ICE: Item Concept Embedding via Textual Information

Ting-Hsiang Wang
CITI, Academia Sinica, Taiwan

Joint work with Prof. Chuan-Ju Wang, Hsiu-Wei Yang,
Bo-Sin Chang, Prof. Ming-Feng Tsai

August 8, 2017
SIGIR, Tokyo, Japan

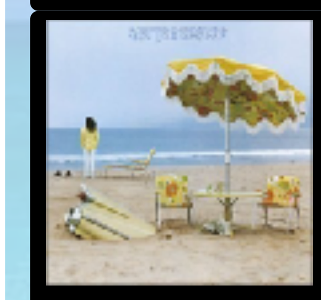
Normal search only retrieve the concept “beach”

beach

beach songs

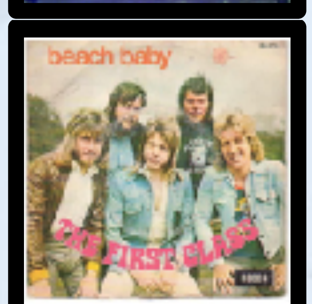
Girls on the **Beach**
Album: All Summer Long(1964)
Artist: The **Beach** Boys
… On the **beach** you’ll find them there …

Rockaway **Beach**
Album: Rocket to Russia (1977)
Artist: Ramones
… Rock-rock, Rockaway **Beach** …



On the **Beach**
Album: On the **Beach** (1974)
Artist: Neil Young
… out here on the **beach** …

Private **Beach** Party
Album: Private **Beach** Party (1985)
Artist: Gregory Isaacs
… At the private **beach** party…

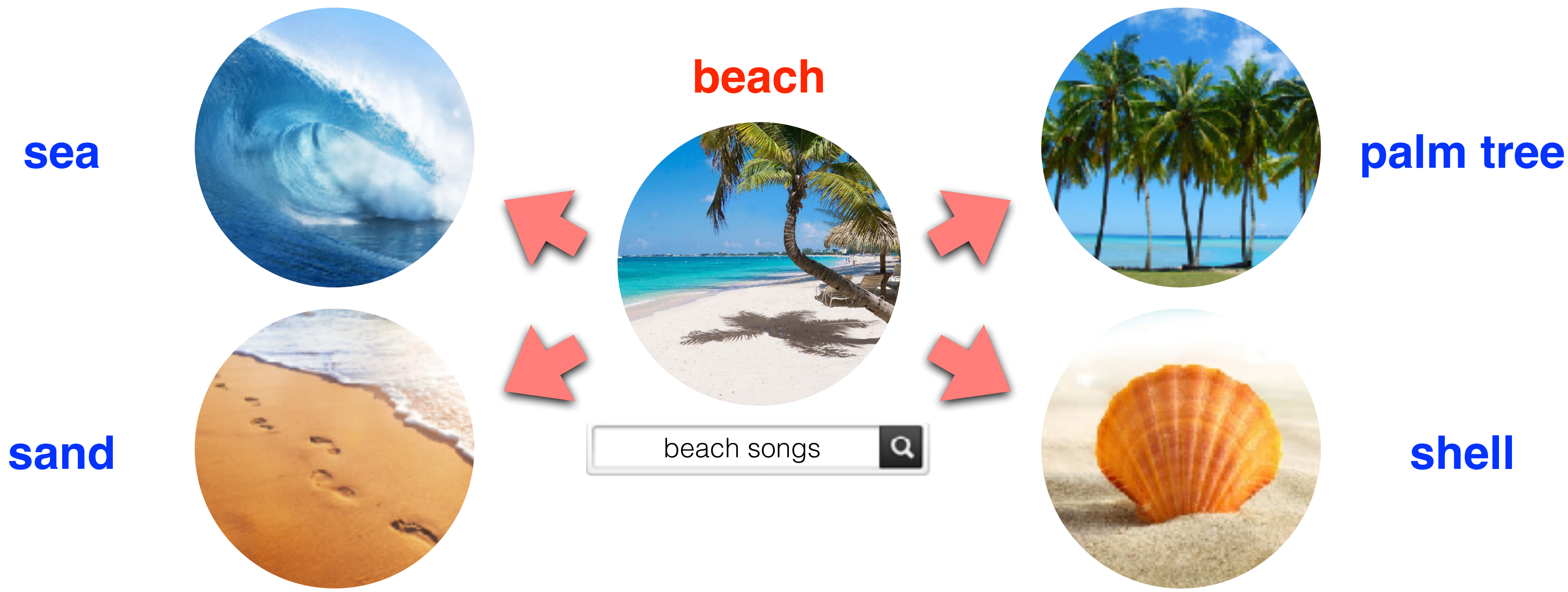


Private **Beach** Baby
Album: single (1974)
Artist: The First Class
… **Beach** baby, **beach** baby…

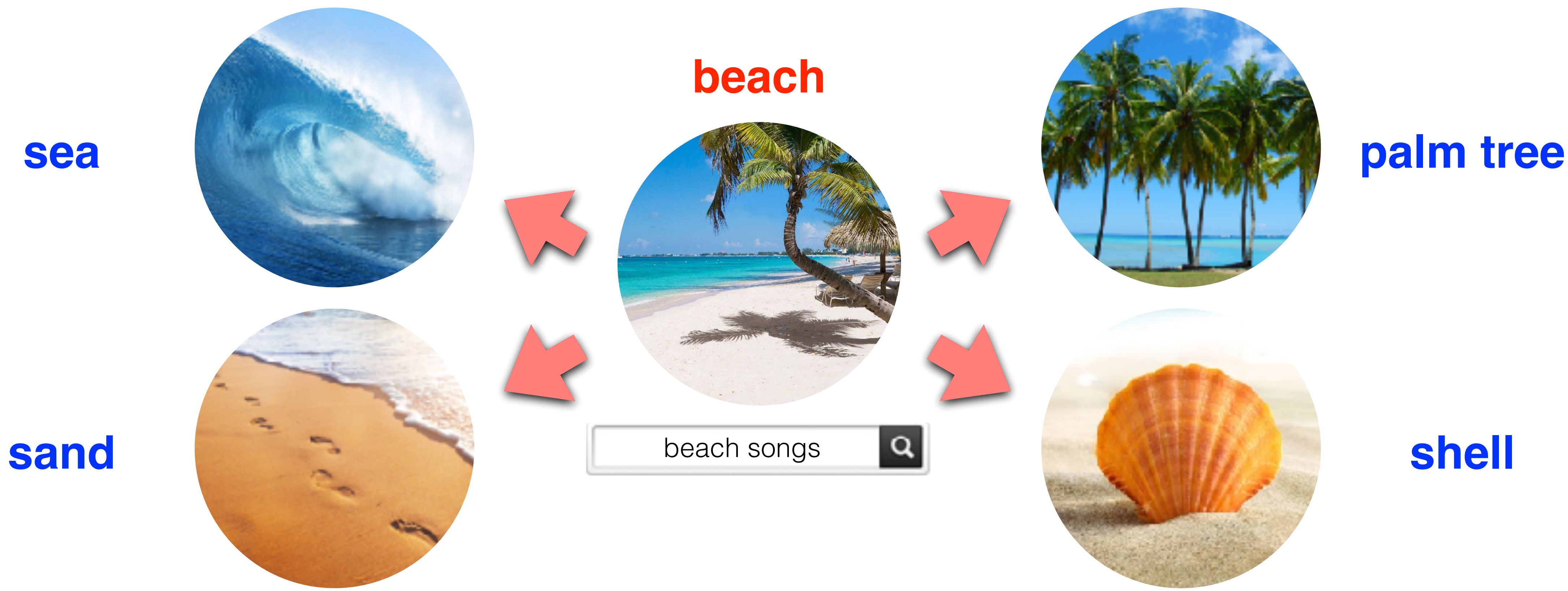
"Beach" has many correlated concepts

beach

Expand concept "beach" to sea, sand, …



Capture similar concepts = diverse AND relevant



Girls on the **Beach**
Album: All Summer Long (1964)
Artist: The **Beach** Boys
… On the **beach** you'll find them there …

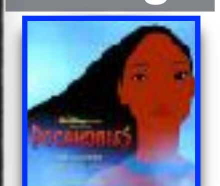
Palmtree
Album: single (2015)
Artist: Mandelbarth
… Under the **palm trees** is where we …

Sand And **Sea**
Album: That's Life (1966)
Artist: Frank Sinatra
… **Sand** and **sea**, **sea** and **sand** …

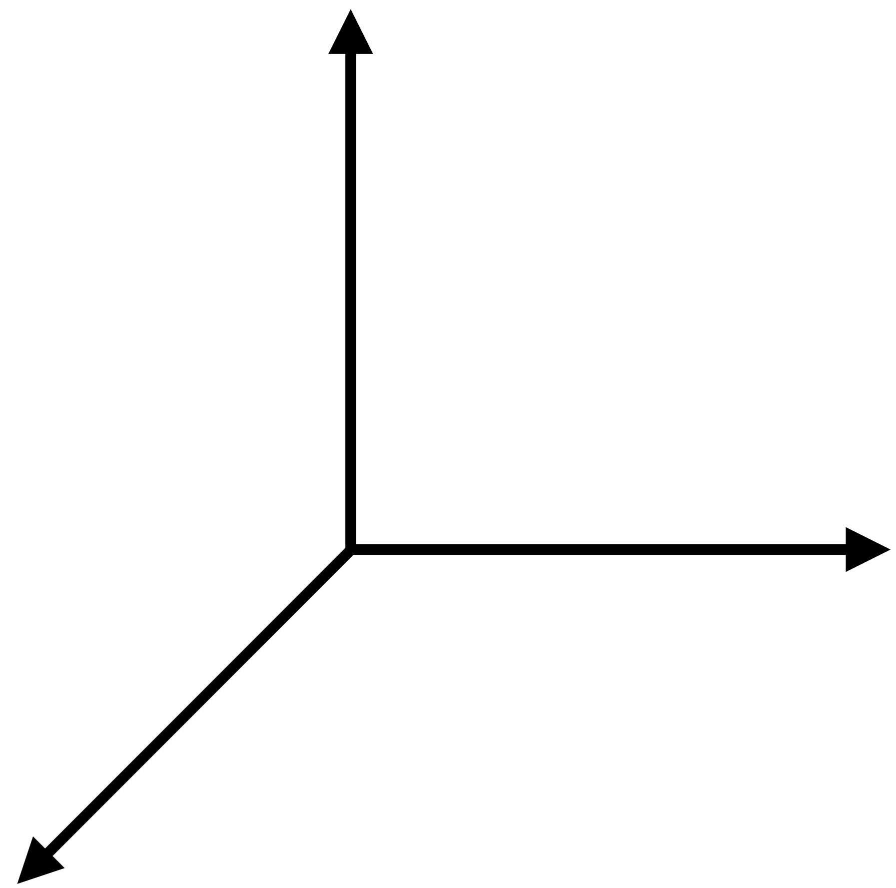
Sea **Shells**
Album: Lipslide (1997)
Artist: Sarah Cracknell
… Hey little **sea shell**, I need a cue…

Embed items and concepts in space such that…

Song	Lyrics	Word
	… On the **beach** you'll find them there …	**beach**
	… **Sand** and **sea**, **sea** and **sand** …	**sand** **sea**
	… Hey little **sea shell**, I need a cue…	**sea** **shell**
	… Under the **palm trees** is where we …	**palm tree**

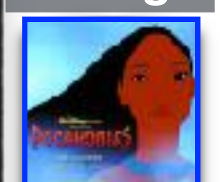
Song	Lyrics	Word
	… all the voices of the **mountains** … All you own is **earth** until …	**mountain** **earth**
	… Far away o'er the **mountains**, … with the **cliffs** of Doneen …	**mountain** **cliff**

Embedding Space

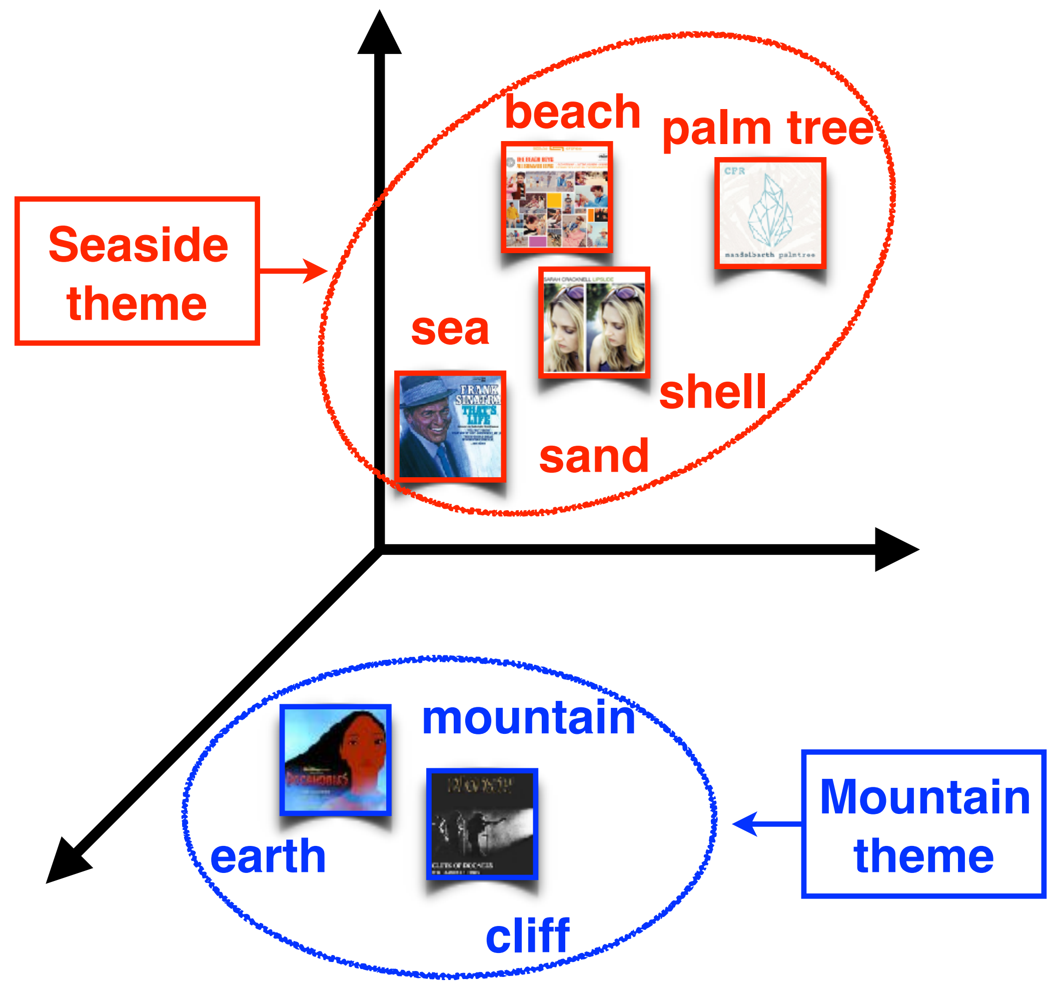


… similar items and concepts flock together

Song	Lyrics	Word
	… On the **beach** you'll find them there …	**beach**
	… **Sand** and **sea**, **sea** and **sand** …	**sand** **sea**
	… Hey little **sea shell**, I need a cue…	**sea shell**
	… Under the **palm trees** is where we …	**palm tree**

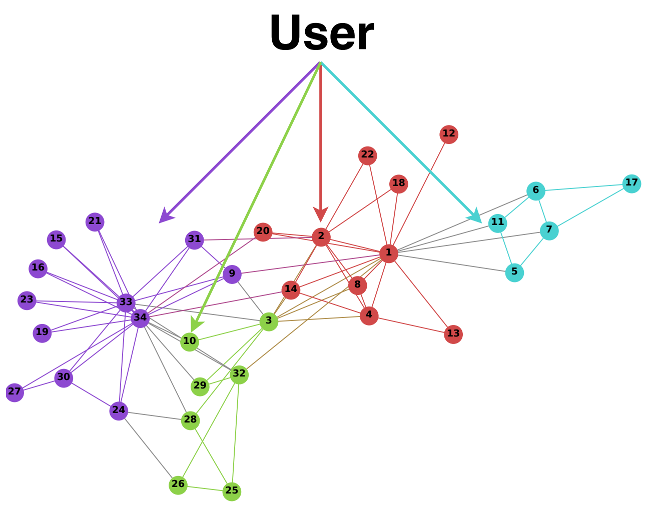
Song	Lyrics	Word
	… all the voices of the **mountains** … All you own is **earth** until …	**mountain** **earth**
	… Far away o'er the **mountains**, … with the **cliffs** of Doneen …	**mountain** **cliff**

Embedding Space

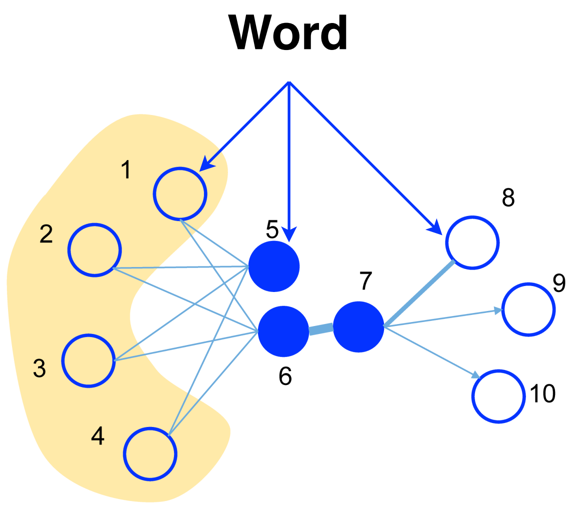


… and different ones separate.

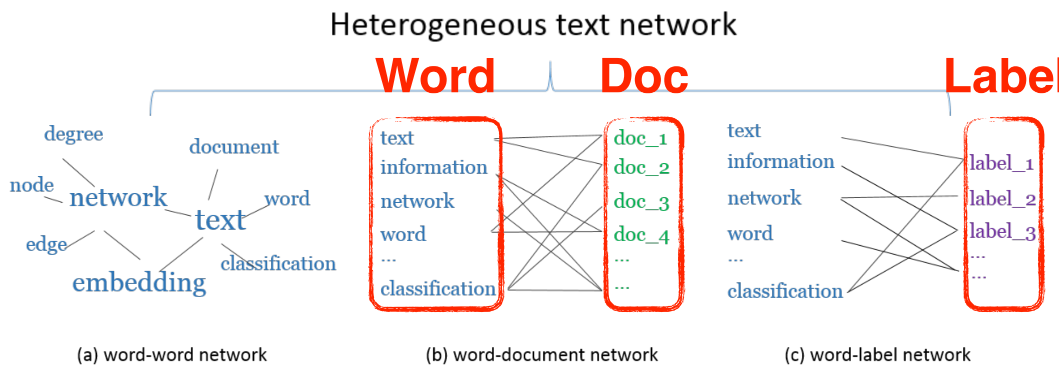
Related works in graph embedding



DeepWalk (Perozzi et al., 2014)



LINE (Tang, et al., 2015)



PTE (Tang, et al. 2015)

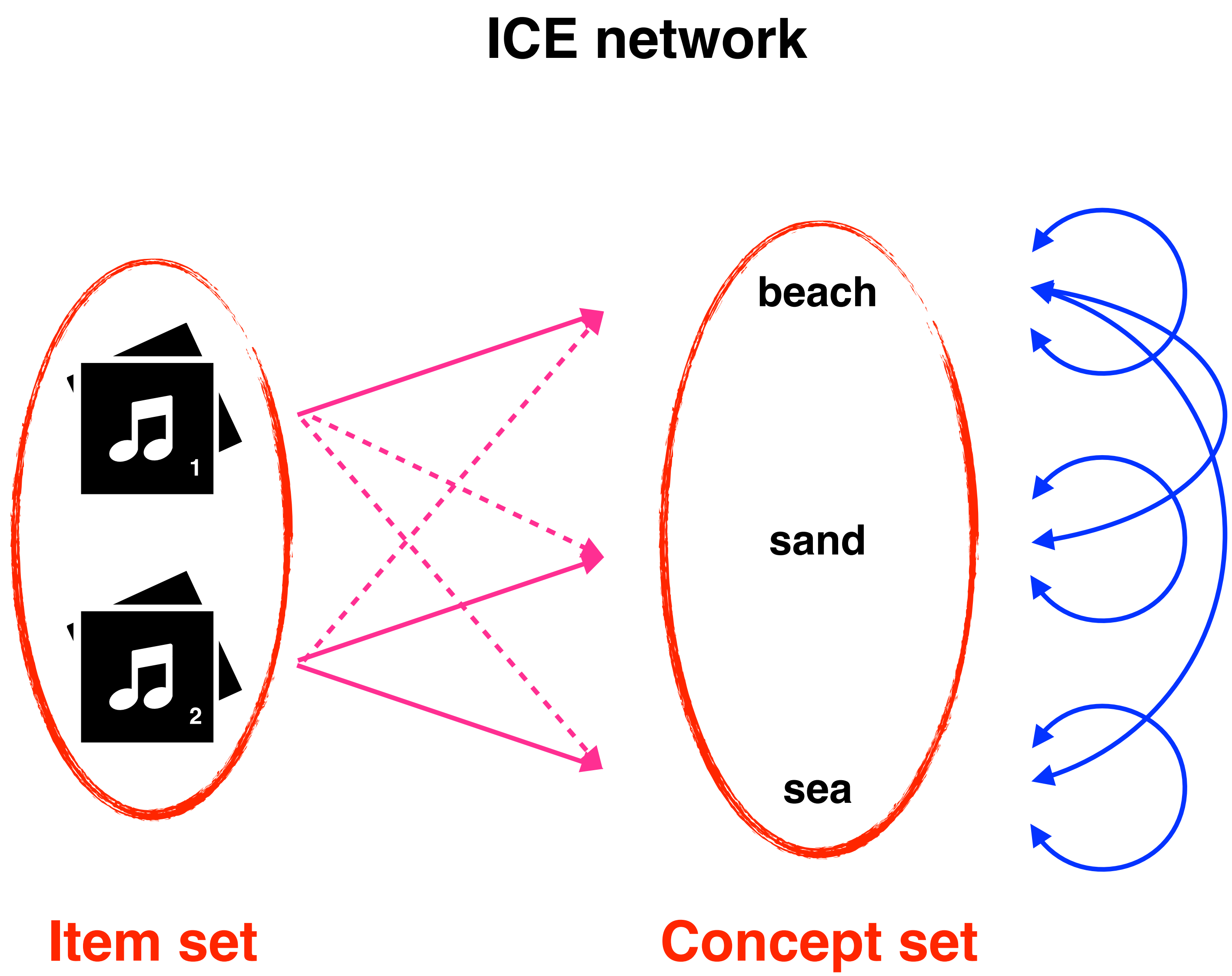
- All the above-mentioned methods focus on **homogeneous** tasks:
  - **DeepWalk**: Homogeneous social networks (users with social relations).
  - **LINE**: Homogeneous social networks or word-word networks, etc.
  - **PTE**: Heterogeneous text network but still for homogeneous tasks, such as document classification.
- However, the inter-retrieval task between concepts and items is **heterogeneous**:
  - e.g., word-to-song retrieval, movie-to-word retrieval, etc.

Our Proposal: Item Concept Embedding (ICE)

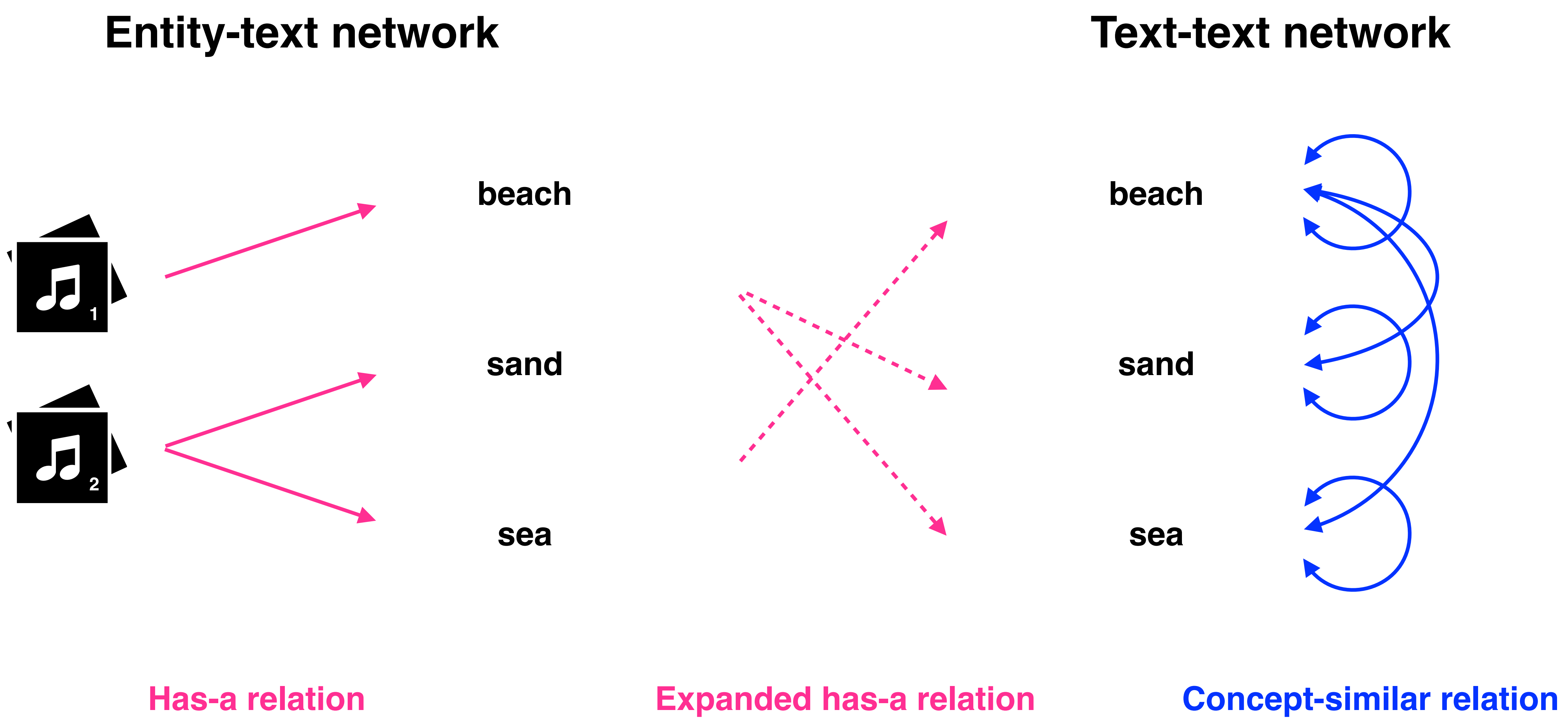
- Main Contributions:

  1. Propose item concept embedding (ICE) approach to model the concepts of items via associated textual information.
  2. Integrate heterogeneous nodes and relations in network using generalized matrix operations.
  3. Learn embeddings capable to retrieve conceptually diverse and relevant results that support both homogeneous and heterogeneous tasks.

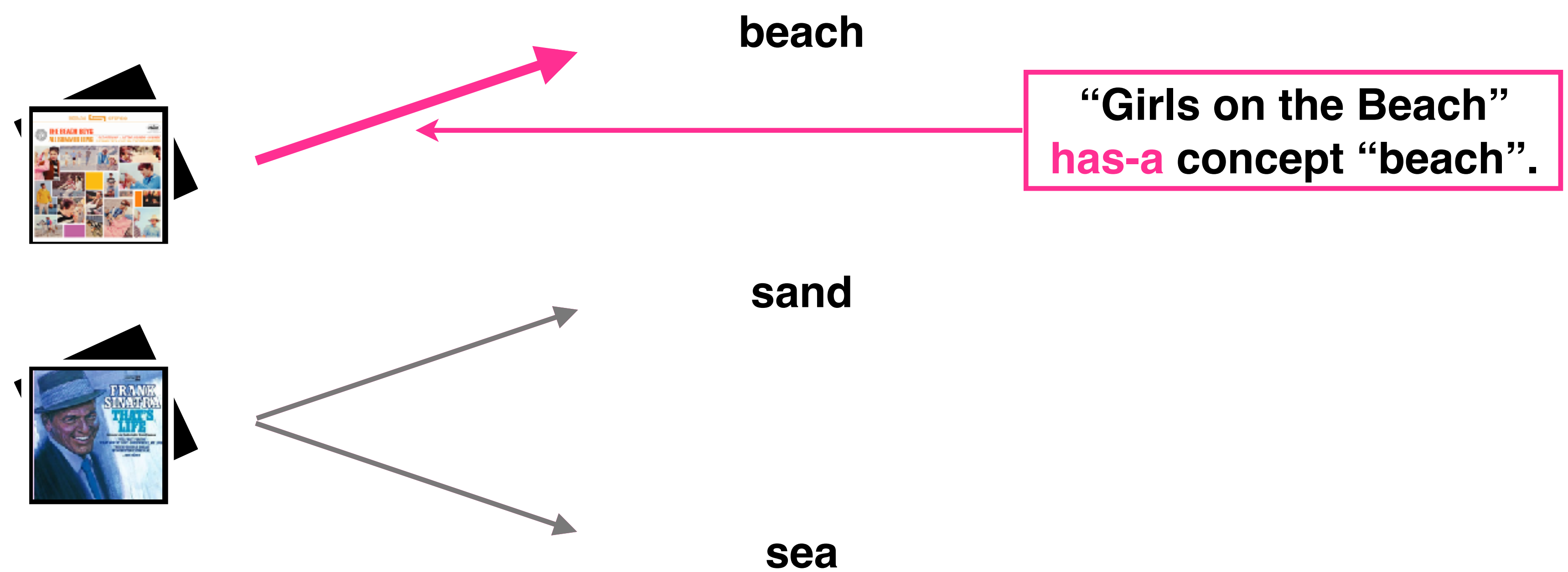
ICE network is an unified network composed of…



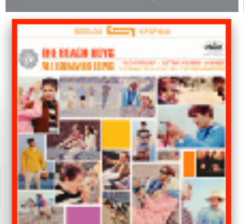
… 2 basis networks and 3 relations



Entity-text network manages item concepts

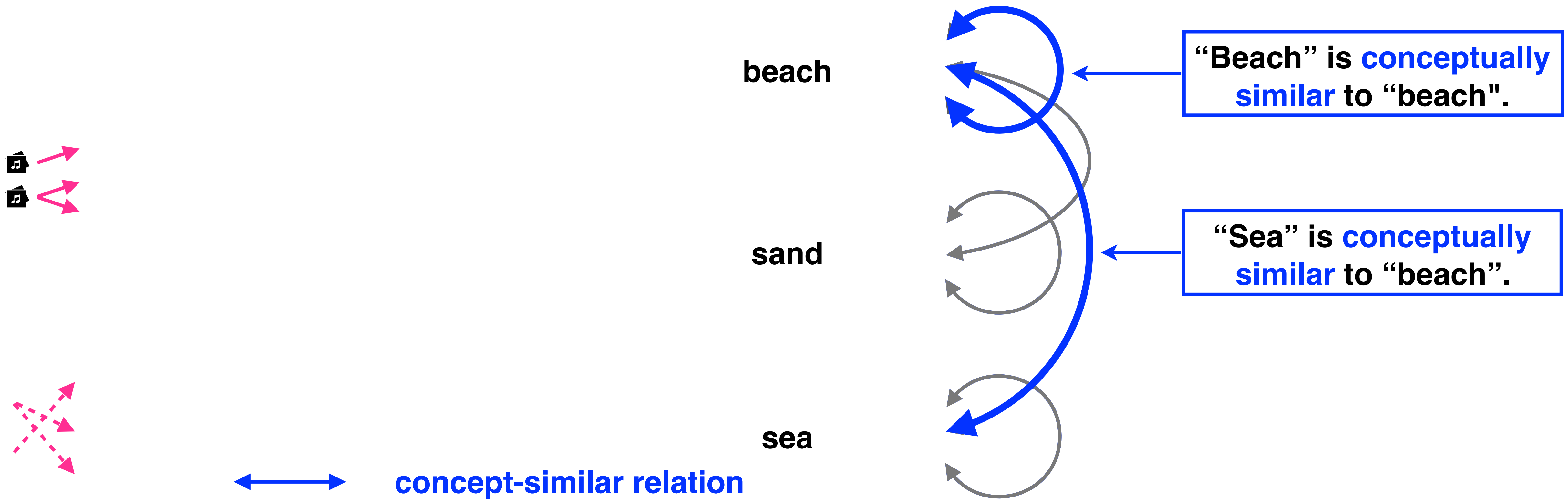


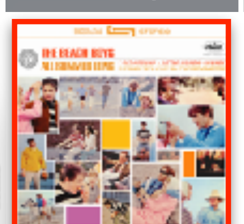
has-a relation

Songs	Lyrics	Words
	… On the **beach** you'll find them there …	**beach**
	… **Sand** and **sea**, **sea** and **sand** …	**sand** **sea**

- Manage the has-a relation between each item and their representative concept words.
- Concept words for each item are picked according to the TF-IDF score.
- Heterogeneous, directed, and bipartite.

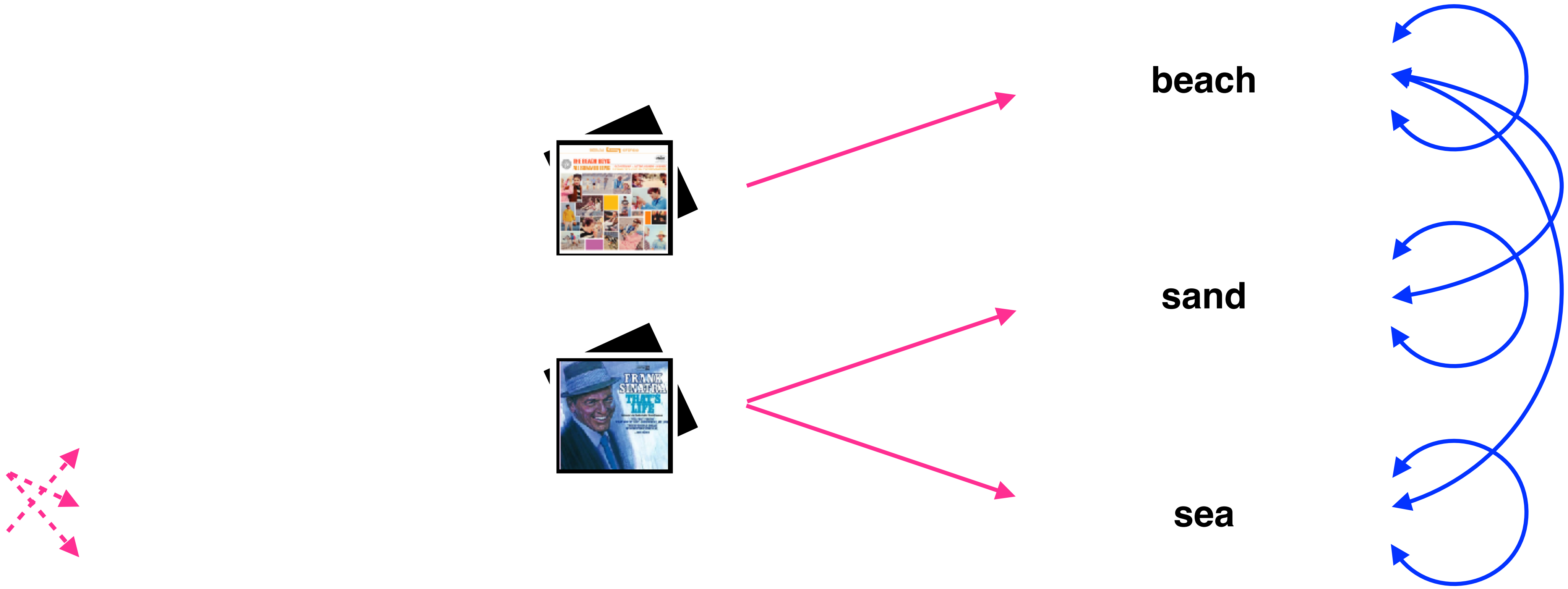
Text-text network manages concept similarity

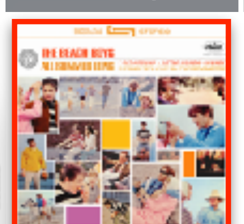


Songs	Lyrics	Words
	… On the **beach** you'll find them there …	beach
	… **Sand** and **sea**, **sea** and **sand** …	sand sea

- Manage the concept-similar relation between each concept words.
- Conceptually similar words are connected according to the cosine similarity between word embeddings.
- Homogeneous and bi-directed.

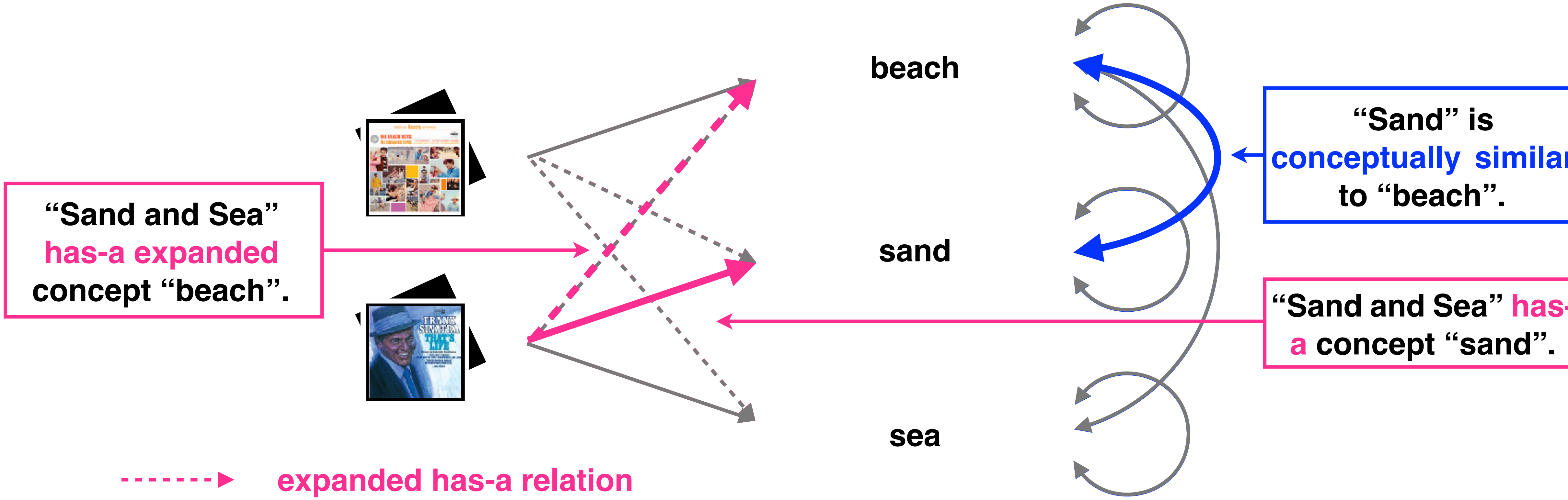
ICE network combines E-T and T-T network and …



Songs	Lyrics	Words
	… On the **beach** you'll find them there …	**beach**
	… **Sand** and **sea**, **sea** and **sand** …	**sand** **sea**

- Combine entity-text network, text-text network, and expanded has-a relation.
- Manage the expanded has-a relation between each item and their expanded concept words.
- Establish relation to expanded concept words via the conceptually similar words of each item.
- Heterogeneous nodes and relations.

… manages the expanded has-a relation

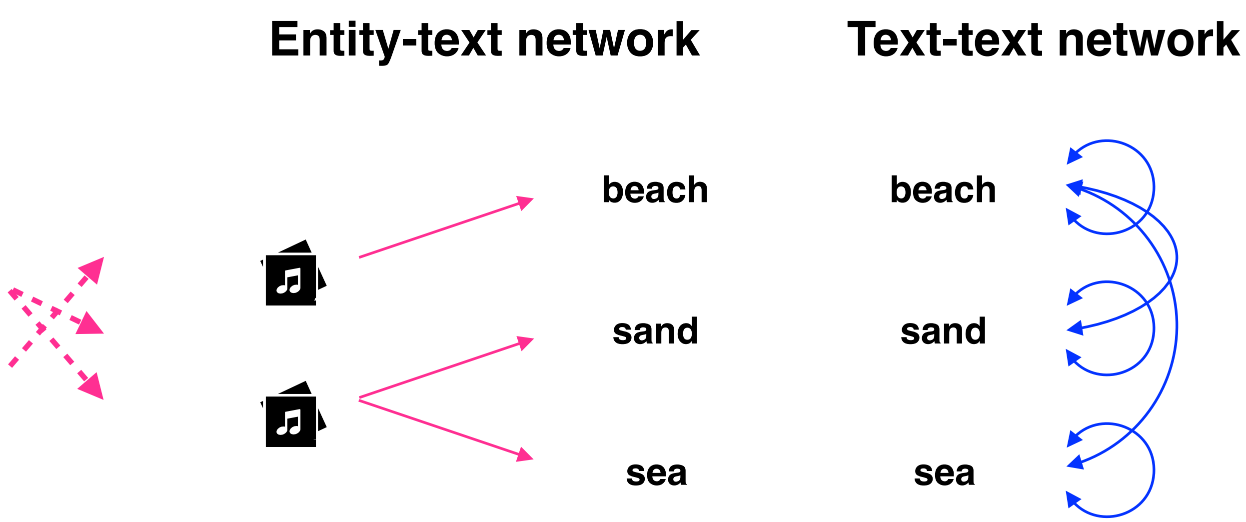


Songs	Lyrics	Words
	… On the **beach** you’ll find them there …	**beach**
	… **Sand** and **sea**, **sea** and **sand** …	**sand** **sea**

- Combine entity-text network, text-text network, and expanded has-a relation.
- Manage the expanded has-a relation between each item and their expanded concept words.
- Establish relation to expanded concept words via the conceptually similar words of each item.
- Heterogeneous nodes and relations.

Construct graph via generalized matrix operation

- Step 1: Establish expanded has-a relation in ET network.



Construct graph via generalized matrix operation

- Step 1: Establish expanded has-a relation in ET network.

Entity-text network

$$\begin{array}{c c} & \begin{array}{ccc} W_1 & W_2 & W_3 \end{array} \\ \begin{array}{c} I_1 \\ I_2 \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \end{array}$$

$$M_{G_{et}}$$

Text-text network

$$\begin{array}{c c} & \begin{array}{ccc} W_1 & W_2 & W_3 \end{array} \\ \begin{array}{c} W_1 \\ W_2 \\ W_3 \end{array} & \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \end{array}$$

$$M_{G_{tt}}$$

Construct graph via generalized matrix operation

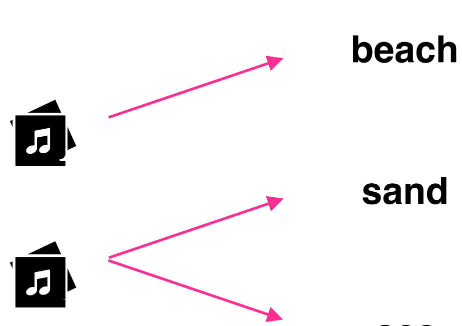
- Step 1: Establish expanded has-a relation in ET network.

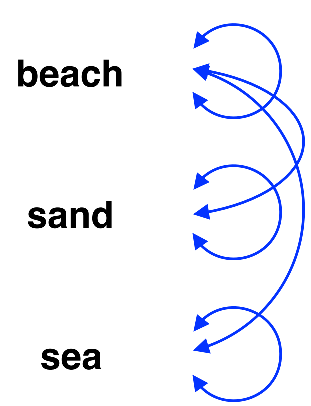
Entity-text network | **Text-text network** | $A = M_{G_{et}} \cdot M_{G_{tt}}$

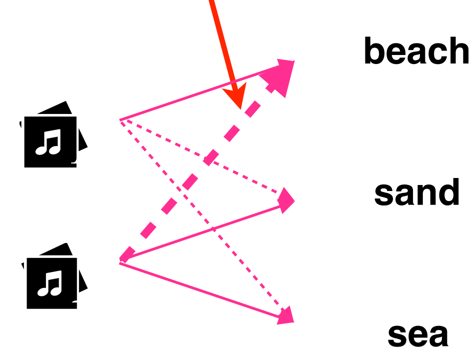
$$\begin{array}{c c} & \begin{matrix} W_1 & W_2 & W_3 \end{matrix} \\ \begin{matrix} I_1 \\ I_2 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \end{array} \cdot \begin{array}{c c} & \begin{matrix} W_1 & W_2 & W_3 \end{matrix} \\ \begin{matrix} W_1 \\ W_2 \\ W_3 \end{matrix} & \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \end{array} = \begin{array}{c c} & \begin{matrix} W_1 & W_2 & W_3 \end{matrix} \\ \begin{matrix} I_1 \\ I_2 \end{matrix} & \begin{bmatrix} 1 & 1 & 1 \\ \boxed{2} & 1 & 1 \end{bmatrix} \end{array}$$

$M_{G_{et}}$ $\qquad$ $M_{G_{tt}}$

More encompassing concept







Construct graph via generalized matrix operation

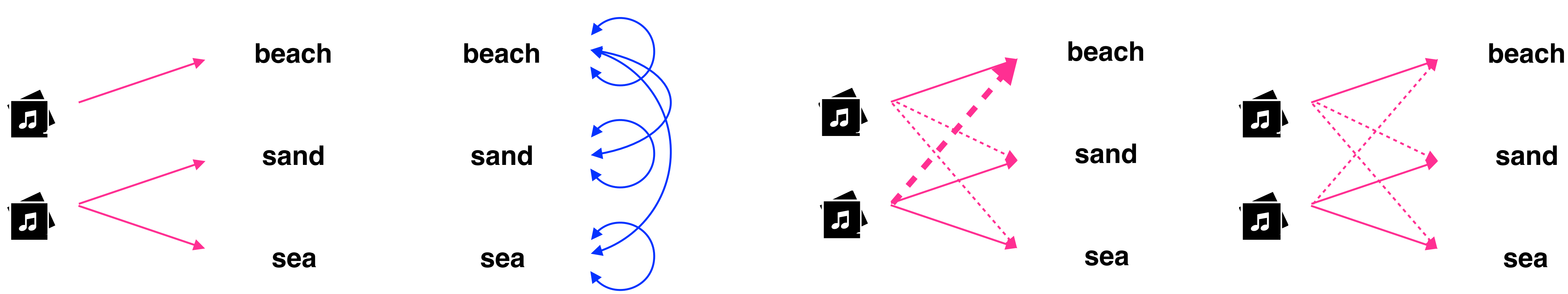
- Step 2: Convert the dot product to a binary matrix $\tilde{A}$.

Entity-text network **Text-text network**

$$A = M_{G_{et}} \cdot M_{G_{tt}} \qquad \tilde{A} = \left(\mathbb{1}_{\{a_{ij}>0\}}\right)$$

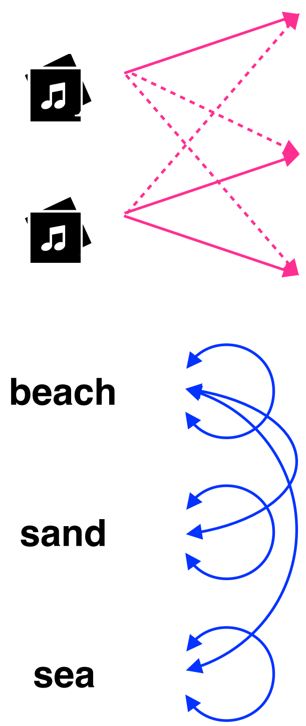
$$\begin{array}{c} \\ I_1 \\ I_2 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} W_1 & W_2 & W_3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \end{array} \cdot \begin{array}{c} \\ W_1 \\ W_2 \\ W_3 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} W_1 & W_2 & W_3 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \end{array} = \begin{array}{c} \\ I_1 \\ I_2 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} W_1 & W_2 & W_3 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} \end{array} \qquad \begin{array}{c} \\ I_1 \\ I_2 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} W_1 & W_2 & W_3 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \end{array}$$

$M_{G_{et}}$ $\qquad$ $M_{G_{tt}}$



Construct graph via generalized matrix operation

- Step 3: Augment binary matrix with the text-text matrix.



$$\tilde{A}$$

$$\begin{array}{c} \\ I_1 \\ I_2 \end{array} \begin{array}{c} \begin{array}{ccc} W_1 & W_2 & W_3 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \end{array}$$

$$\begin{array}{c} \\ W_1 \\ W_2 \\ W_3 \end{array} \begin{array}{c} \begin{array}{ccc} W_1 & W_2 & W_3 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \end{array}$$

$$M_{G_{tt}}$$

Construct graph via generalized matrix operation

- Step 3: Augment binary matrix with the text-text matrix.

$$\tilde{A}$$

$$\begin{array}{c} \\ I_1 \\ I_2 \end{array} \begin{array}{c} \begin{array}{ccc} W_1 & W_2 & W_3 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \end{array}$$

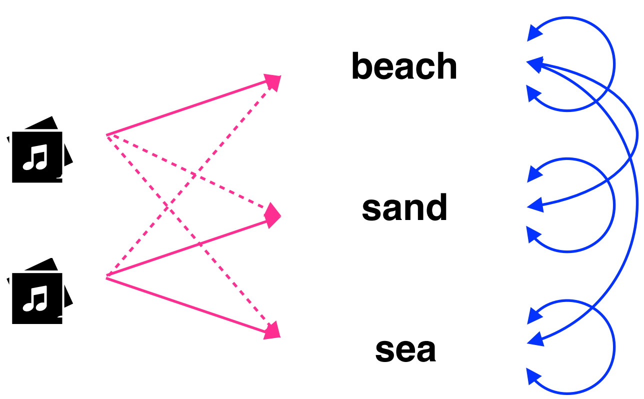
$$\begin{array}{c} \\ W_1 \\ W_2 \\ W_3 \end{array} \begin{array}{c} \begin{array}{ccc} W_1 & W_2 & W_3 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \end{array}$$

$$M_{G_{tt}}$$

$$M_{G_{ice}} = \begin{bmatrix} \tilde{A} \\ M_{G_{tt}} \end{bmatrix}$$

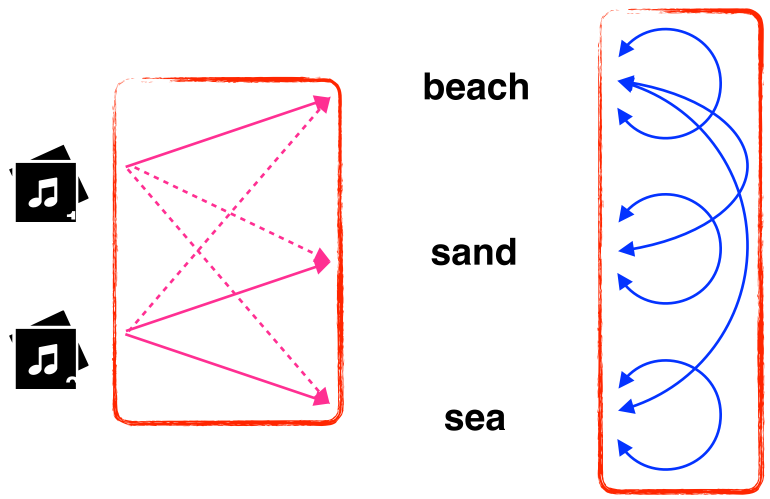
$$\begin{array}{c} \\ I_1 \\ I_2 \\ W_1 \\ W_2 \\ W_3 \end{array} \begin{array}{c} \begin{array}{ccc} W_1 & W_2 & W_3 \end{array} \\ \left[\begin{array}{ccc} 1 & 1 & 1 \\ 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right] \end{array}$$

G_{ice}: ICE Network

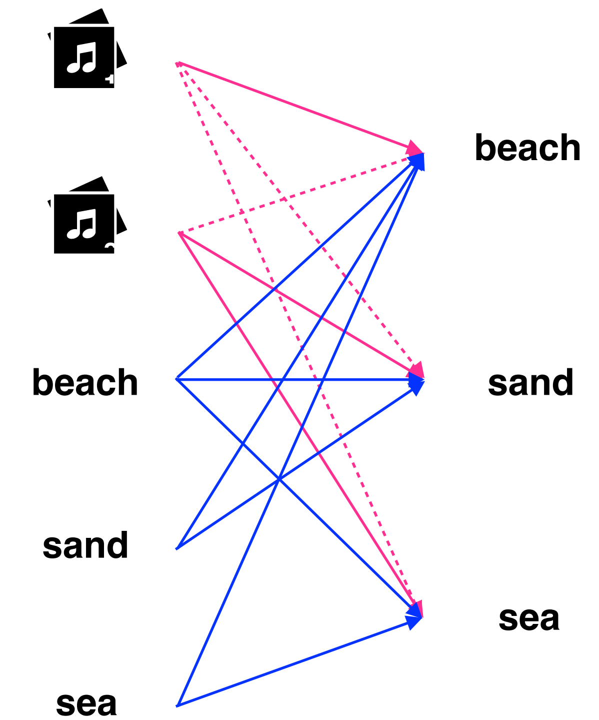


Modeling of neighborhood proximity

- Intuition: Maintain homogeneous neighborhood.



(a) ICE network



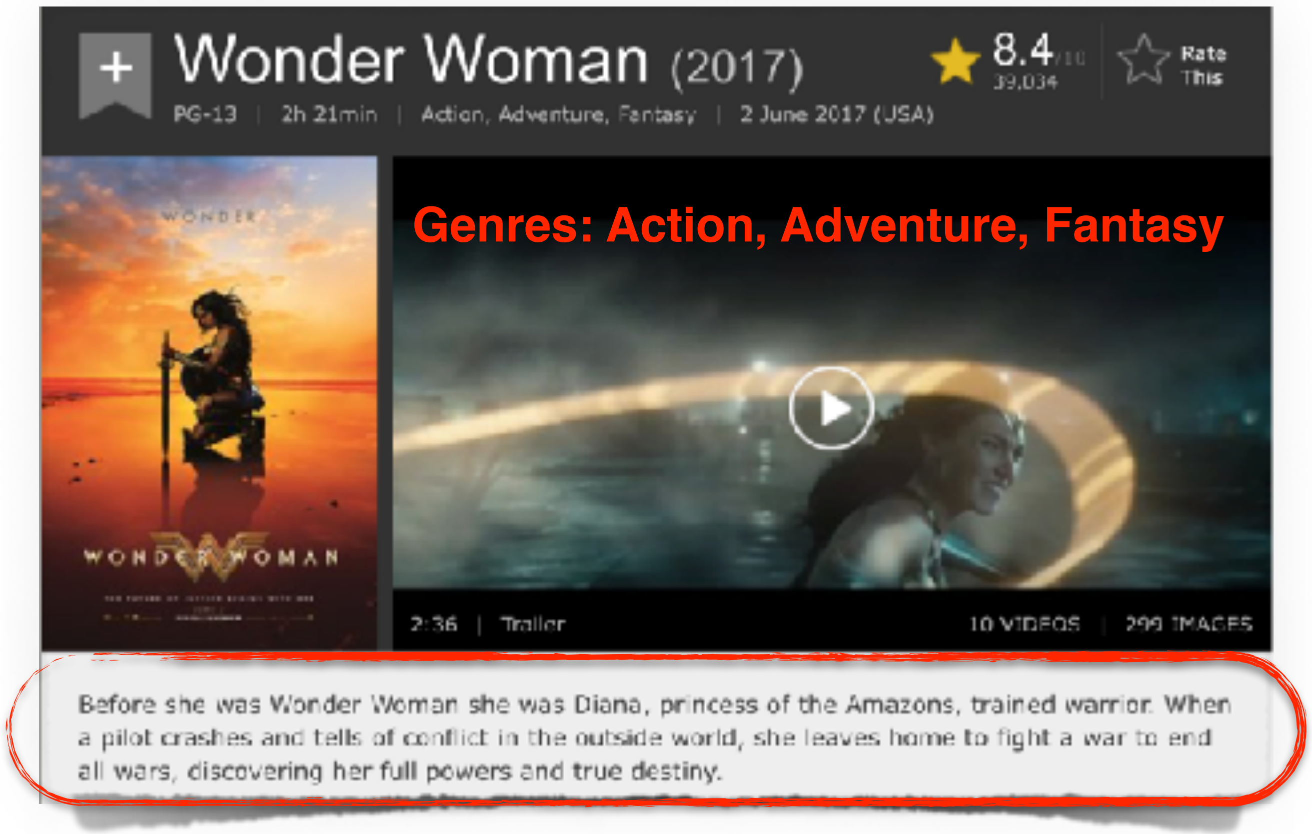
(b) Homogeneous context

- Jointly minimize the KL divergence of objective functions:

$$O_{ice} = -\left(\sum_{(n_i, n_\ell) \in \tilde{E}_{et}} x_{i\ell} \log P(n_\ell \mid n_i) + \sum_{(n_w, n_\ell) \in E_{tt}} x_{w\ell} \log P(n_\ell \mid n_w) \right)$$

Datasets: Real-world movie and music datasets

- IMDB (movie) dataset:
  - Movie, plots, and genres
- KKBOX (music) dataset:
  - Song and lyrics



	IMDB	KKBOX
# movies/songs	36,586	33,106
Average text length	65.0	215.24
Average # unique words	47.8	81.37
Vocabulary size	66,924	101,395
# single genres	28	-
# multi-label genres	915	-

稻香 — 周杰倫（Jay Chou）

對 / 這個 / 世界 / 如果/ 你 / 有 / 太多 / 的 / 抱怨
跌倒 / 了 / 就 / 不敢 / 繼續/ 往前 / 走
為什麼 / 人 / 要 / 這麼 / 的 / 脆弱/ / 墮落
請 / 你 / 打開 / 電視 / 看看
多少 / 人 / 為 / 生命 / 在 / 努力 / 勇敢 / 的 / 走 / 下去
我們 / 是不是 / 該 / 知足
珍惜 / 一切 / 就算 / 沒有 / 擁有
…

Experiment — Tasks and baselines

- Two types of tasks:
  1. Homogeneous:
     - Movie classification.
     - Movie-to-movie retrieval.
  2. Heterogeneous:
     - Word-to-movie retrieval. (Ex: Using “Killer” in Thriller movies.)
     - Movie-to-word retrieval.
     - Word-to-song retrieval. (Ex: Using contextual words.)
- Baselines:
  1. Traditional: Keyword-based (**KBR**), bag-of-words (**BOW**)
  2. Embedding: Bipartite (**BPT**), average embedding (**AVGEMB**)

Homogeneous: Movie genre classification

- Multi-label Movie Genre Classification (homogeneous):

Table 4: Movie genre classification task

W = # of rep words per item	$\|W\| = 10$					$\|W\| = 20$			
exp = # of exp. words per concept word	BOW	BPT	ICE (exp-3)	ICE (exp-5)		BOW	BPT	ICE (exp-3)	ICE (exp-5)
Exact match ratio	0.136	**0.160**	0.156	0.157		0.162	**0.182**	**0.182**	0.181
Micro-average F-measure	0.365	0.401	0.408	**0.410**	<	0.415	**0.464**	0.462	0.463
Macro-average F-measure	0.087	0.166	**0.170**	**0.170**		0.156	**0.229**	0.223	0.222

- Increasing the number of concept words used to represent an item improves the performance of the item embedding.

Comparable performance in homogeneous tasks

- Multi-label Movie Genre Classification (homogeneous):

Table 4: Movie genre classification task

W = # of concept words per item	$\vert W\vert = 10$				$\vert W\vert = 20$			
exp = # of exp. words per concept word	BOW	BPT	ICE (exp-3)	ICE (exp-5)	BOW	BPT	ICE (exp-3)	ICE (exp-5)
Exact match ratio	0.136	**0.160**	0.156	0.157	0.162	**0.182**	**0.182**	0.181
Micro-average F-measure	0.365	0.401	0.408	**0.410**	0.415	**0.464**	0.462	0.463
Macro-average F-measure	0.087	0.166	**0.170**	**0.170**	0.156	**0.229**	0.223	0.222

- ICE embeddings are suitable for homogeneous tasks.

Heterogeneous: Word-to-movie retrieval

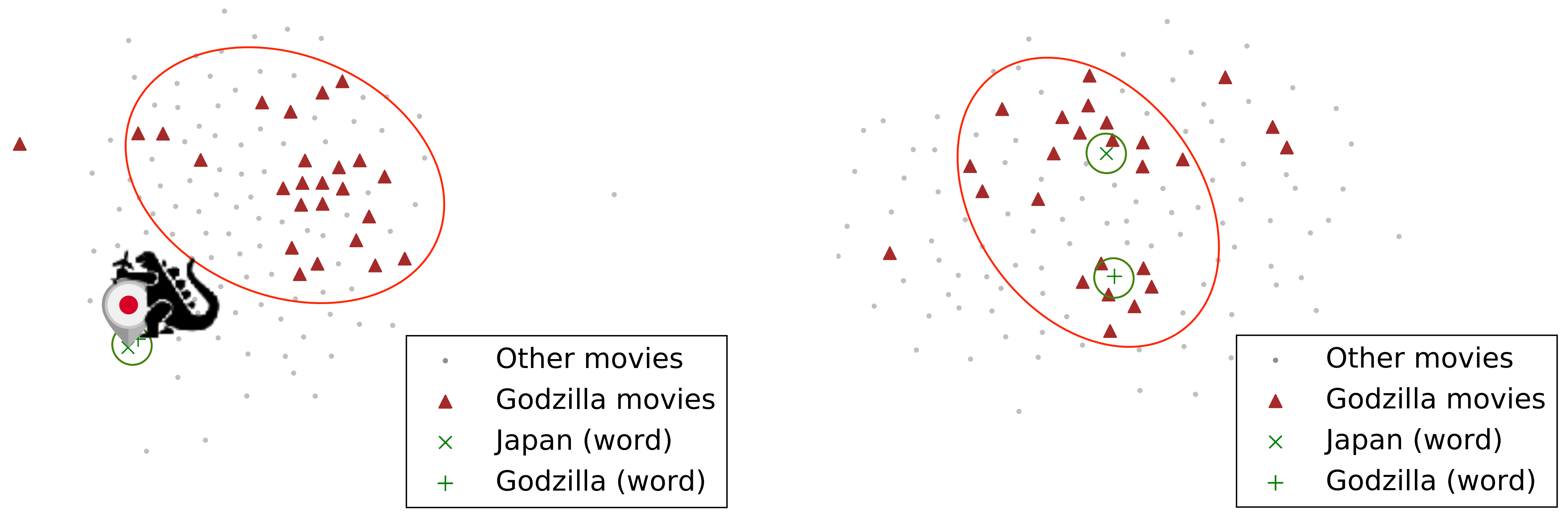
- Word-to-movie Retrieval (heterogeneous):

Table 5: Word-to-movie retrieval task

$\lvert W \rvert = 20$	Horror (3754/36586)	Thriller (4636/36586)	Western (751/36586)	Action (5029/36586)	Short (1094/36586)	Sci-Fi (2004/36586)	Average
		"Killer"		P@50		"Alien"	
RAND	0.080	0.080	0.060	0.080	0.000	0.120	0.070
KBR	0.324	**0.230**	**0.321**	0.418	0.062	0.373	0.288
AVGEMB	0.322	0.212	0.316	0.406	0.092	0.392	0.290
AVGEMB (all)	0.324	0.225	0.304	0.366	0.089	**0.401**	0.285
BPT	0.096	0.104	0.010	0.154	0.032	0.086	0.080
ICE (exp-5)	**0.354**	0.204	0.294	**0.444**	**0.142**	0.392	$\boxed{\mathbf{0.305}}$
				P@100			
RAND	0.050	0.100	0.030	0.110	0.000	0.060	0.058
KBR	**0.327**	**0.224**	0.236	0.395	0.057	0.307	0.258
AVGEMB	0.324	0.215	0.266	0.385	0.074	0.372	0.273
AVGEMB (all)	0.314	0.208	**0.269**	0.376	0.074	**0.382**	0.270
BPT	0.088	0.116	0.012	0.156	0.034	0.086	0.082
ICE (exp-5)	0.321	0.193	0.264	**0.421**	**0.109**	0.362	$\boxed{\mathbf{0.278}}$

Movies flock to concepts with high similarity

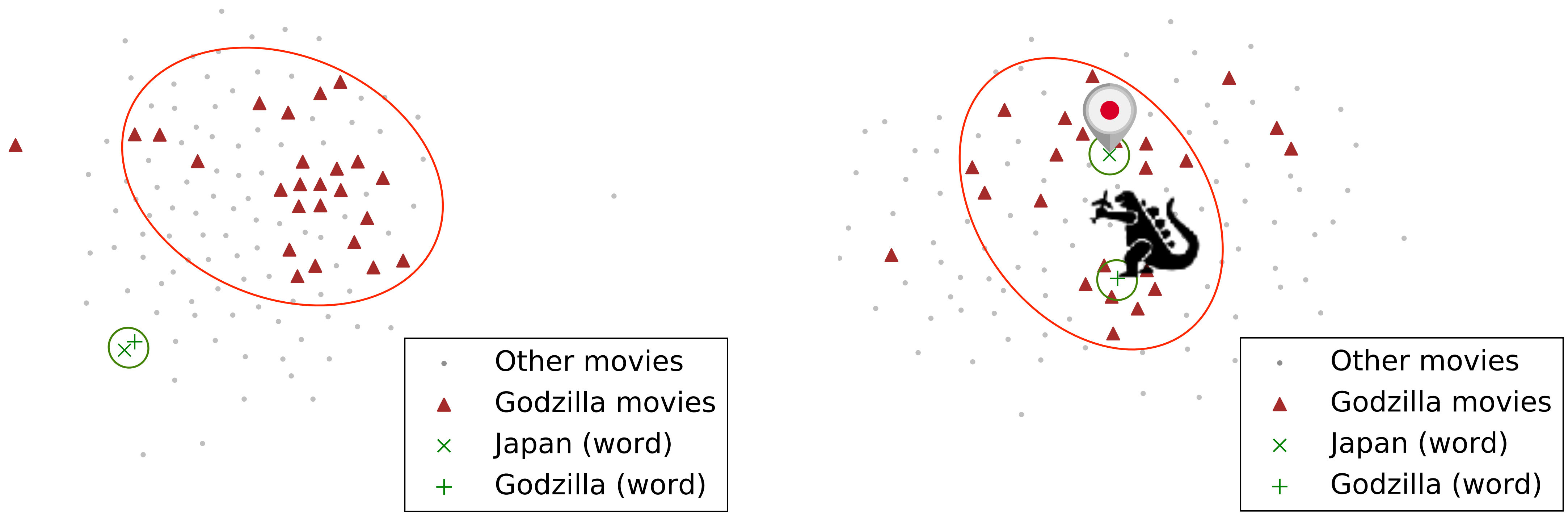
Figure 4: Visualization of the Representations of the Godzilla-related Movies and Two Related Keywords



(a) BPT with $|W| = 20$ **(b) ICE (exp-5) with** $|W| = 20$

Movies flock to concepts with high similarity

Figure 4: Visualization of the Representations of the Godzilla-related Movies and Two Related Keywords



(a) BPT with $|W| = 20$ **(b) ICE (exp-5) with** $|W| = 20$

- ICE concept embeddings can retrieve movies of similar concepts, and vice versa.
- Therefore, ICE embeddings are suitable for heterogeneous tasks.

Heterogeneous: Word-to-song retrieval

- Word-to-song Retrieval (heterogeneous):

Context types

Table 6: Performance comparison on the 15 keywords

$\|W\| = 10$		Keyword				Concept-similar word			
			P@100				P@100		
	Query	# keyword songs	BPT	AVGEMB	ICE (exp-3)	# concept-similar songs	BPT	AVGEMB	ICE (exp-3)
Mood	失落 (lost)	516	0.000	0.160	**0.470**	403	0.030	**0.120**	0.050
	心痛 (heartache)	824	0.050	0.080	**0.250**	4,075	0.170	0.500	**0.610**
	想念 (pining)	1,729	0.050	0.250	**0.700**	1,176	0.080	**0.180**	0.060
	深爱 (affectionate)	380	0.000	0.090	**0.550**	442	0.020	0.110	**0.250**
	難過 (sad)	1678	0.040	0.200	**0.530**	1,781	0.080	**0.320**	0.070
Location	回家 (home)	934	0.040	0.310	**0.900**	1,190	0.020	**0.340**	0.160
	房間 (room)	610	0.000	0.420	**0.510**	28	0.000	0.010	**0.060**
	海邊 (seaside)	264	0.000	0.230	**0.360**	91	0.000	0.070	**0.080**
	火車 (train)	151	0.010	0.330	**0.510**	20	0.000	**0.040**	0.020
	花園 (garden)	139	0.000	0.160	**0.390**	2	0.000	0.000	0.000
Time	夕陽 (dusk)	387	0.010	0.180	**0.360**	307	0.020	**0.100**	0.070
	日出 (sunrise)	240	0.000	0.290	**0.430**	390	0.060	0.380	**0.690**
	日落 (sunset)	226	0.030	0.380	**0.590**	407	0.010	0.270	**0.530**
	月亮 (moon)	598	0.000	0.360	**0.930**	1,608	0.030	0.320	**0.350**
	黑夜 (dark night)	1,189	0.030	0.140	**0.510**	279	**0.030**	**0.030**	0.010
	Total/Avg. P@100	9,865	0.017	0.239	**0.533**	12,199	0.037	0.186	**0.201**

Diverse and relevant by ConceptNet

Table 7: Performance evaluated by ConceptNet Human-labeled semantic knowledge graph

$\lvert W \rvert = 10$		P@10		Diversity@10		P@100		Diversity@100	
Query	# words in ConceptNet	KBR	ICE (exp-3)	KBR	ICE (exp-3)	KBR	ICE (exp-3)	KBR	ICE (exp-3)
夕陽 (dusk)	11	0.00	**0.20**	0.00	0.00	**0.25**	0.08	0.00	**0.75**
房間 (room)	39	**0.60**	0.10	0.00	0.00	**0.36**	0.16	0.00	**0.69**
日出 (sunrise)	17	0.40	**1.00**	0.00	**0.70**	**0.30**	0.24	0.00	**0.75**
花園 (garden)	33	**0.30**	0.10	0.00	0.00	**0.34**	0.08	0.00	**0.50**
黑夜 (dark night)	17	0.50	**1.00**	0.00	0.00	0.50	**0.57**	0.00	**0.68**
Average	23.4	0.36	**0.48**	0.00	**0.14**	**0.35**	0.23	0.00	**0.67**

Relevance Diversity

- Songs retrieved using ICE word embeddings have high diversity and relevance by human standard.

Case Study

Table 8: An example for movie-to-word retrieval

Query movie: Toy Story, 1995 (Animation, Adventure, Comedy)		
BPT	ICE (exp-5)	
manias	andy	**Protagonist**
entraineuse	gave	
taddeo	give	
anuelo	sid	**Antagonist**
portico	tabbed	
bep	robertson	
meanness	Named	
zanchi	stuffed_animals	
sarti	toys	**Generic toys**
raffin	Toys	

Table 10: An example for word-to-movie retrieval

Word query: alien **Representative concept for Sci-Fi**

BPT	ICE (exp-5)
The Blue Lagoon, 1949 (Adventure, Drama, Romance)	Coneheads, 1993 (Comedy, **Sci-Fi**)
Turner & Hooch, 1989 (Comedy, Crime, Drama)	Without Warning, 1980 (**Sci-Fi**, Horror)
Only the Young, 2012 (Documentary, Comedy, Romance)	They Came from Beyond Space, 1967 (Adventure, **Sci-Fi**)
Brute Force, 1947 (Crime, Drama, Film-Noir)	Battle of the Stars, 1978 (**Sci-Fi**)
Home, 2015 (Animation, Adventure, Comedy)	Howard the Duck, 1986 (Action, Adventure, Comedy)

Conclusions

1. Propose the ICE framework, which models item concepts using textual information.
2. Propose a generalized network construction method based on matrix operations.
3. Leverage neighborhood proximity to learn embeddings capable to be used in both homogeneous and heterogeneous tasks.
4. Resulted embeddings can be used to retrieve conceptually diverse an relevant items.

Release: ICE API and dataset

github.com/
cnclabs/ICE

- ICE API:
    - Repo: https://github.com/cnclabs/ICE
    - Demo: https://cnclabs.github.io/ICE/
- IMDB dataset:
    - MovieLens 10/2016 Full dataset.
    - 36,586 movies with plot descriptions and genres.
- Special thanks to Chen Chih-Ming for his help to the development of the API.

Appendix

github.com/
cnclabs/ICE

Diversity Measure

- An item is considered diverse if and only if it contains at least one expanded concept words:

$$\text{Diversity@}n = \frac{|R \cap \overline{S}_k|}{|R|},$$

where R denotes the set of relevant songs and S_k denotes the set of songs containing the given keyword k, from the n retrieved songs.