School of
Information Systems

SMU
SINGAPORE MANAGEMENT UNIVERSITY

Correlation-Sensitive Next-Basket Recommendation

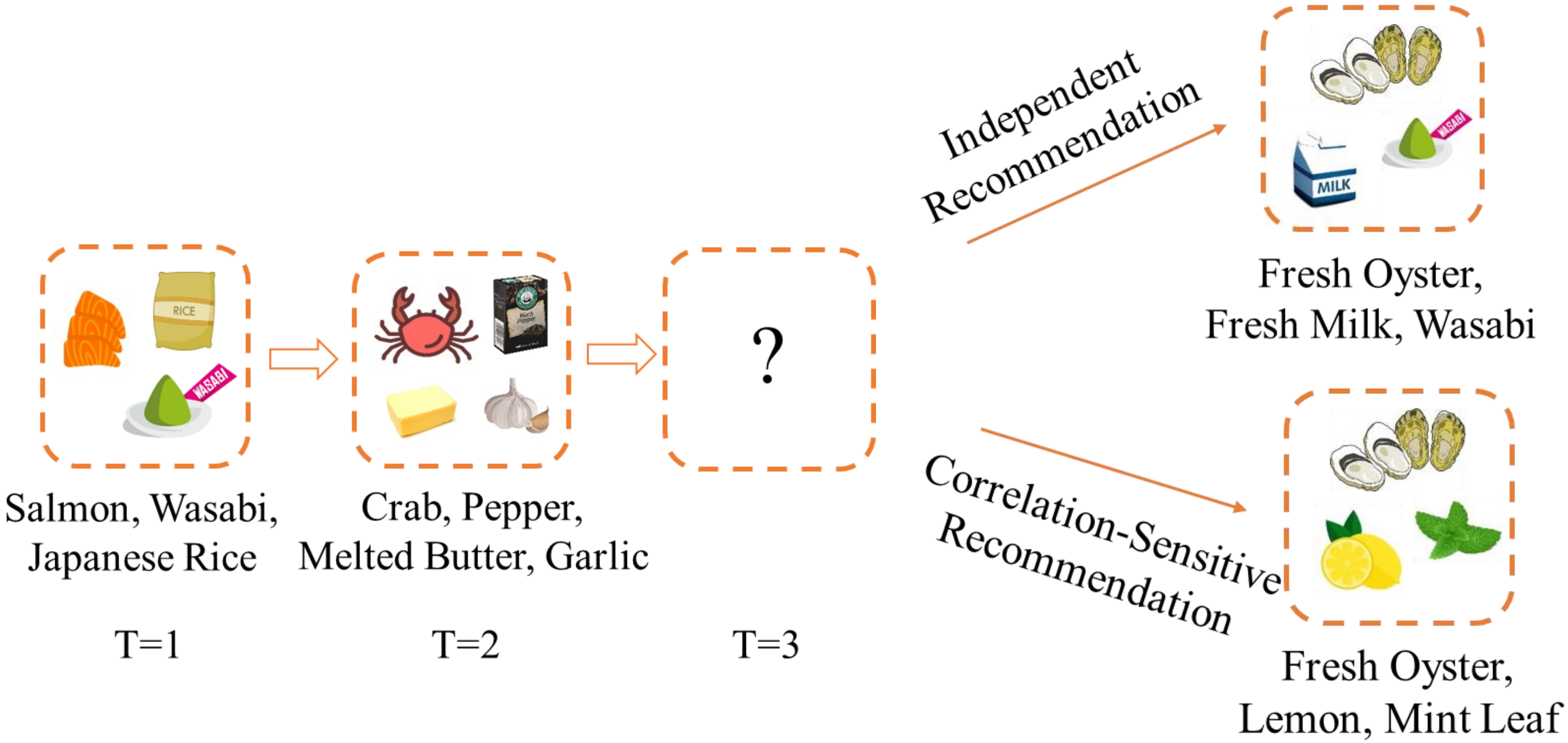
Duc-Trong Le, Hady W. Lauw, Yuan Fang

IJCAI 2019
Macau, China

Outline

- **Correlation-Sensitive Recommendation**
- **Basket-Sequence Correlation Networks (Beacon)**
  - Item-Item Correlation Matrix
  - Correlation-Sensitive Components
- **Experiments** on TaFeng, Delicious, Foursquare

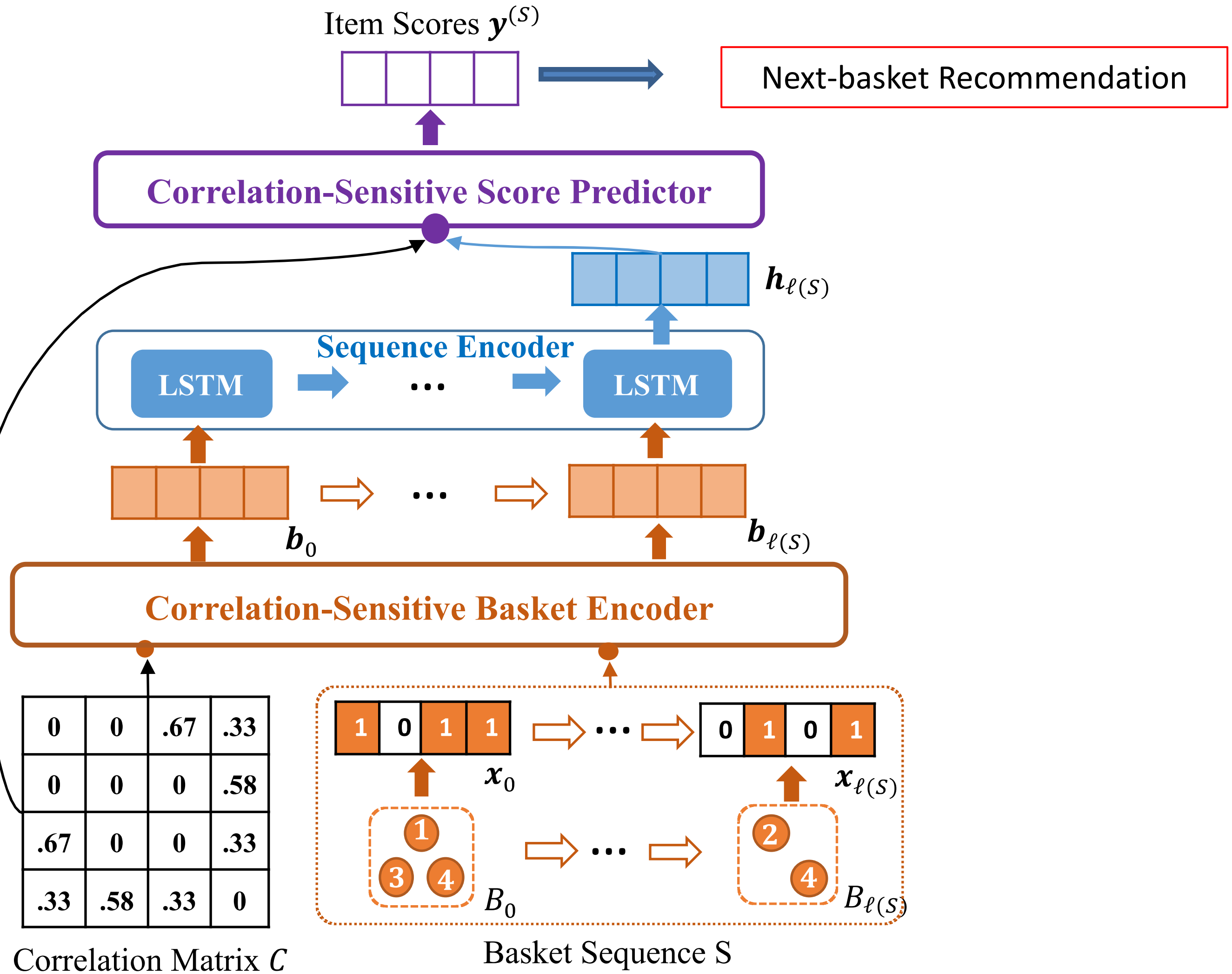
Correlation-Sensitive Recommendation



TASK:

Modeling concurrently **correlative** & **sequential** associations in basket sequences to predict **the next basket of correlated items.**

Correlation-Sensitive Next-Basket Recommendation with Basket-Sequence Correlation Networks (*Beacon*)

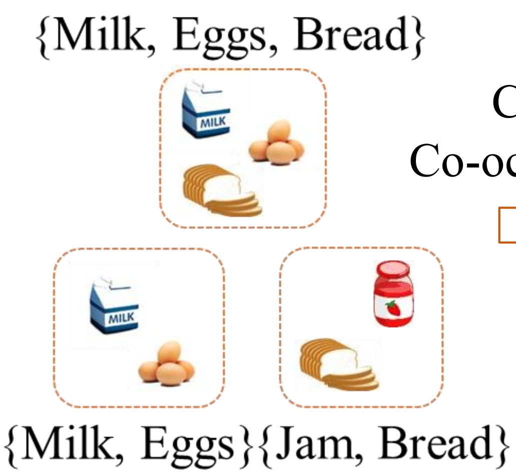


School of Information Systems

SMU SINGAPORE MANAGEMENT UNIVERSITY

Item-Item Correlation Matrix C

Objective: Leverage correlations between item-item pairs



Count Co-occurrence →

	0	0	2	1
	0	0	0	1
	2	0	0	1
	1	1	1	0

Co-occurrence Matrix F

Normalize →

	0	0	.67	.33
	0	0	0	.58
	.67	0	0	.33
	.33	.58	.33	0

Correlation Matrix C

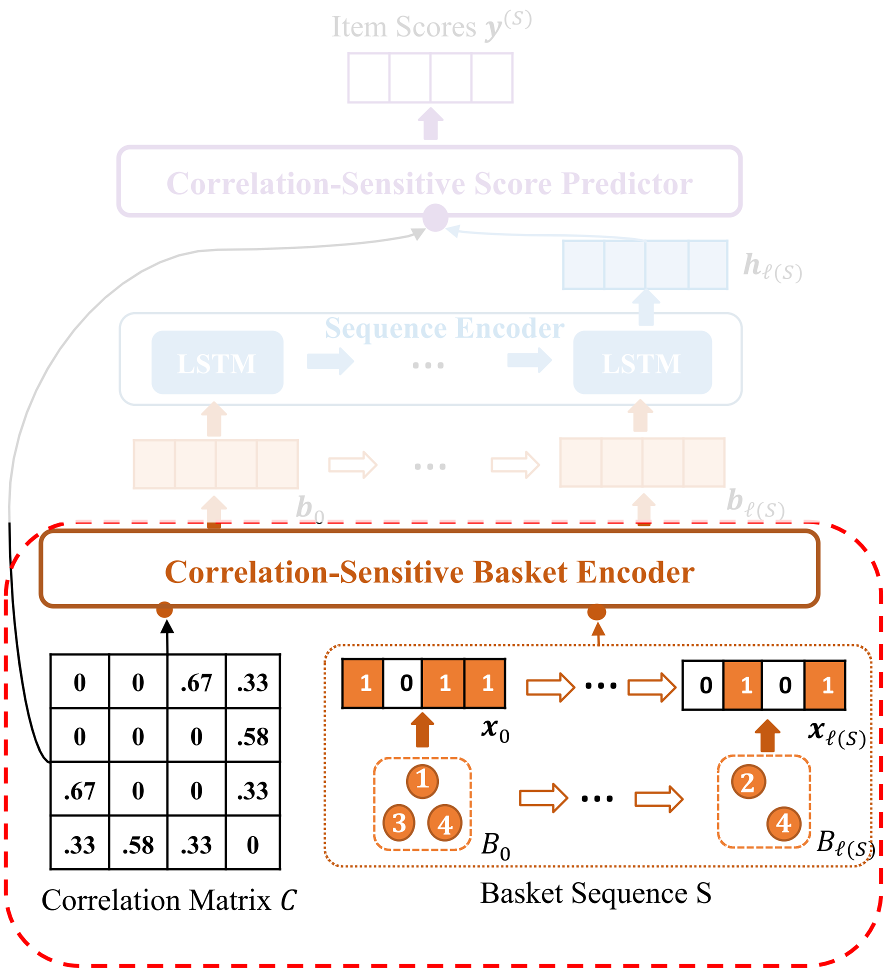
- A pair with frequent co-occurrence has a higher score than less frequent pairs.
- A pair with exclusive connection have a higher score than non-exclusive pairs.

$$C = D^{-\frac{1}{2}} F D^{-\frac{1}{2}},$$
$$D_{ii} = \sum_j C_{ij}$$

$$C \rightarrow C + \sum_{n=2}^{N} \mu^{n-1} \mathrm{Norm}(C^n),$$

N-th order matrix $\quad \mu \in (0, 1)$

Correlation-Sensitive Basket Encoder



- Input: Given a basket B_t

$$\mathbf{x}_t \in \{0,1\}^{|V|} \quad C \in \mathbb{R}^{|V|\times|V|}$$

- The **immediate representation** of B_t:

$$\boxed{\mathbf{z}_t = \underbrace{\mathbf{x}_t \circ \boldsymbol{\omega}}_{\text{Item Importance Parameters}} + \mathrm{ReLU}(\mathbf{x}_t C - \underbrace{\eta \mathbf{1}}_{\text{Noise-Canceling Parameter}}),}$$

$$\mathbf{z}_t, \boldsymbol{\omega} \in \mathbb{R}^{|V|} \qquad \eta \in \mathbb{R}^{+}$$

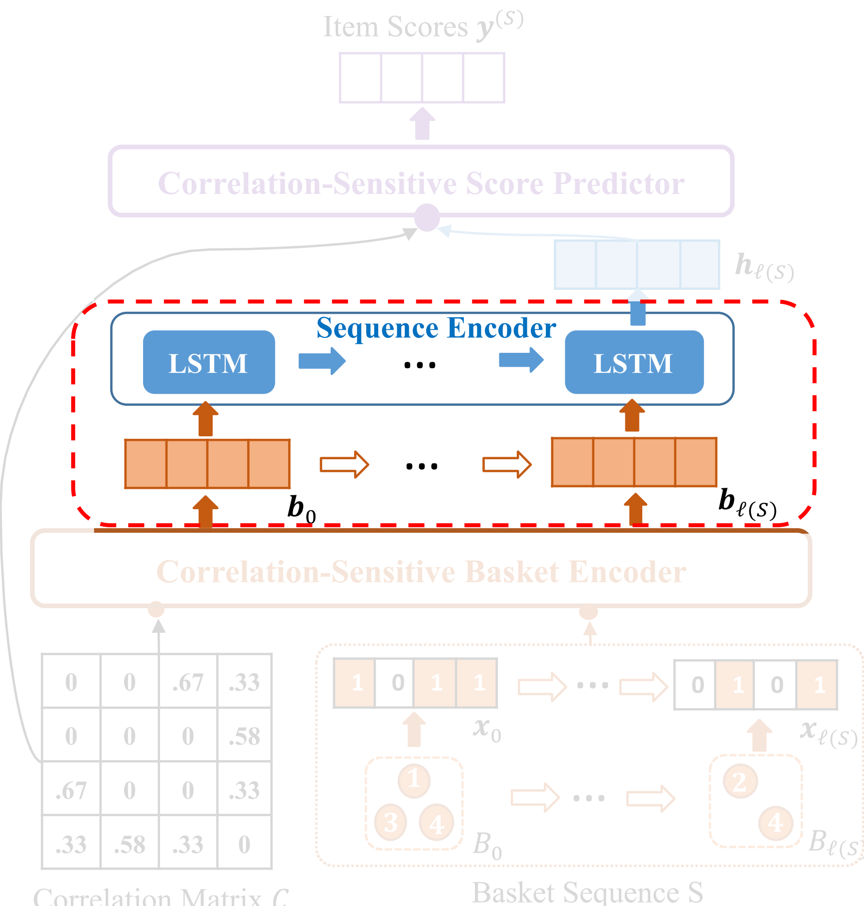
- The **L-dimensional latent representation**:

$$\boxed{\mathbf{b}_t = \mathrm{ReLU}(\mathbf{z}_t \Phi + \phi),}$$

$$\Phi \in \mathbb{R}^{|V|\times L} \quad \phi \in \mathbb{R}^{L}$$

$$\mathbf{b}_t \in \mathbb{R}^{L}$$

Sequence Encoder



- Input: Given the L-dimensional latent representation of basket B_t at time t

$$\mathbf{b}_t \in \mathbb{R}^L$$

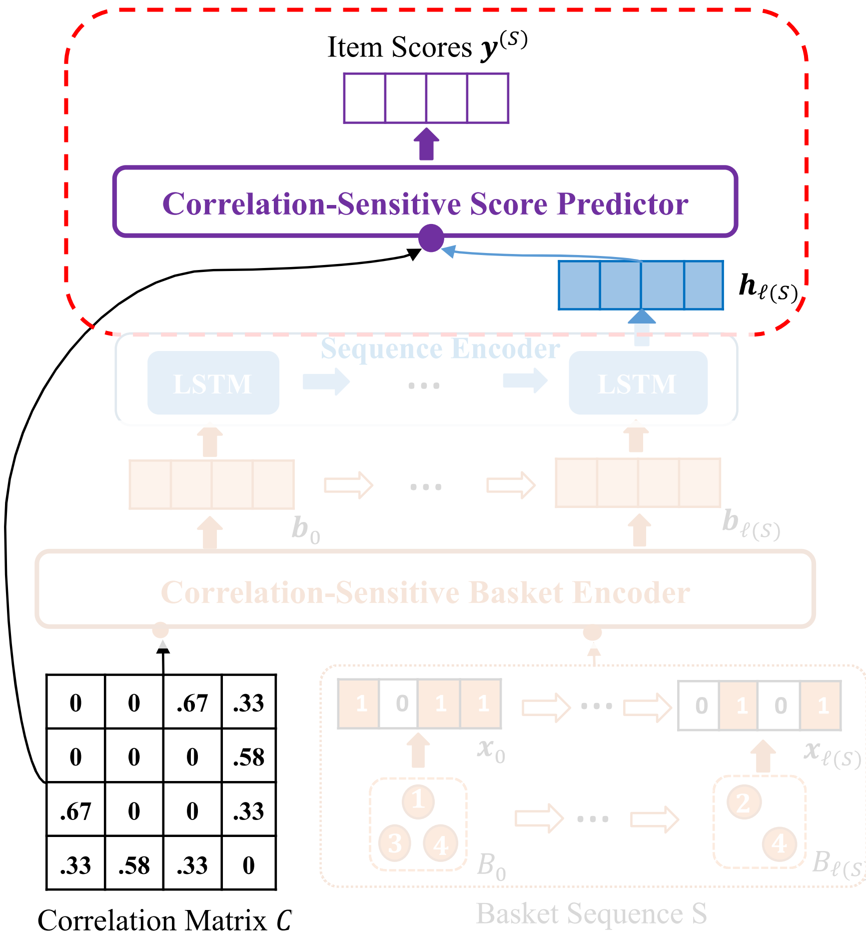
- The **recurrent hidden output** h_t at time t:

$$\boxed{\mathbf{h}_t = \tanh(\mathbf{b}_t\Psi + \mathbf{h}_{t-1}\Psi' + \psi)}$$

$$\Psi \in \mathbb{R}^{L\times H}, \Psi' \in \mathbb{R}^{H\times H}$$

$$\psi \in \mathbb{R}^H \quad \mathbf{h}_t \in \mathbb{R}^H$$

Correlation-Sensitive Score Predictor



- Input: Given the hidden output of the last basket B_T:

$$\mathbf{h}_t \in \mathbb{R}^H$$

- The **sequential signal** for next-item adoptions:

$$\boxed{\mathbf{s}^{(S)} = \sigma(\mathbf{h}_{\ell(S)}\Gamma),}$$

$$\mathbf{s}^{(S)} \in \mathbb{R}^{|V|} \quad \Gamma \in \mathbb{R}^{H \times |V|}$$

- The **correlation-sensitive score** predictor:

$$\boxed{\mathbf{y}^{(S)} = \underbrace{\alpha(\mathbf{s}^{(S)} \circ \boldsymbol{\omega} + \mathbf{s}^{(S)}C)}_{\text{Basket-Sensitive}} + \underbrace{(1-\alpha)\mathbf{s}^{(S)}}_{\text{Sequence-Sensitive}}}$$

$$\alpha \in [0,1] \quad \mathbf{y}^{(S)} \in \mathbb{R}^{|V|}$$

Next-Basket Recommendation Strategy

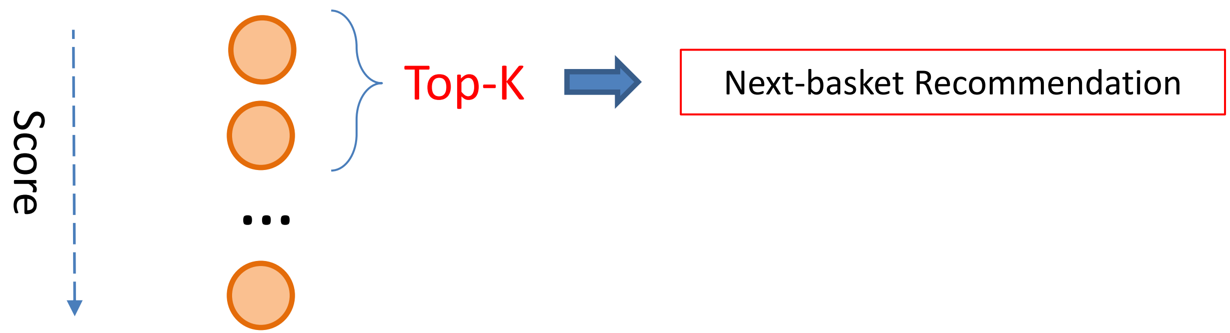
- The predicted ranking $\mathrm{r}^{(S)}$ of the item set based on $\mathrm{y}^{(S)}$

$$\mathrm{r}^{(S)}\colon \mathrm{y}^{(S)} \rightarrow \{1, 2, \dots, N\}$$

  where $\mathrm{r}_i^{(S)}$ is the ranking of item i.

- Approximately recommended basket B_{T+1} of size K:

$$B_{T+1}|K = \{i|\mathrm{r}_i^{(S)} \leq K\}$$



Ranked Item List

Experimental Setup

- **Task**: Next-basket recommendation
  - For each testing sequence S, hide the last target basket B.
  - Given $S' = S \backslash \mathrm{B}$, require each model to the next-basket recommendation with the ground-truth basket.
- **Datasets**

Dataset	#Sequence	#Item	#Average Length	#Average Basket Size
TaFeng (E-commerce)	77209	9964	7.0	5.9
Delicious (Bookmark Tag)	61908	6520	21.4	3.8
Foursquare (Check-ins)	100980	5527	22.2	1.8

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

- **Metric**: *F1@K, Half-life utility (HLU)*

RQ1: Does *Beacon* outperform against baselines?

Dataset	Model	L	H	F1@K (%) @5	F1@K (%) @10	HLU
TaFeng	*POP*	-	-	4.66	4.02	6.64
	MC	-	-	4.11	3.61	5.78
	MCN	8	-	4.56	4.02	6.34
	DREAM	8	-	5.85	4.90	6.96
	BSEQ	32	16	4.48	4.04	6.34
	triple2vec	64	-	4.66	3.88	4.85
	Beacon	8	64	**6.36**†	**5.26**†	**7.83**†
Delicious	*POP*	-	-	3.88	4.04	6.05
	MC	-	-	4.27	4.59	6.52
	MCN	32	-	4.20	4.59	6.50
	DREAM	32	-	3.13	3.47	4.93
	BSEQ	64	32	3.86	3.97	5.95
	triple2vec	32	-	3.76	4.04	5.16
	Beacon	64	64	**4.93**†	**5.47**†	**7.76**†
Foursquare	*POP*	-	-	2.73	2.90	4.84
	MC	-	-	3.58	3.43	5.53
	MCN	64	-	3.09	2.89	5.08
	DREAM	64	-	2.84	3.00	4.98
	BSEQ	64	32	2.80	2.89	4.82
	triple2vec	64	-	2.73	2.90	4.53
	Beacon	64	64	**3.61**	**3.59**†	**6.32**†

†, ‡, § denote statistically significant improvements ($p<0.05$) of *Beacon* over MCN, DRM, BSEQ

RQ2: Is the learning of item importance $\boldsymbol{\omega}$ useful?

Dataset	Model	F1@K (%) @5	F1@K (%) @10	HLU
TaFeng	*Beacon*$_{\text{corr-impt-}}$	3.87	3.44	5.13
	Beacon$_{\text{corr-}}$	5.78†	4.86†	7.18†
	Beacon (full)	**6.36**§	**5.26**§	**7.83**§
Delicious	*Beacon*$_{\text{corr-impt-}}$	4.02	4.43	6.38
	Beacon$_{\text{corr-}}$	4.67†	5.10†	7.15†
	Beacon (full)	**4.94**§	**5.47**§	**7.76**§
Foursquare	*Beacon*$_{\text{corr-impt-}}$	2.98	3.29	5.39
	Beacon$_{\text{corr-}}$	3.58†	3.52†	6.16†
	Beacon (full)	**3.61**	**3.59**§	**6.32**§

$\dagger$: denotes statistically significant improvements ($p<0.05$)

Beacon$_{\text{corr-impt-}}$ Without correlations, fix $\boldsymbol{\omega} = \mathbf{1}$
Beacon$_{\text{corr-}}$ Without correlations, learn $\boldsymbol{\omega}$
Beacon (full) With correlations, learn $\boldsymbol{\omega}$

RQ3: Is the modeling of correlations useful?

Dataset	Model	F1@K (%) @5	F1@K (%) @10	HLU
TaFeng	*Beacon*$_{\text{corr-impt-}}$	3.87	3.44	5.13
	Beacon$_{\text{corr-}}$	5.78†	4.86†	7.18†
	Beacon (full)	**6.36**§	**5.26**§	**7.83**§
Delicious	*Beacon*$_{\text{corr-impt-}}$	4.02	4.43	6.38
	Beacon$_{\text{corr-}}$	4.67†	5.10†	7.15†
	Beacon (full)	**4.94**§	**5.47**§	**7.76**§
Foursquare	*Beacon*$_{\text{corr-impt-}}$	2.98	3.29	5.39
	Beacon$_{\text{corr-}}$	3.58†	3.52†	6.16†
	Beacon (full)	**3.61**	**3.59**§	**6.32**§

§ denotes statistically significant improvements ($p<0.05$)

Beacon$_{\text{corr-impt-}}$ Without correlations, fix $\boldsymbol{\omega} = \mathbf{1}$
Beacon$_{\text{corr-}}$ Without correlations, learn $\boldsymbol{\omega}$
Beacon (full) With correlations, learn $\boldsymbol{\omega}$

A Qualitative Example on Delicious

Target bookmark	Tag basket prediction ($K = 5$)		
	Beacon	*MC*	*POP*
Manual de jQuery [6]	*web*, design, *programming*, *javascript*, *tools*	digital, sociales, *web*, internet, periodismo	art, design, education, video, *tools*
The $300 Million Button [7]	twitter, *ux*, *propinquity*, *critical*, *writing*	*design*, peace, education, blog, *tips*	art, design, education, video, *tools*

[6] http://www.desarrolloweb.com/manuales/manual-jquery.html
[7] https://articles.uie.com/three_hund_million_button

Conclusion

- Modeling concurrently **correlative** **&** **sequential** associations in basket sequences to predict **next-basket of correlated items.**
- ***Propose*** ***Beacon*** show **statistically significant** improvements over:
    - **Traditional Basket Sequence Models (MCN, DRM, BSEQ, triple2vec)**

  in terms of top-K recommendations.

THANK YOU
Q&A

RQ4: How does α affect the performance?