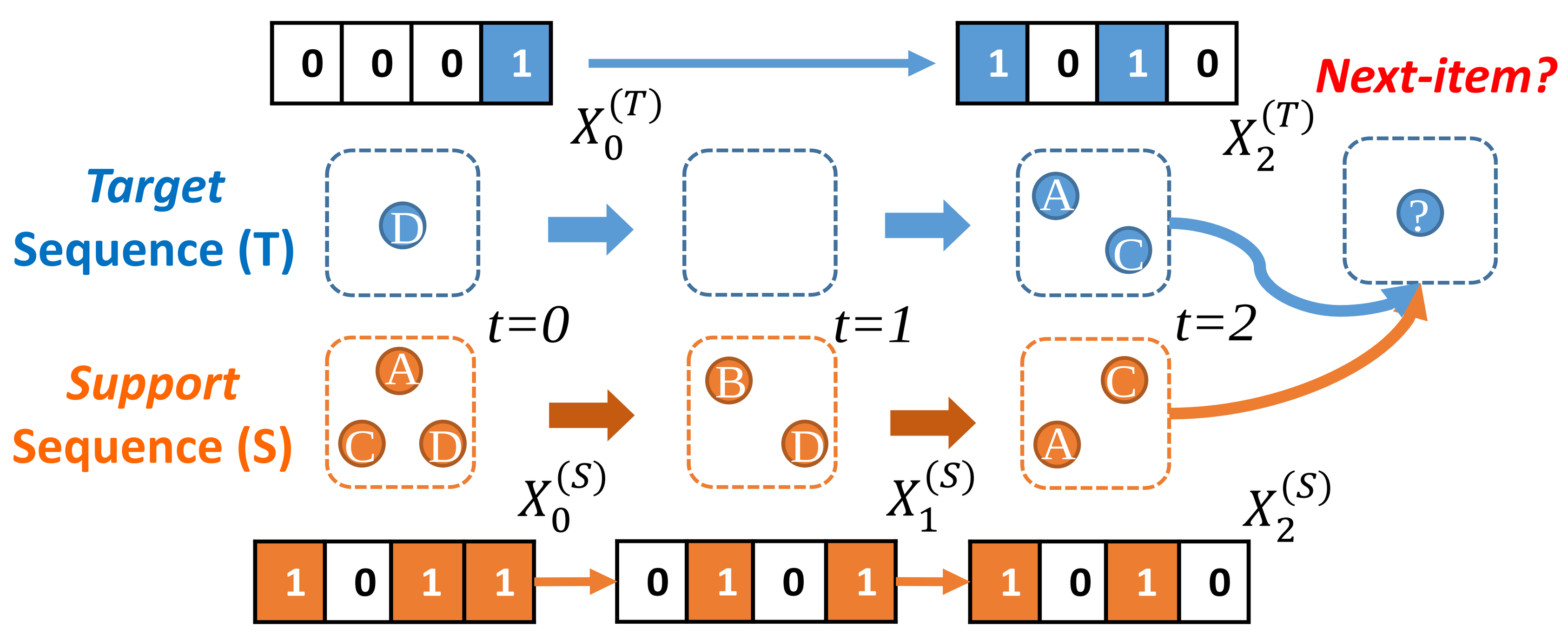
Contemporaneous Basket Sequences (CBS)



Associations:

- ***Correlative*** associations among items of a basket
- ***Sequential*** associations across baskets of a sequence

Problem



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

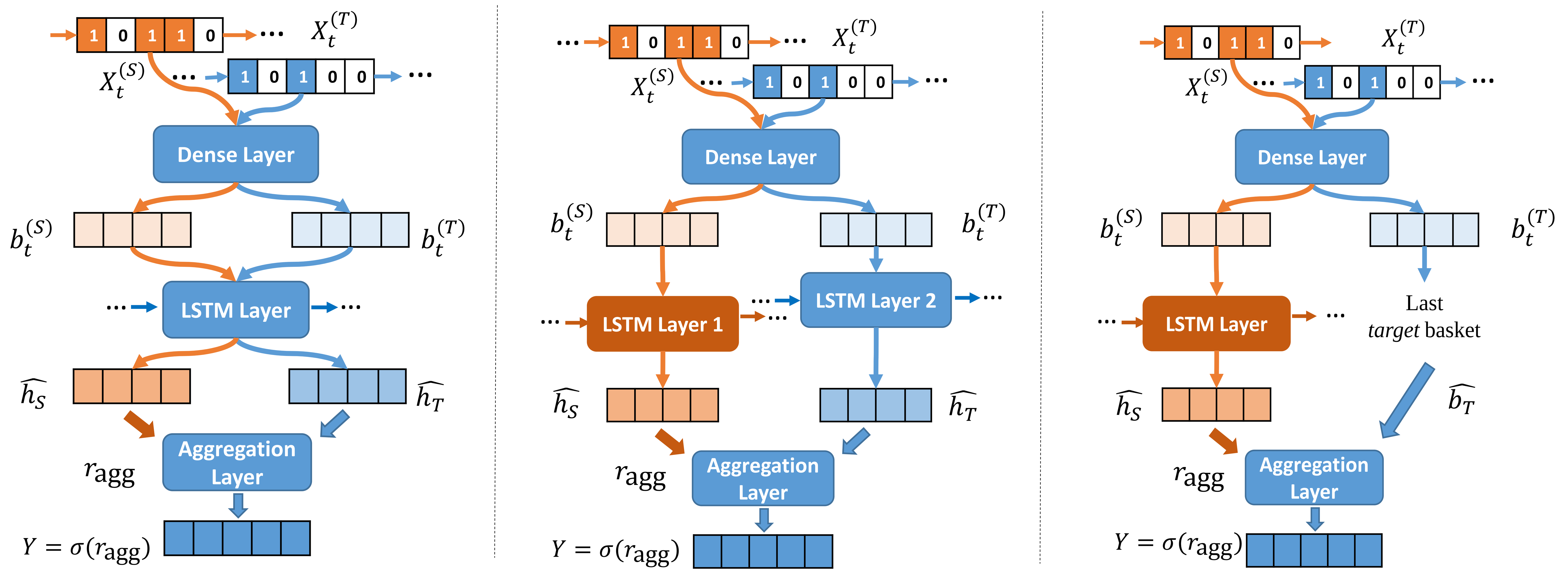
Solution: Given a contemporaneous pair (T, S), estimate conditional probabilities to rank all candidate items v_i:

$$Y_i = P(v_i|T,S) = F\left(v_i, X^{(T)}, X^{(S)}; \Theta\right)$$

where F is a real-valued function, e.g., softmax σ

Modeling CBS with Twin Networks

- **Dense Layer**: $b_t = f(\Theta_b X_t + \Omega_b)$; $\Theta_b \in \mathbb{R}^{L \times N}, \Omega_b \in \mathbb{R}^{L}$
- **LSTM Layer**: $h_t = g(\Phi_b b_t + \Phi_h h_{t-1} + \Omega_h)$; $\Phi_b \in \mathbb{R}^{H \times L}, \Phi_h \in \mathbb{R}^{H \times H}, \Omega_h \in \mathbb{R}^{H}$



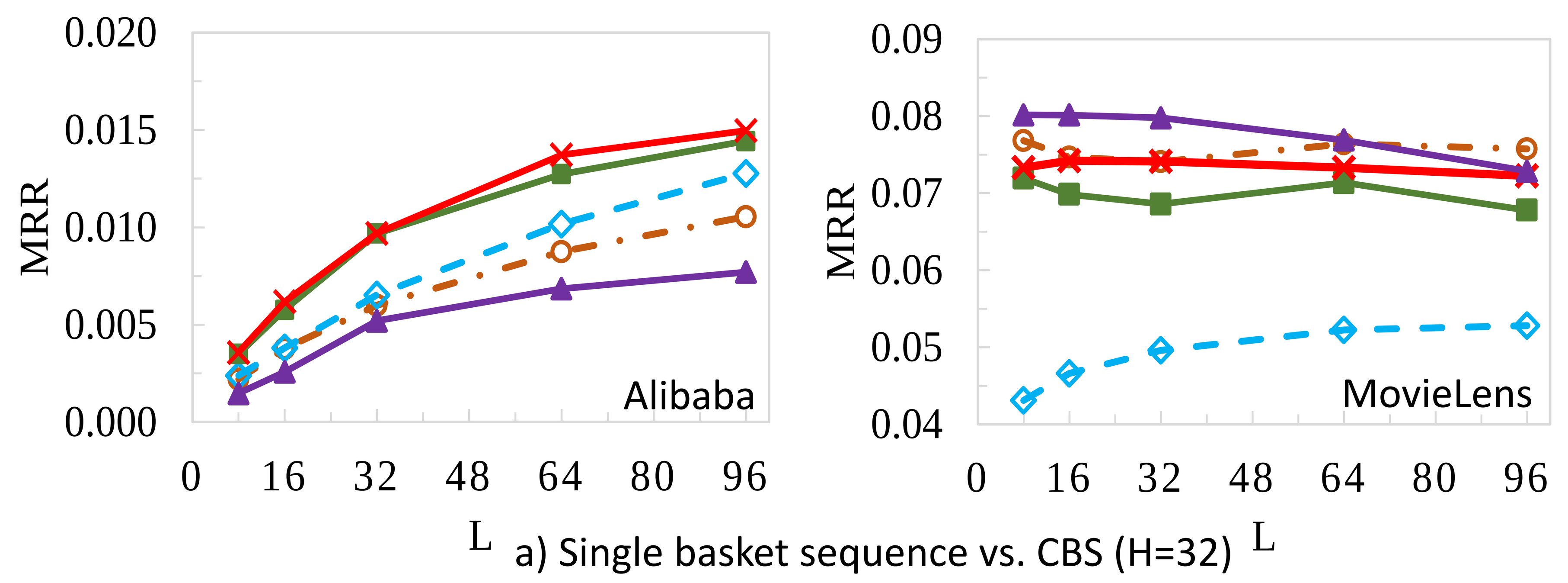
	Siamese Network (CBS-SN)	*Concordant Fraternal Network (CBS-CFN)*	*Discordant Fraternal Network (CBS-DFN)*
Hypothesis	Same underlying behavior in CBS	Different long-term dependencies in CBS	Short-term in *Target* & Long-term in *Support*
Aggregation	$r_{agg} = W.\text{concate}(\widehat{h_T}, \widehat{h_S}) + \Omega_r$ $W \in \mathbb{R}^{N \times 2H}, \Omega_r \in \mathbb{R}^{N}$	$r_{agg} = W_1.\widehat{h_T} + W_2.\widehat{h_S} + \Omega_r$ $W_1, W_2 \in \mathbb{R}^{N \times H}, \Omega_r \in \mathbb{R}^{N}$	$r_{agg} = W_1.\widehat{h_S} + W_2.\widehat{b_T} + \Omega_r$ $W_1 \in \mathbb{R}^{N \times H}, W_2 \in \mathbb{R}^{N \times L}, \Omega_r \in \mathbb{R}^{N}$

Experiments

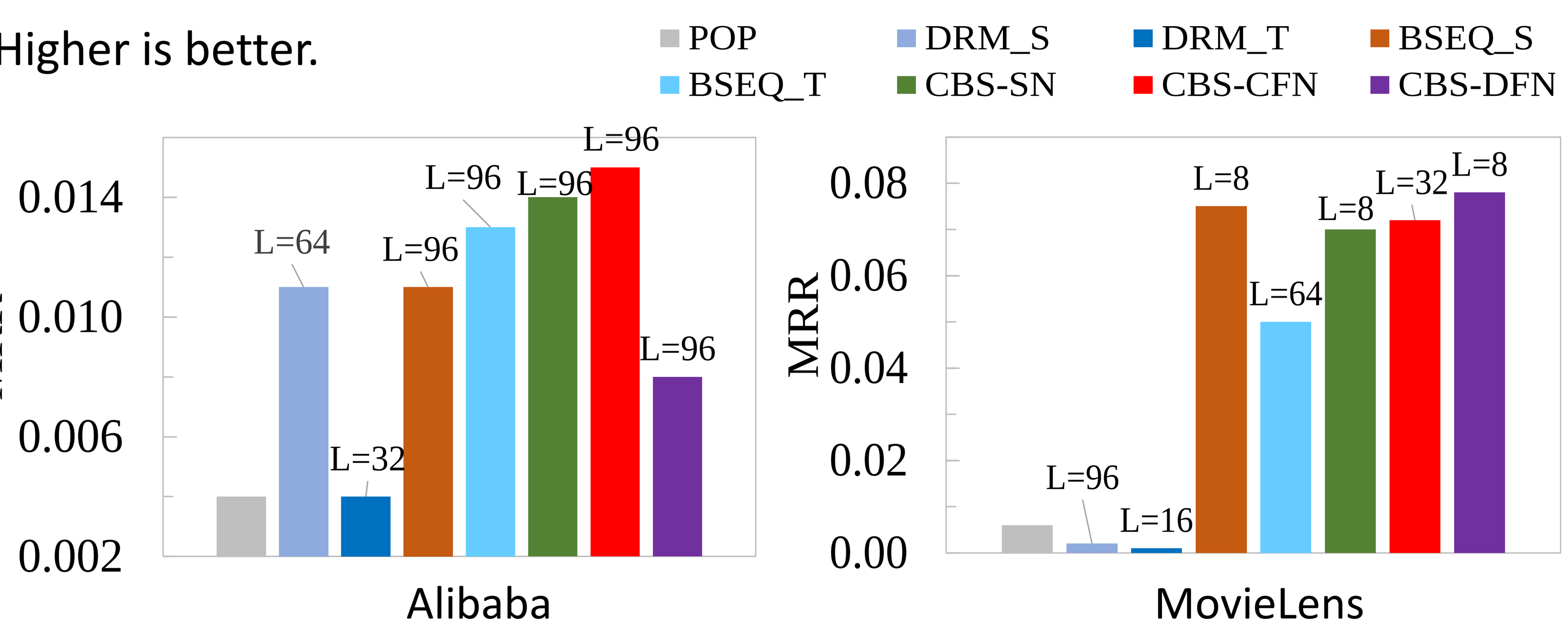
Datasets: **Alibaba** – "*click*" as support, "*purchase*" as target; and **MovieLens** - "*select a movie to rate*" as support, "*highly rate a movie*" as target

Methodology: For a given testing pair $< S, T >$, hide last target basket B and generate the **top-K predictions** given $< S, T \backslash B >$

Metric: Mean reciprocal rank (MRR) measures the overall ranking performance. Higher is better.



a) Single basket sequence vs. CBS (H=32)



b) Best Performance Comparison between the CBS models and baselines
(_T ~ Target only; _S ~ Support only; H =32)

Conclusion: Experiments on the two datasets show that the modeling of **Contemporaneous Basket Sequences with Twin networks** contributes **statistically significant** improvements as compared to **single basket-sequence models** in terms of top-K recommendations.