THE WEB CONFERENCE ACM

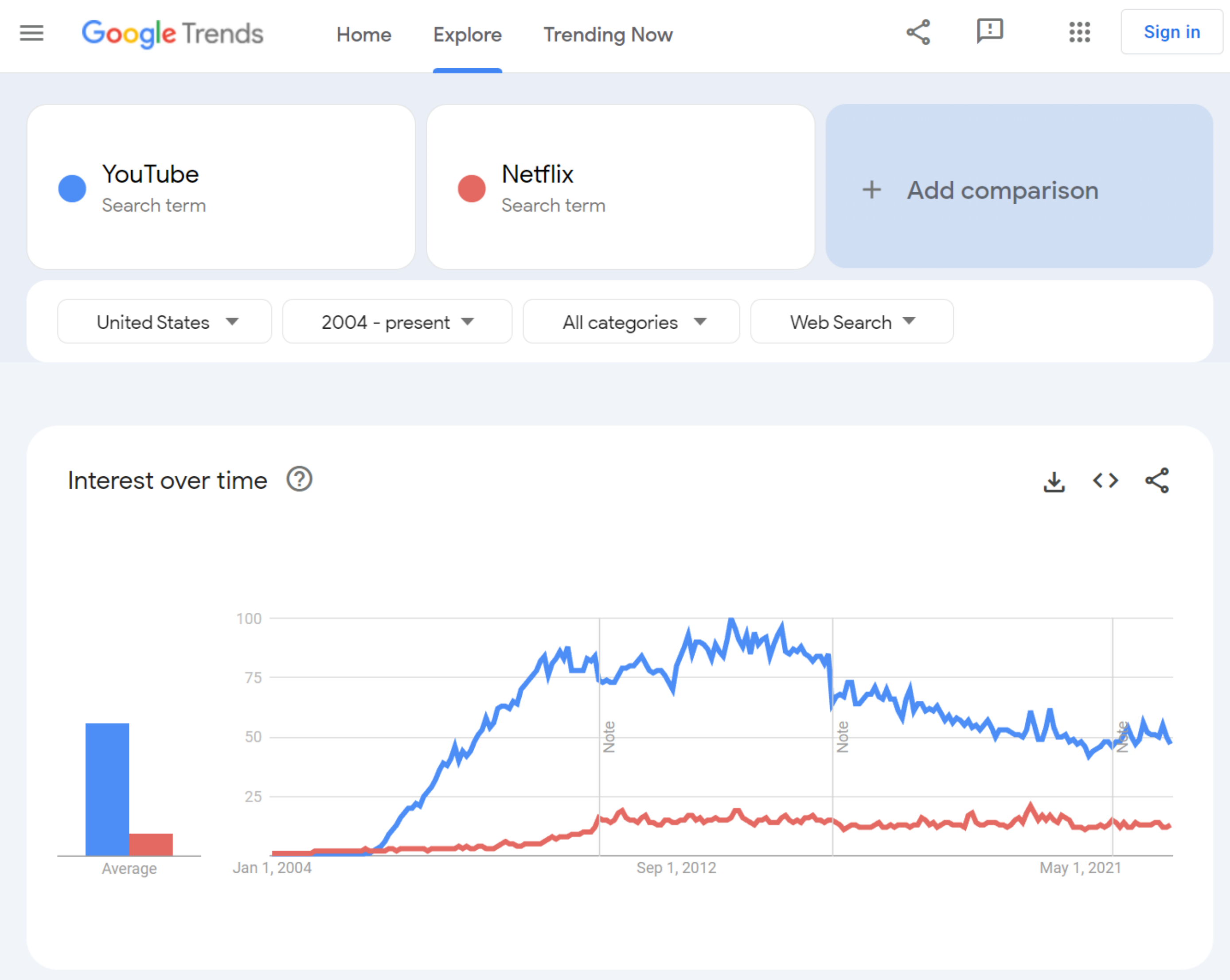
Modeling Dynamic Interactions over Tensor Streams

K. Kawabata Y. Matsubara Y. Sakurai

SANKEN Osaka University Japan

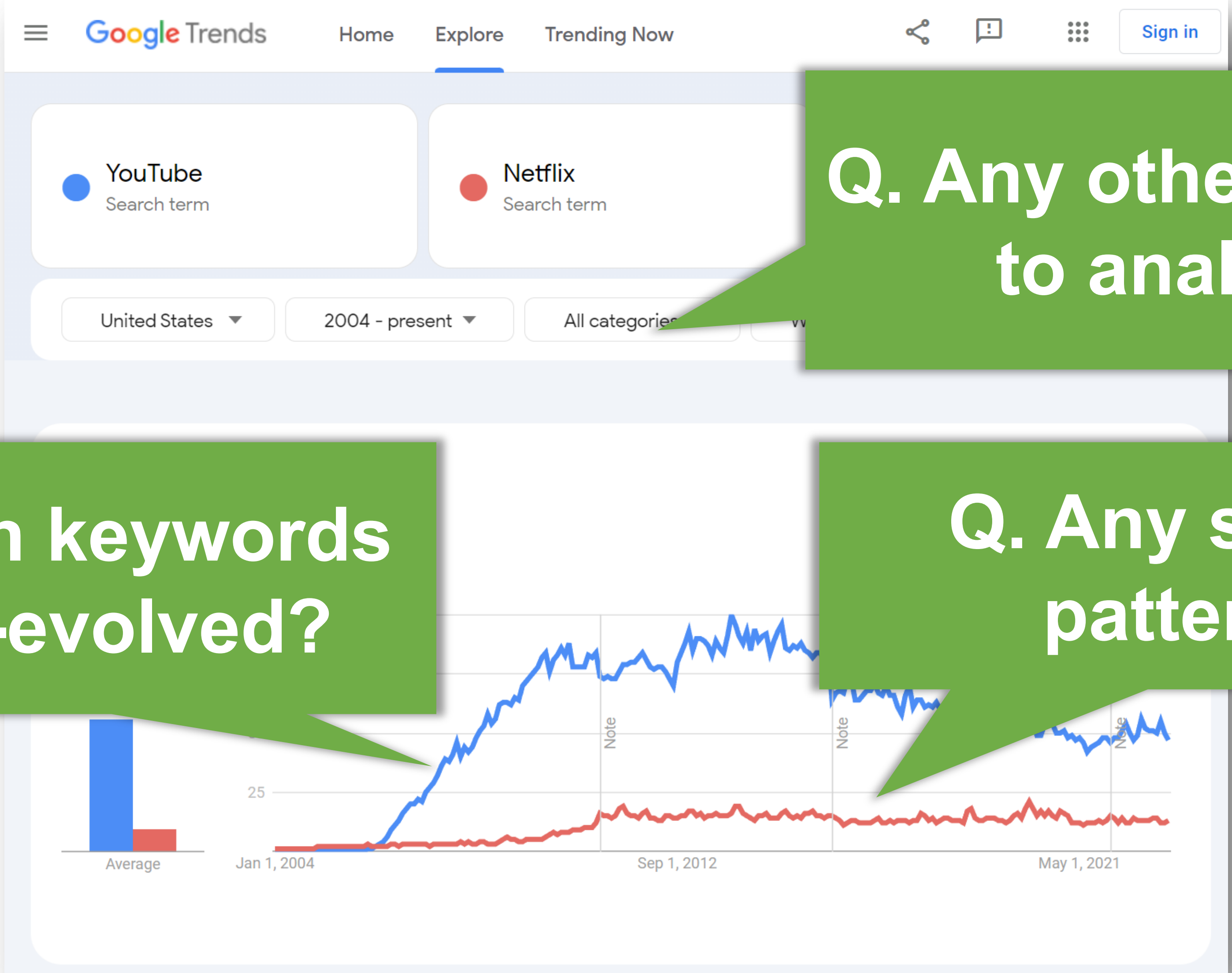
SANKEN OSAKA UNIVERSITY

Motivation



Google Trends
Home
Explore
Trending Now
Sign in
YouTube
Search term
Netflix
Search term
+ Add comparison
United States
2004 - present
All categories
Web Search
Interest over time
100
75
50
25
Note
Note
Note
Average
Jan 1, 2004
Sep 1, 2012
May 1, 2021

Motivation

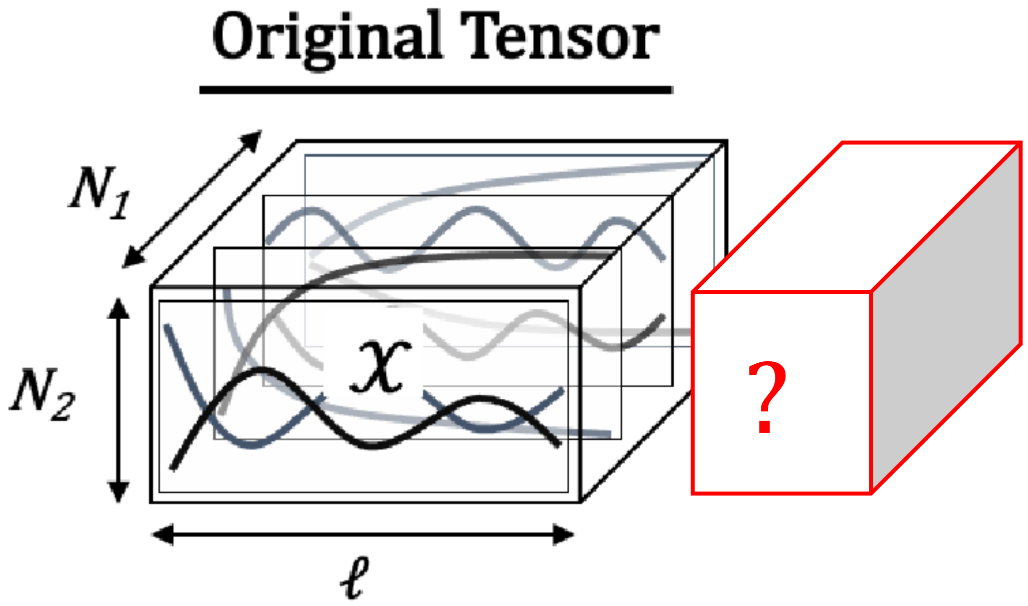


Q. Any other aspects to analyze?

Q. Which keywords are co-evolved?

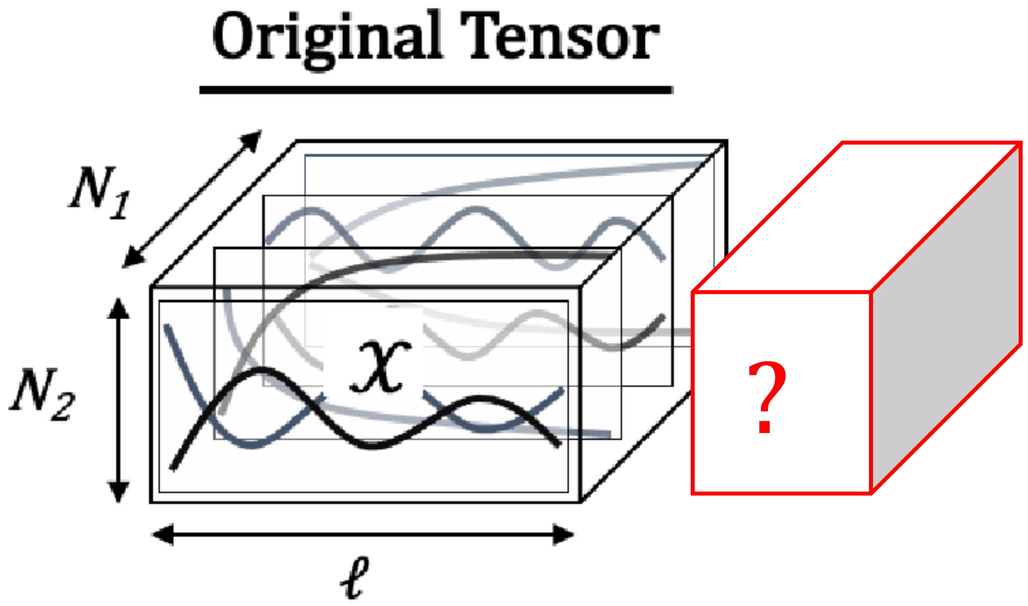
Q. Any shifting patterns?

Problem definition



- **Given:** the most recent ℓ-long tensor
- **Forecast:** ℓ_s-steps ahead tensor continuously

Problem definition

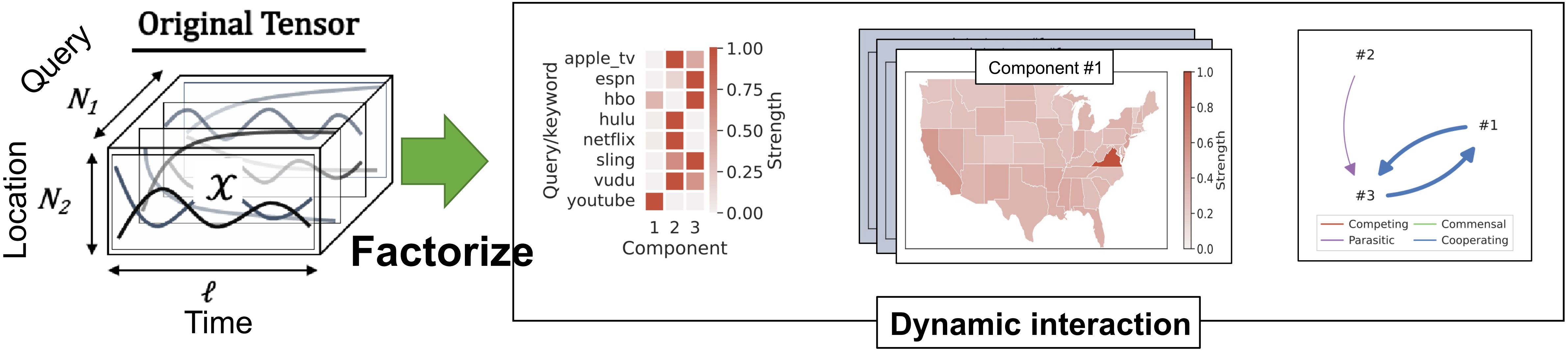


- **Given:** the most recent ℓ-long tensor
- **Forecast:** ℓ_s-steps ahead tensor continuously

DISMO: a new streaming tensor factorization method

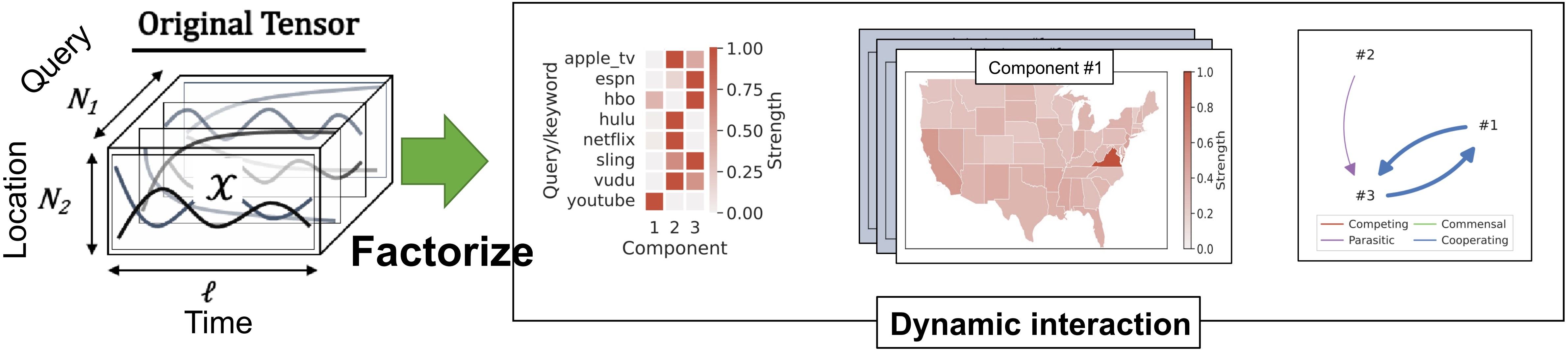
Preview of the results

- Trend analysis on Video-on-demand services



Preview of the results

- Trend analysis on Video-on-demand services

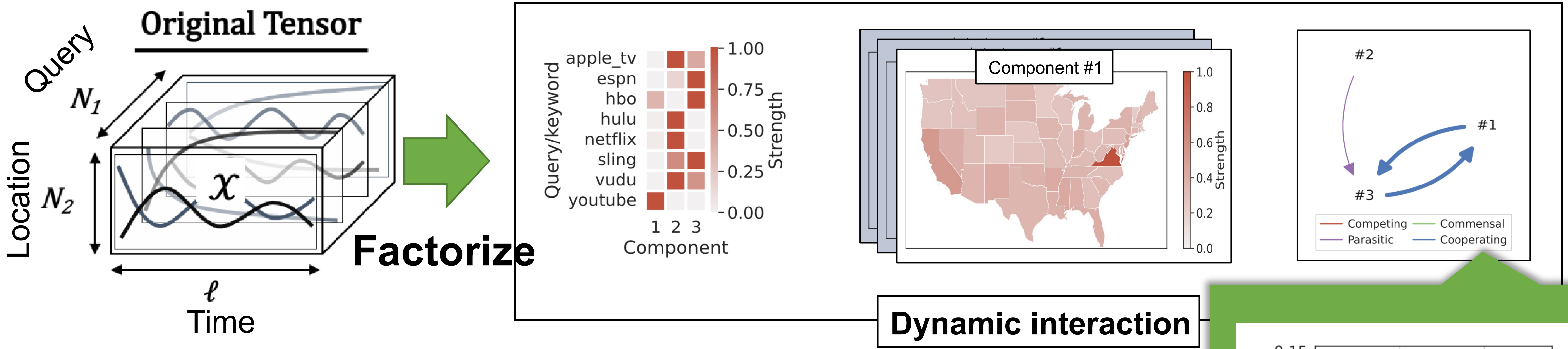


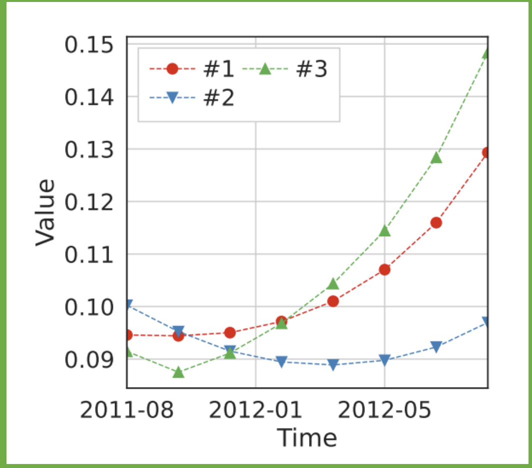
Dynamic interaction:

- Multi-aspect factors that represents latent groups of attributes
- Interaction relationships (network) between latent groups
- All-aspect factors can evolve for recent data

Preview of the results

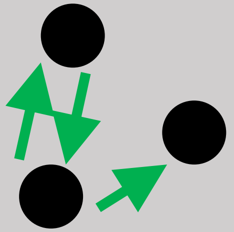
- Trend analysis on Video-on-demand services





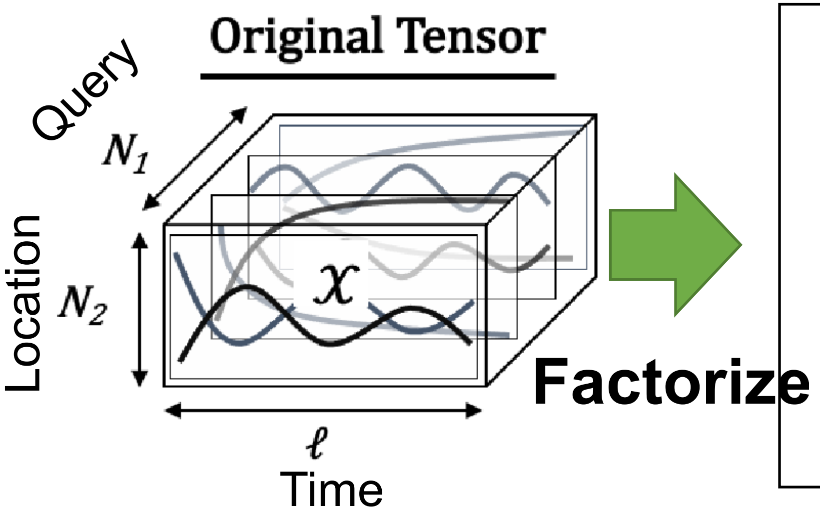
Time components in DISMO:

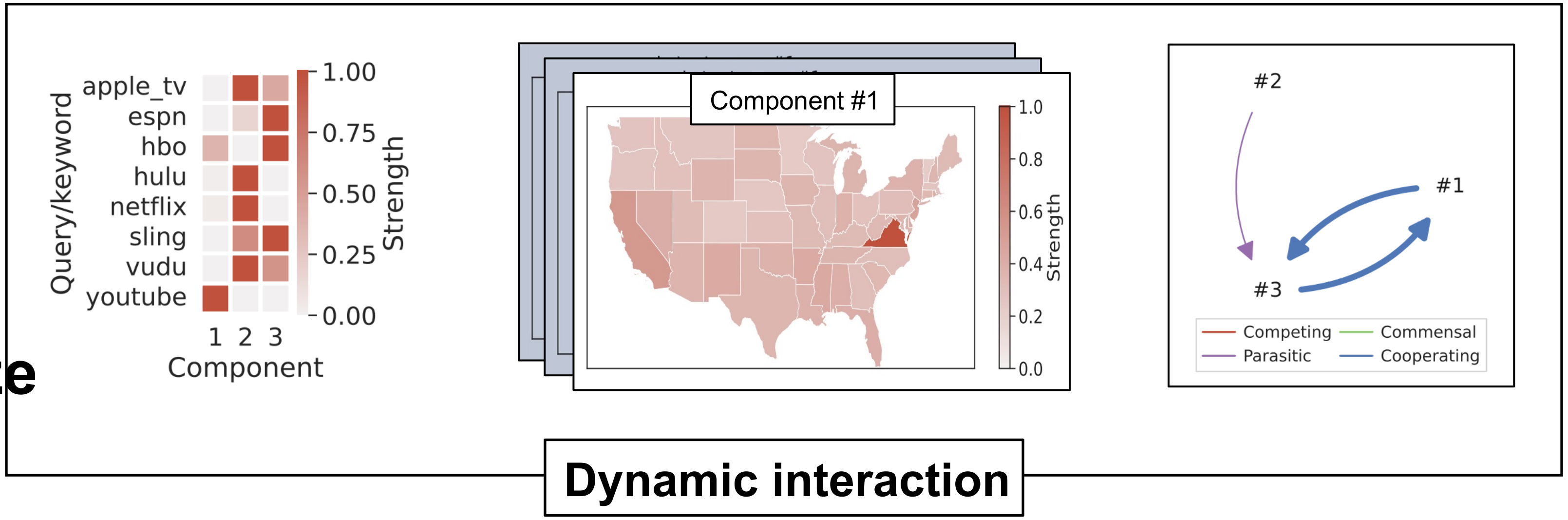
- Interpret 4 types of relationships
- Generate latent dynamics effectively



Preview of the results

- Trend analysis on Video-on-demand services



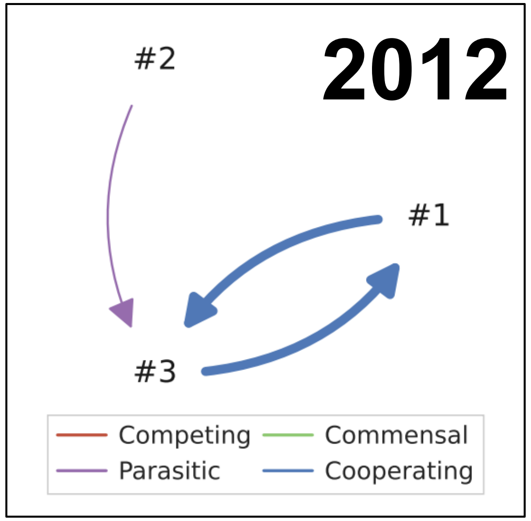


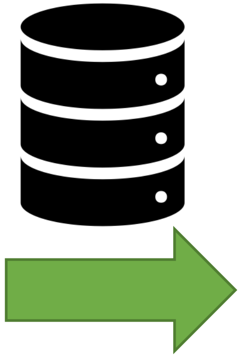
DISMO automatically finds meaningful latent groups

#1 Representative video sharing service	#2 On-demand subscription services	#3 TV-based subscription services

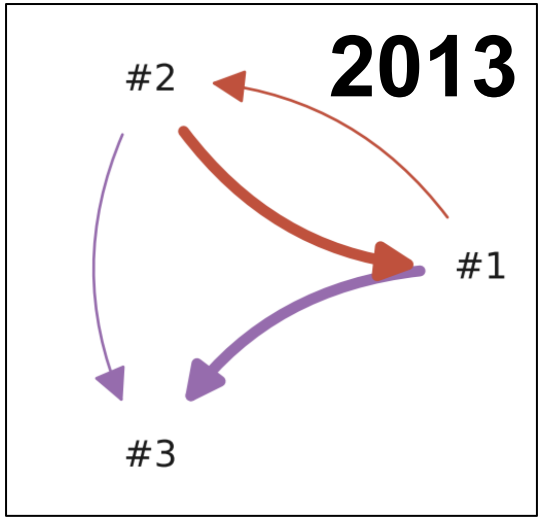
Preview of the results

- Trend analysis on Video-on-demand services





Shift!



DISMO automatically finds shifting patterns

#1 Representative video sharing service

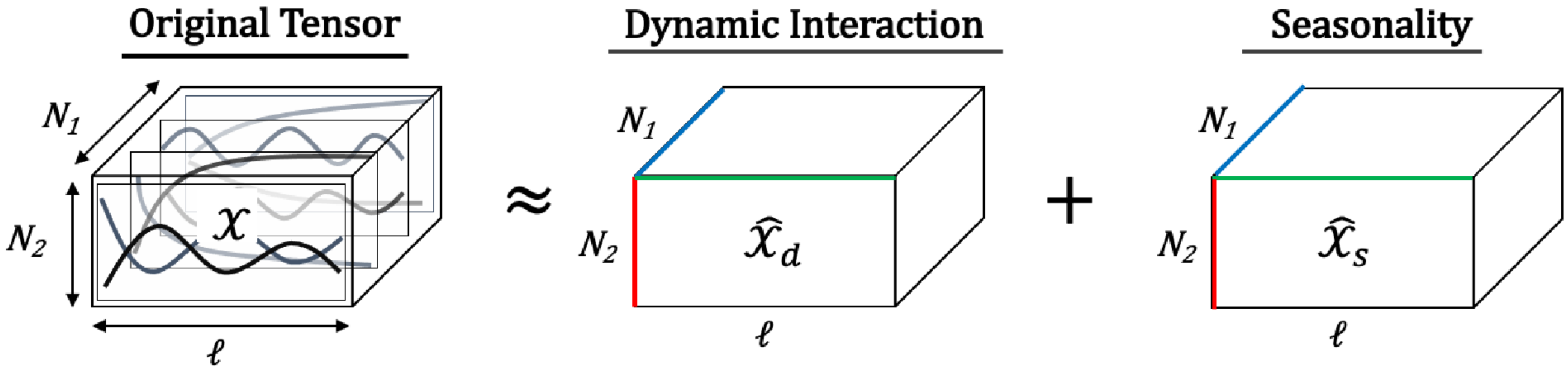
#2 On-demand subscription services

#3 TV-based subscription services

Outline

- **Motivation**
- **Problem definition**
- **Proposed model**
- **Proposed algorithms**
- **Experiments**
- **Conclusion**

Proposed model: overview



- **Key idea: dynamic interaction and seasonality**
  - Step 1. Non-linear dynamical system for interactions
  - Step 2. Interaction-based tensor factorization
  - Step 3. Multi-aspect factor set for shifting trends

1. Non-linear dynamical system

- **Background: Interactions between species determine population dynamics**
  - modeled by a non-linear dynamical system:

Intrinsic growth rate **Interaction strength**

$$\frac{dx_i}{dt} = \boxed{a_i} x_i \left(1 - \frac{\sum_{j=1}^{k} \boxed{c_{ij}} x_j}{\boxed{b_i}} \right)$$

Carrying capacity

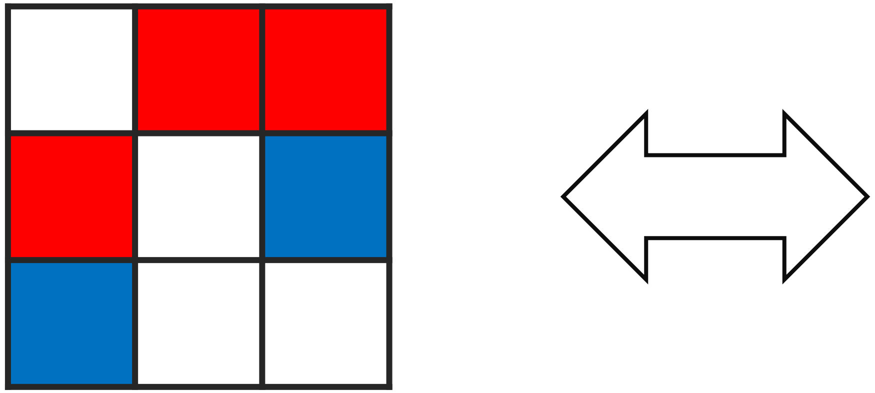
1. Non-linear dynamical system

- **Q. How to model interactions?**
- **Idea: Latent Interaction system (LIS):** $\theta = \{a, b, C\}$
  - Species = Users
  - Population dynamics = Trends (e.g., user attention)

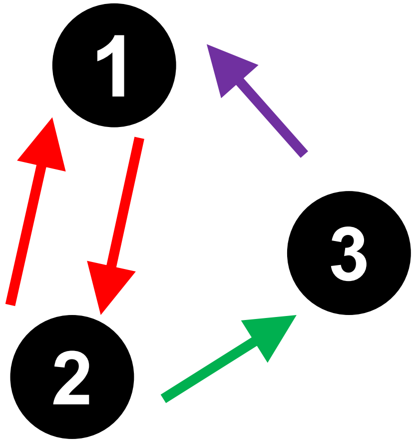
c_{ij}	c_{ji}	**Relationship**
+	+	Competing
−	−	Cooperating
−	+	Parasitic
−	0	Commensal

1. Non-linear dynamical system

- **Q. How to model interactions?**
- **Idea: Latent Interaction system (LIS):** $\theta = \{a, b, C\}$
  - Example: interaction coefficients as a graph



Matrix C



Nodes: latent groups
Edges: Relation types

2. Multi-aspect mining for latent interactions

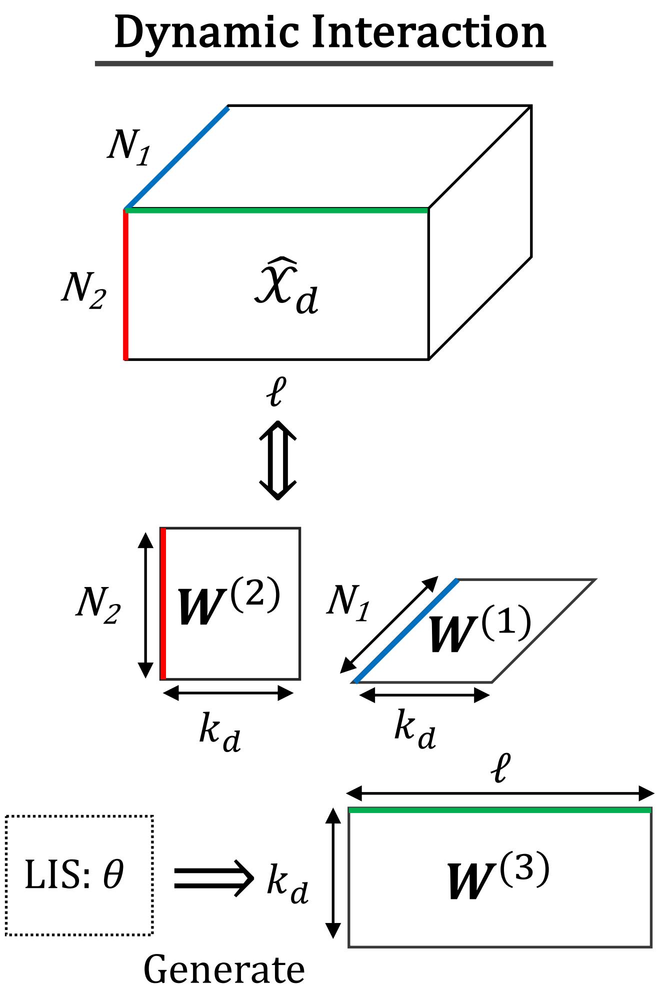
- **Q. How to find latent groups?**
- **Idea: Interaction factor set:**

$$\mathcal{D} = \{W^{(1)}, \dots, W^{(M)}, \theta\}$$

Non-negative elements

$$\hat{\mathcal{X}}_d = \sum_{i=1}^{k_d} \boldsymbol{w}_i^{(1)} \circ \cdots \circ \boldsymbol{w}_i^{(M)} \circ \boldsymbol{w}_i^{(M+1)}$$

Time components are generated by θ



2. Multi-aspect mining for latent interactions

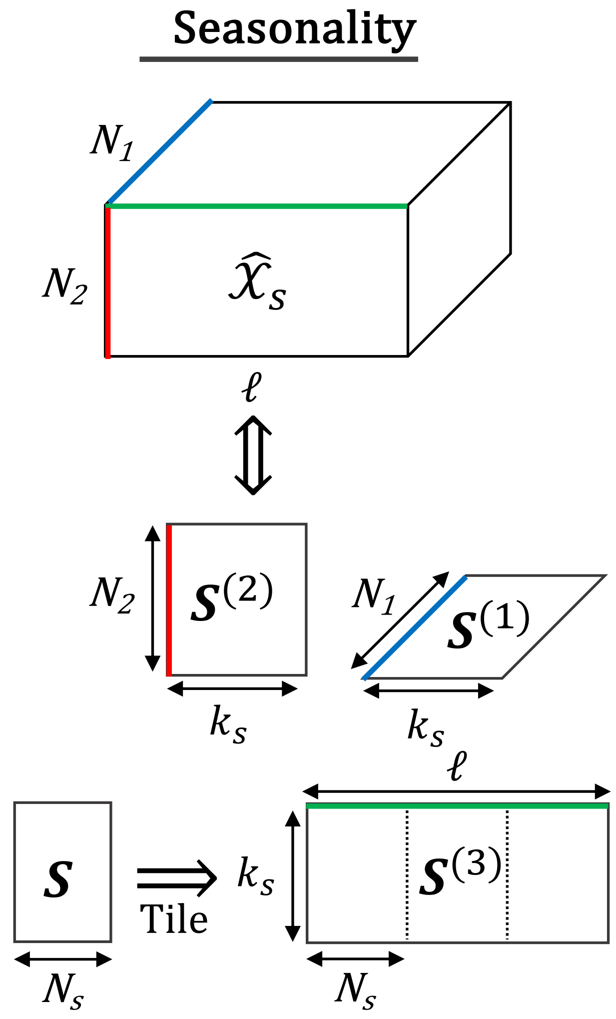
- **Q. How to extract seasonality?**
- **Idea: Seasonal factor set:**

$$\mathcal{S} = \{S^{(1)}, \ldots, S^{(M)}, S\}$$

k_s is independent of k_d

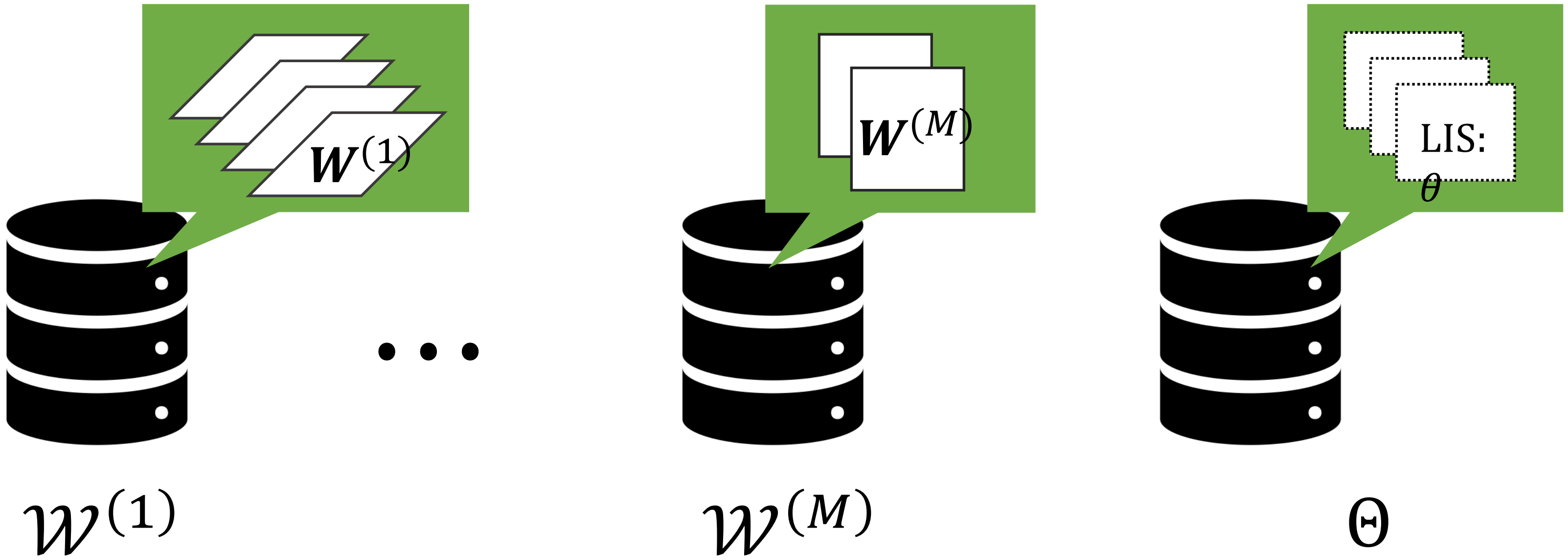
$$\hat{\mathcal{X}}_s = \sum_{i=1}^{k_s} \boldsymbol{s}_i^{(1)} \circ \cdots \circ \boldsymbol{s}_i^{(M)} \circ \boldsymbol{s}_i^{(M+1)}$$

S represents cyclic patterns in N_s periods



3. Dynamic interactions

- **Q. How to detect shifting trends?**
- **Idea: Sets of multi-aspect components:** $\mathcal{W}^{(1)}, \ldots, \mathcal{W}^{(M)}, \Theta$



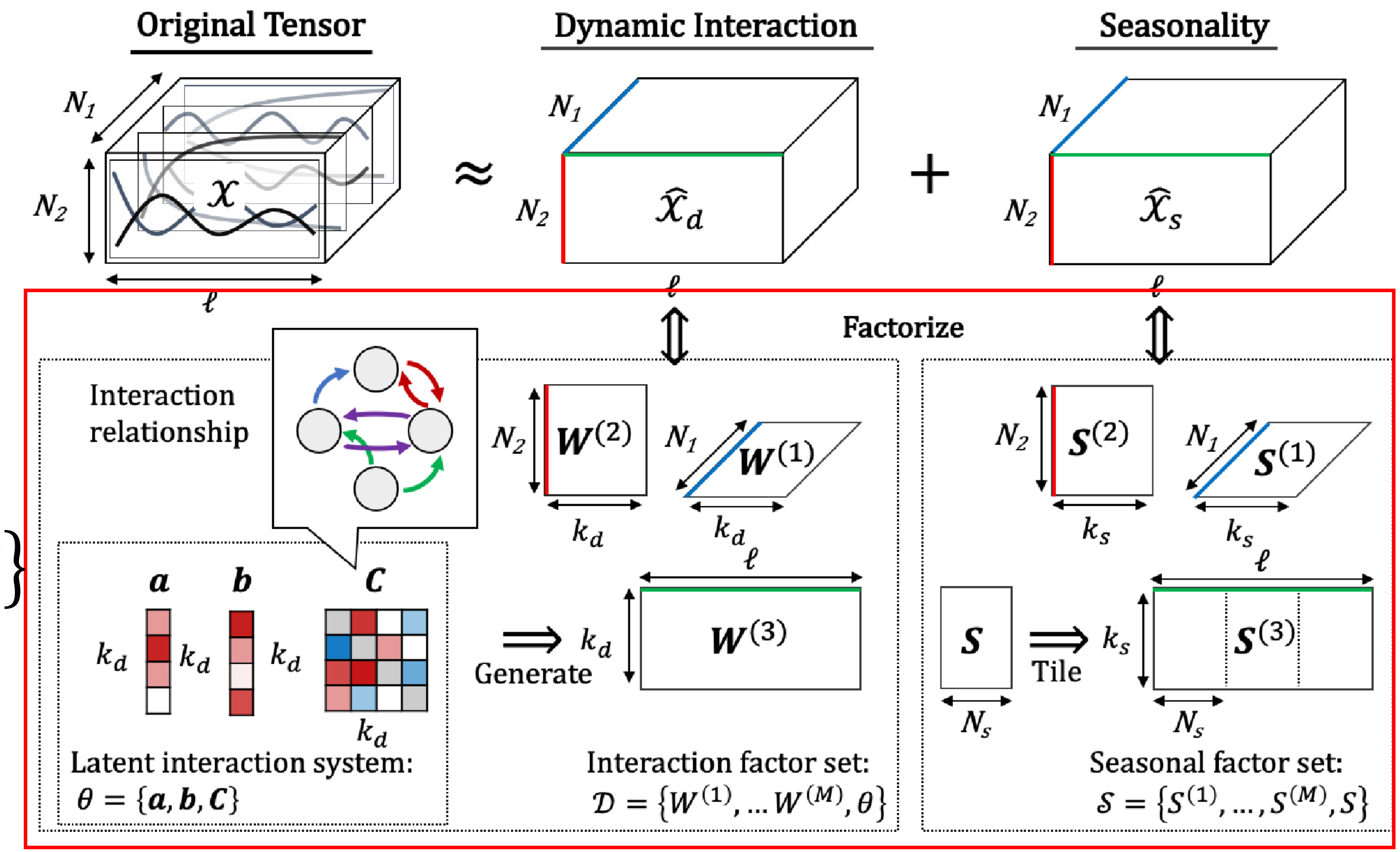
DISMO employs new factors for each aspect if required

Proposed model

- **Goal:** Estimate and update a full parameter set $\mathcal{F}$ for $\mathcal{X}$

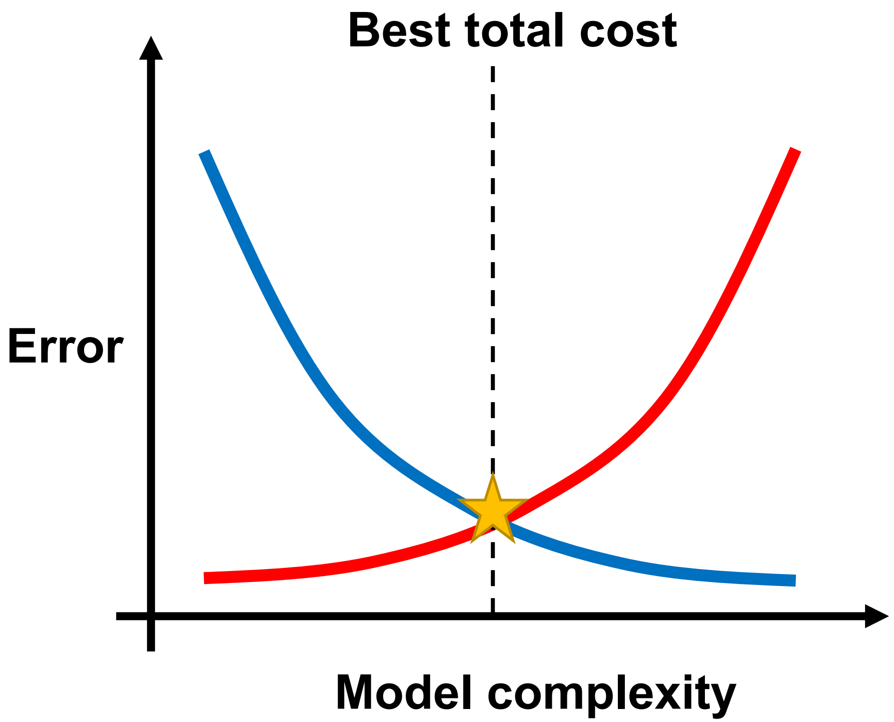
Full model set of DISMO

$$\mathcal{F} = \{\mathcal{W}^{(1)}, \dots, \mathcal{W}^{(M)}, \Theta, \mathcal{S}\}$$



Automatic tensor compression

- Q. How to estimate $\mathcal{F}$ automatically?
- Idea: Minimum description length (MDL) principle



$$\underbrace{<\mathcal{X};\mathcal{F}>}_{\text{Total cost}} = \underbrace{<\mathcal{F}>}_{\text{Model cost}} + \underbrace{<\mathcal{X}|\mathcal{F}>}_{\text{Encoding cost}}$$

DISMO determines # of components automatically

Automatic tensor compression

- Model cost: $< \mathcal{F} > = < \Theta > + < \mathcal{W}^{(1)} > + \cdots + < \mathcal{W}^{(M)} > + < \mathcal{S} >$

$$< \Theta > = \sum_{\theta \in \Theta} < \theta >, \ < \theta > = < \boldsymbol{a} > + < \mathbf{b} > + < \mathbf{C} > .$$
$$< \boldsymbol{a} > = |\boldsymbol{a}| \cdot (\log(k_d) + c_F) + \log^*(|\boldsymbol{a}|),$$
$$< \mathbf{b} > = |\mathbf{b}| \cdot (\log(k_d) + c_F) + \log^*(|\mathbf{b}|),$$
$$< \mathbf{C} > = |\mathbf{C}| \cdot (2 \cdot \log(k_d) + c_F) + \log^*(|\mathbf{b}|),$$

$$< \mathcal{W}^{(m)} > = \sum_{\mathbf{W}^{(m)} \in \mathcal{W}^{(m)}} < \mathbf{W}^{(m)} >, \ < \mathcal{S} > = \sum_{\mathbf{S}^{(m)} \in \mathcal{S}} < \mathbf{S}^{(m)} >,$$
$$< \mathbf{W}^{(m)} > = |\mathbf{W}^{(m)}| (k_d / N_m)(\log(N_m) + \log(k_d) + c_F) + \log^*(|\mathbf{W}^{(m)}|),$$
$$< \mathbf{S}^{(m)} > = |\mathbf{S}^{(m)}| (k_s / N_m)(\log(N_m + \log(k_s) + c_F) + \log^*(|\mathbf{S}^{(m)}|).$$

Normalized for skewed # of dimensions of each aspect

- Data encoding cost: Negative log-likelihood

$$< \mathcal{X} | \mathcal{F} > = \sum_{x \in \mathcal{X}} -\log_2 p_{\mu,\sigma}(x - \hat{x}_d - \hat{x}_s).$$

Outline

- Motivation
- Problem definition
- Proposed model
- **Proposed algorithms**
- **Experiments**
- **Conclusion**

DISMO factorization

- **Given:** a multi-order tensor $\mathcal{X} \in \mathbb{N}^{N_1 \times \cdots \times N_M \times \ell}, k_d, k_s$
- **Object:** $\min_{\mathcal{D},\mathcal{S}} \|\mathcal{X} - \hat{\mathcal{X}}_d - \hat{\mathcal{X}}_s\|$
- **Idea:** Alternating updates of multi-aspect factors

DISMO factorization

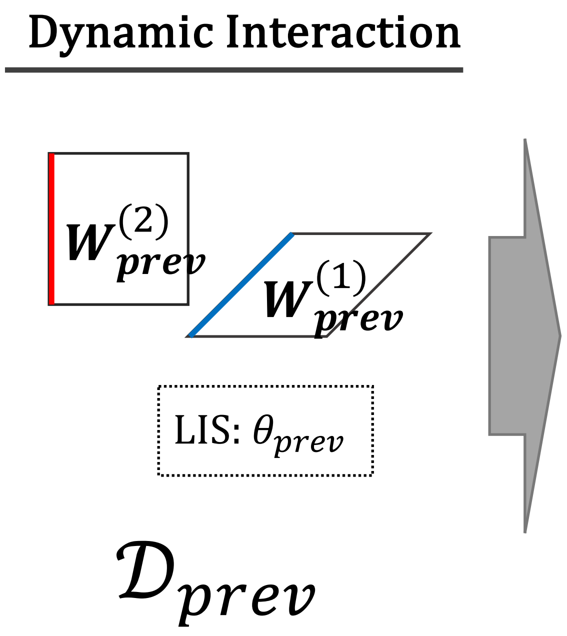
- **Given:** a multi-order tensor $\mathcal{X} \in \mathbb{N}^{N_1 \times \cdots \times N_M \times \ell}$
- **Object:** $\min_{\mathcal{D},\mathcal{S}} \left\| \mathcal{X} - \widehat{\mathcal{X}}_d - \widehat{\mathcal{X}}_s \right\|$
- **Idea:** Alternating updates of multi-aspect factors

Streaming DISMO factorization

- **Given:** a multi-order tensor $\mathcal{X} \in \mathbb{N}^{N_1 \times \cdots \times N_M \times \ell}$
- **Object:** Estimate $\mathcal{D}, \mathcal{S}$ that minimizes $\Delta < \mathcal{X}; \mathcal{D}, \mathcal{S} >$
- **Idea1:** Alternating shifts in dynamic interaction set $\mathcal{D}$

Streaming DISMO factorization

- **Given:** a multi-order tensor $\mathcal{X} \in \mathbb{N}^{N_1 \times \cdots \times N_M \times \ell}$
- **Object:** Estimate $\mathcal{D}, \mathcal{S}$ that minimizes $\Delta < \mathcal{X}; \mathcal{D}, \mathcal{S} >$
- **Idea1:** Alternating shifts in dynamic interaction set $\mathcal{D}$



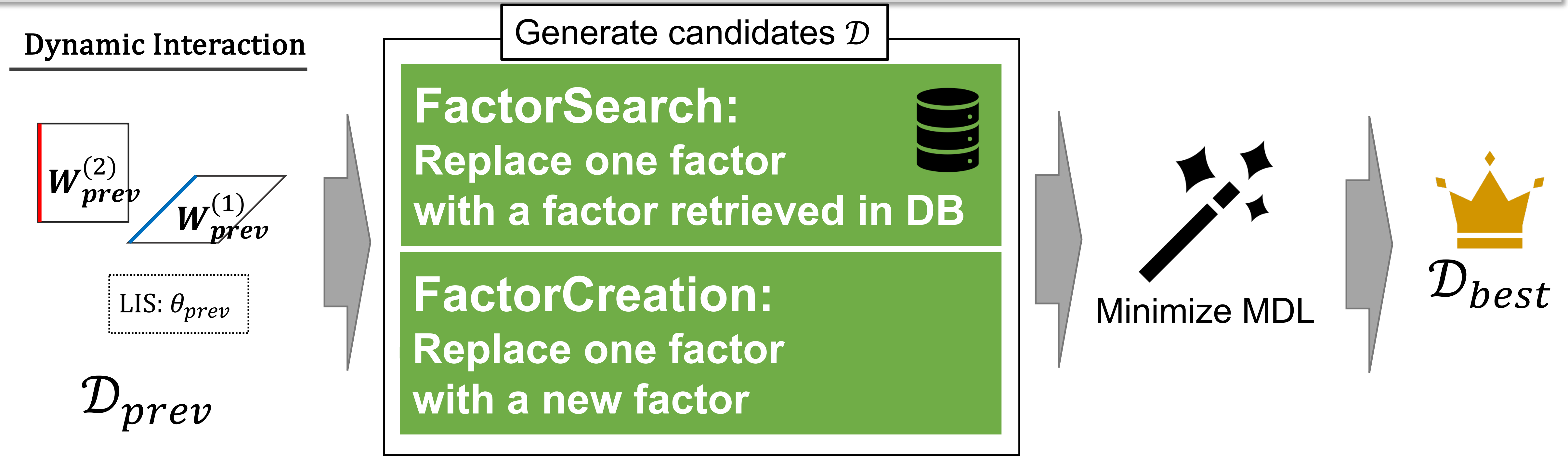
Generate candidates $\mathcal{D}$

FactorSearch:
Replace one factor with a factor retrieved in DB

FactorCreation:
Replace one factor with a new factor

Streaming DISMO factorization

- **Given:** a multi-order tensor $\mathcal{X} \in \mathbb{N}^{N_1 \times \cdots \times N_M \times \ell}$
- **Object:** Estimate $\mathcal{D}, \mathcal{S}$ that minimizes $\Delta < \mathcal{X}; \mathcal{D}, \mathcal{S} >$
- **Idea1:** Alternating shifts in dynamic interaction set $\mathcal{D}$



Streaming DISMO factorization

- **Given:** a multi-order tensor $\mathcal{X} \in \mathbb{N}^{N_1 \times \cdots \times N_M \times \ell}$
- **Object:** Estimate $\mathcal{D}, \mathcal{S}$ that minimizes $\Delta < \mathcal{X}; \mathcal{D}, \mathcal{S} >$
- **Idea2:** Smoothly update $\mathcal{S}$ to extract long-term patterns

$$\mathbf{P}^{(m)}_{new} \leftarrow \mathbf{P}^{(m)} + (\mathbf{X}^{(m)} - \boxed{\hat{\mathbf{X}}^{(m)}_{d}})(\odot^{M+1}_{i \neq m} \mathbf{S}^{(i)}),$$
$$\mathbf{Q}^{(m)}_{new} \leftarrow \mathbf{Q}^{(m)} + (\otimes^{M+1}_{i \neq m} \mathbf{S}^{(i)\top} \mathbf{S}^{(i)}),$$
$$\mathbf{S}^{(m)}_{new} \leftarrow \mathbf{P}^{(m)} (\mathbf{Q}^{(m)})^{\dagger}.$$

Incremental CP decomposition

Computed with $\mathcal{D}_{best}$

Outline

- Motivation
- Problem definition
- Proposed model
- Proposed algorithms
- **Experiments**
- **Conclusion**

Experiments

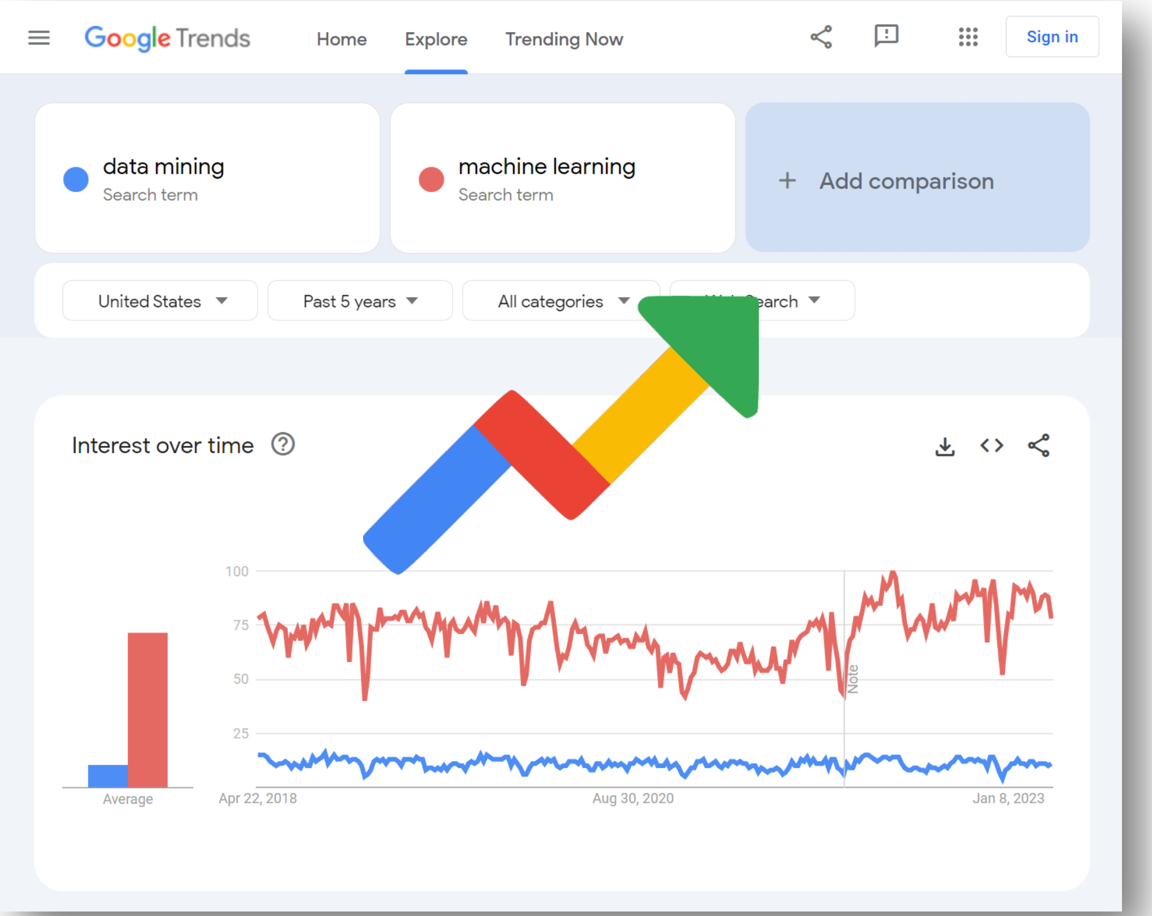
- Q1. Accuracy
- Q2. Scalability
- Q3. Effectiveness

Experiments

- Datasets
  - 13 years (2008 ~ 2020)
  - 50 states of the US

Table 3: GoogleTrends query sets.

Name	Query
Ecommerce	Amazon/Apple/BestBuy/Costco/Craigslist/Ebay/Etsy/HomeDepot/Kohls/Macys/Target/Walmart
VoD	AppleTV/Disney/ESPN/HBO/Hulu/Netflix/Sling/YouTube
Facilities	Aquarium/Bookstore/Gym/Library/Museum/Theater/Zoo
Sweets	Cake/Candy/Chocolate/Cookie/Cupcake/Gum/Icecream/Pie/Pudding



https://trends.google.com/home

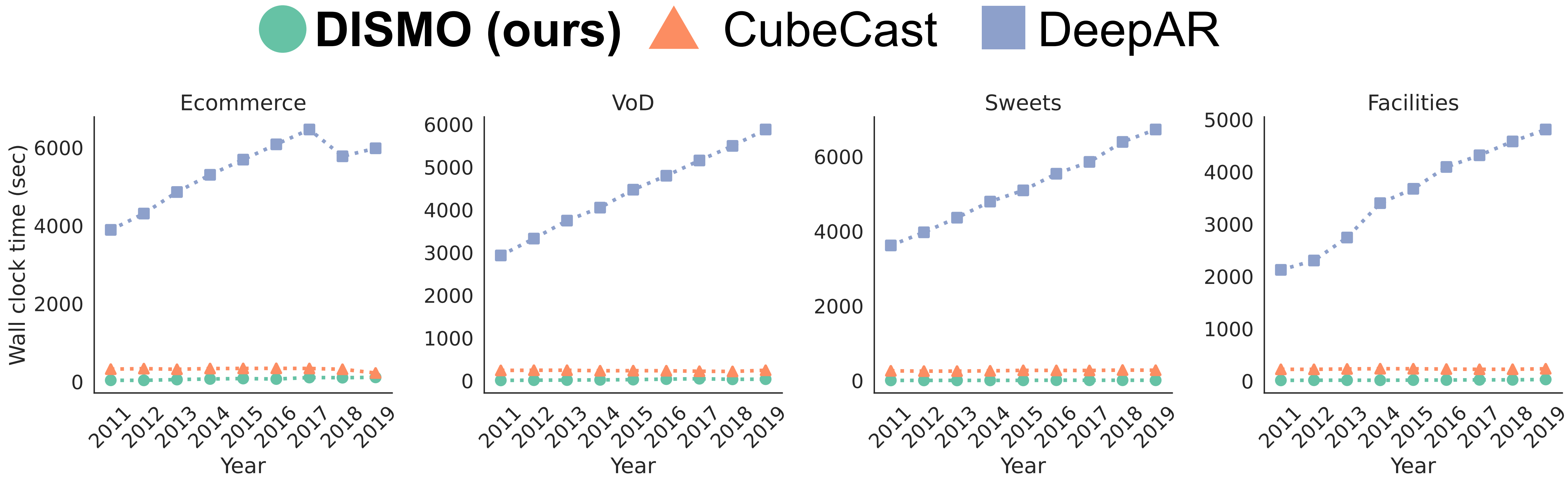
Q1. Accuracy

Table 2: Forecasting performance comparison. DISMO outperformed its competitors in terms of RMSE.

data	ℓ_s	DISMO	DISMO-naive	CubeCast	DeepAR	SMF	TRMF
Ecommerce	13	**0.0316 ± 0.0081**	0.0881 ± 0.1239	<u>0.0492 ± 0.0339</u>	0.0638 ± 0.0149	0.0591 ± 0.0163	0.1755 ± 0.0207
	26	**0.0368 ± 0.0103**	0.1122 ± 0.1230	<u>0.0455 ± 0.0269</u>	0.0721 ± 0.0159	0.0604 ± 0.0164	0.1758 ± 0.0198
	39	**0.0425 ± 0.0147**	0.1613 ± 0.1420	<u>0.0431 ± 0.0219</u>	0.0776 ± 0.0167	0.0615 ± 0.0166	0.1781 ± 0.0203
Facilities	13	**0.0356 ± 0.0062**	<u>0.0445 ± 0.0076</u>	0.0890 ± 0.0089	0.0593 ± 0.0146	0.0472 ± 0.0115	0.1390 ± 0.0183
	26	**0.0383 ± 0.0108**	<u>0.0458 ± 0.0093</u>	0.0883 ± 0.0119	0.0666 ± 0.0156	0.0471 ± 0.0125	0.1388 ± 0.0161
	39	**0.0406 ± 0.0131**	<u>0.0466 ± 0.0105</u>	0.0865 ± 0.0137	0.0704 ± 0.0155	0.0482 ± 0.0130	0.1381 ± 0.0152
Sweets	13	**0.0276 ± 0.0146**	0.0297 ± 0.0144	0.0422 ± 0.0209	0.0340 ± 0.0167	<u>0.0280 ± 0.0148</u>	0.0823 ± 0.0124
	26	**0.0279 ± 0.0150**	0.0298 ± 0.0146	0.0405 ± 0.0183	0.0357 ± 0.0167	<u>0.0286 ± 0.0151</u>	0.0826 ± 0.0127
	39	<u>0.0283 ± 0.0149</u>	0.0299 ± 0.0146	0.0393 ± 0.0172	0.0371 ± 0.0166	**0.0275 ± 0.0153**	0.0830 ± 0.0128
VoD	13	**0.0293 ± 0.0121**	0.0558 ± 0.0136	0.0479 ± 0.0294	0.1233 ± 0.0438	<u>0.0447 ± 0.0161</u>	0.2297 ± 0.0489
	26	**0.0336 ± 0.0155**	0.0578 ± 0.0145	<u>0.0423 ± 0.0248</u>	0.1433 ± 0.0435	0.0452 ± 0.0158	0.2280 ± 0.0511
	39	<u>0.0384 ± 0.0194</u>	0.0592 ± 0.0150	**0.0380 ± 0.0203**	0.1505 ± 0.0419	0.0450 ± 0.0160	0.2275 ± 0.0604

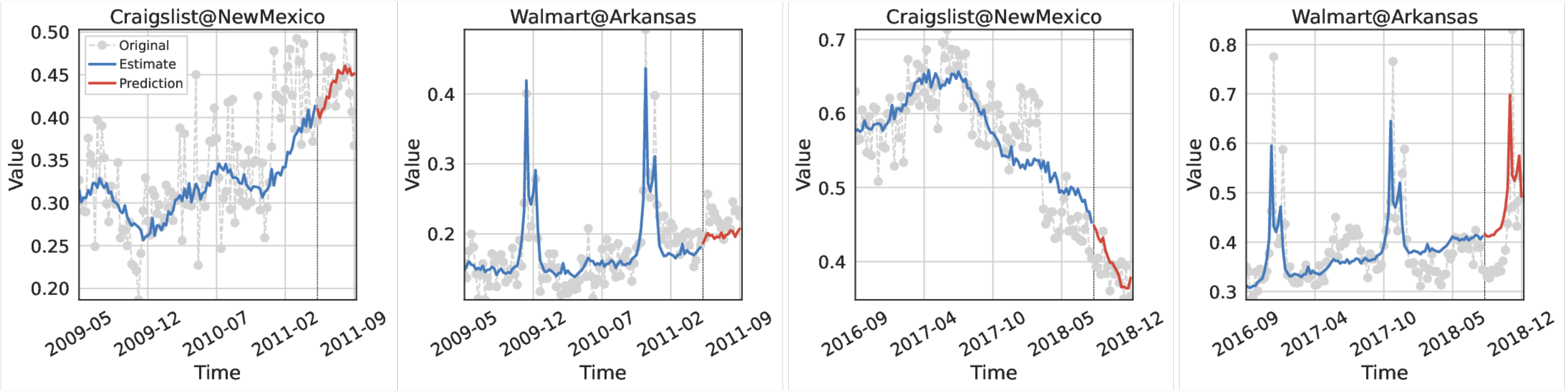
Accurate to forecast multi-steps ahead tensors

Q2. Scalability



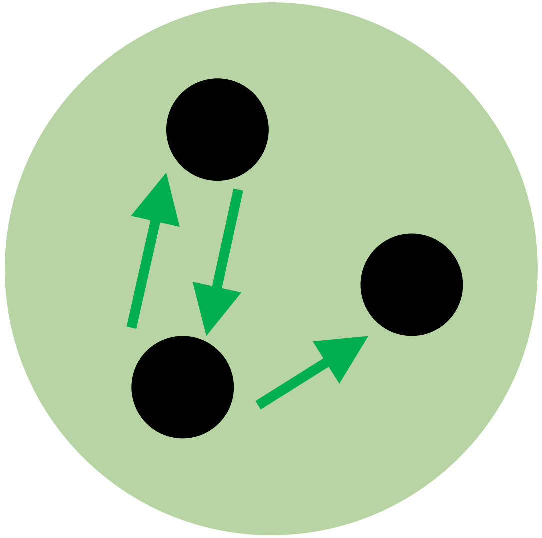
DISMO: scalable at any time

Q3. Effectiveness



DISMO can effectively model dynamic interactions and seasonality

Conclusion



Interpretable

Our non-linear model reveals important relationships between latent groups in multi-order tensor streams

Dynamic

Shifting trends can be detected in real time by maintaining multiple multi-aspect factors

Automatic

The number of any factors composed by our method are determined without any parameter tuning

Thank you!

Code & datasets