Empowering Citizens. Smarter Societies.

Insight
Centre for Data Analytics

Transfer Learning for Item Recommendations and Knowledge Graph Completion in Item Related Domains via a Co-Factorization Model

Guangyuan Piao, John G. Breslin

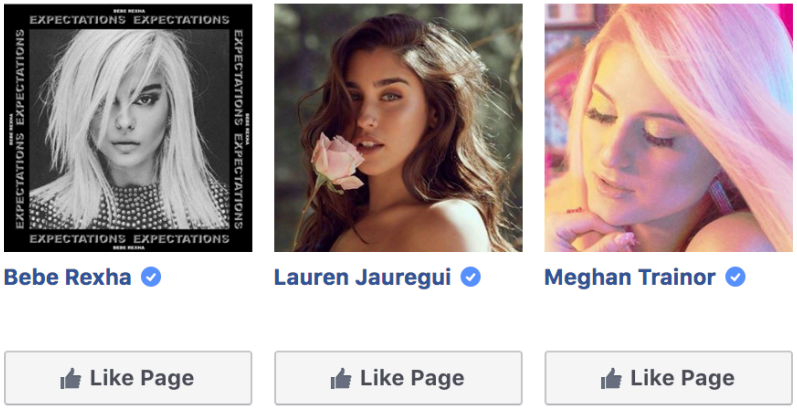
Insight Centre for Data Analytics @NUI Galway

A World Leading SFI Research Centre

03/06/2018 – 07/06/2018 ESWC2018

Contents

Recommender Systems (RecSys)



amazon

NETFLIX

RecSys Approaches

collaborative filtering approaches

- recommend items to a user by
- identifying other users with a similar profile

✗ new item

✗ new user

content-based recommender systems

- recommend items to a user based on
- their **description** and on the profile of the user's interests

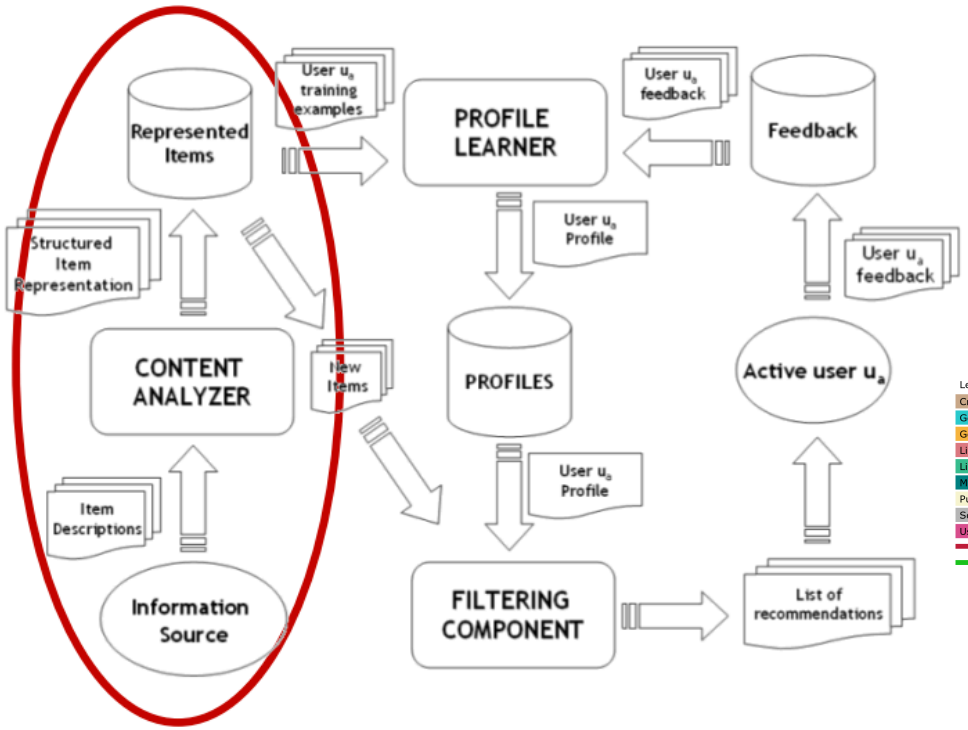
✓ new item

✗ new user

hybrid approaches

- combine several RecSys approaches
- overcome the disadvantages of a RecSys approach

Linked Open Data (LOD)



content-based recommender systems

LOD dataset(s) provides **domain knowledge** and **rich Information** about items

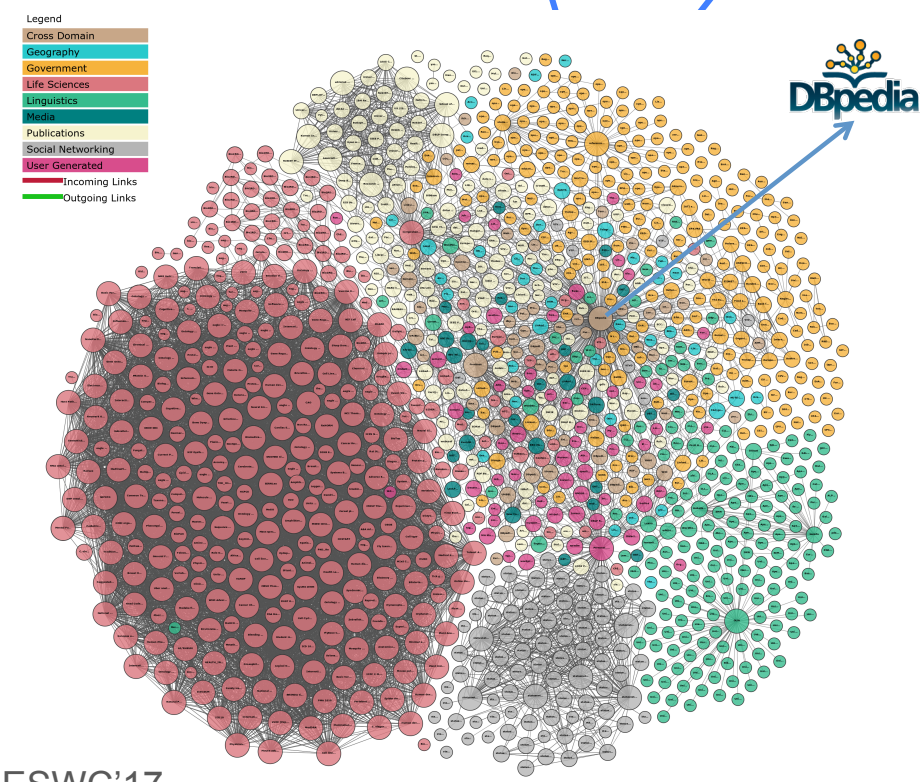
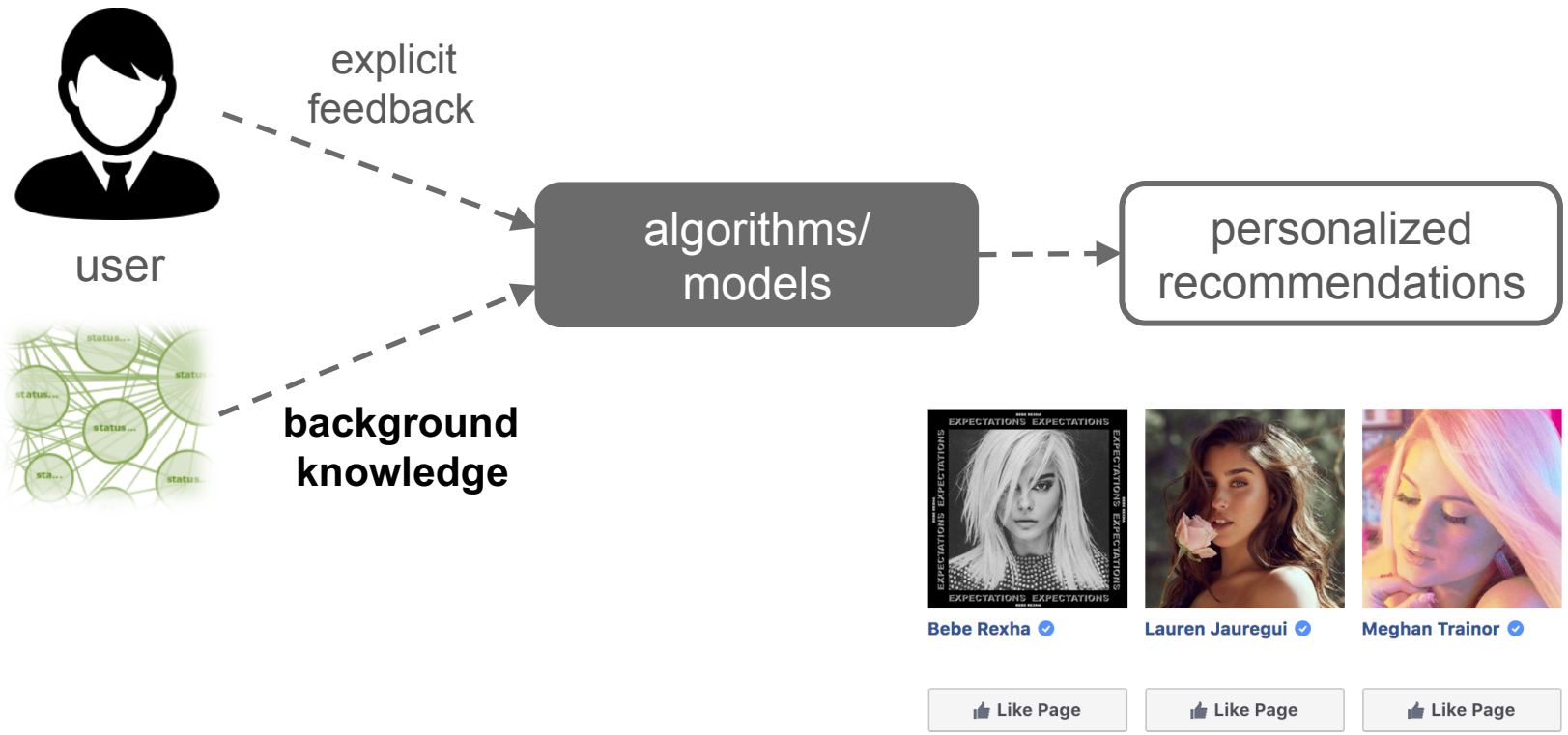


figure: Tutorial on LOD for Semantics-Aware RecSys, ESWC'17

Semantics-Aware RecSys



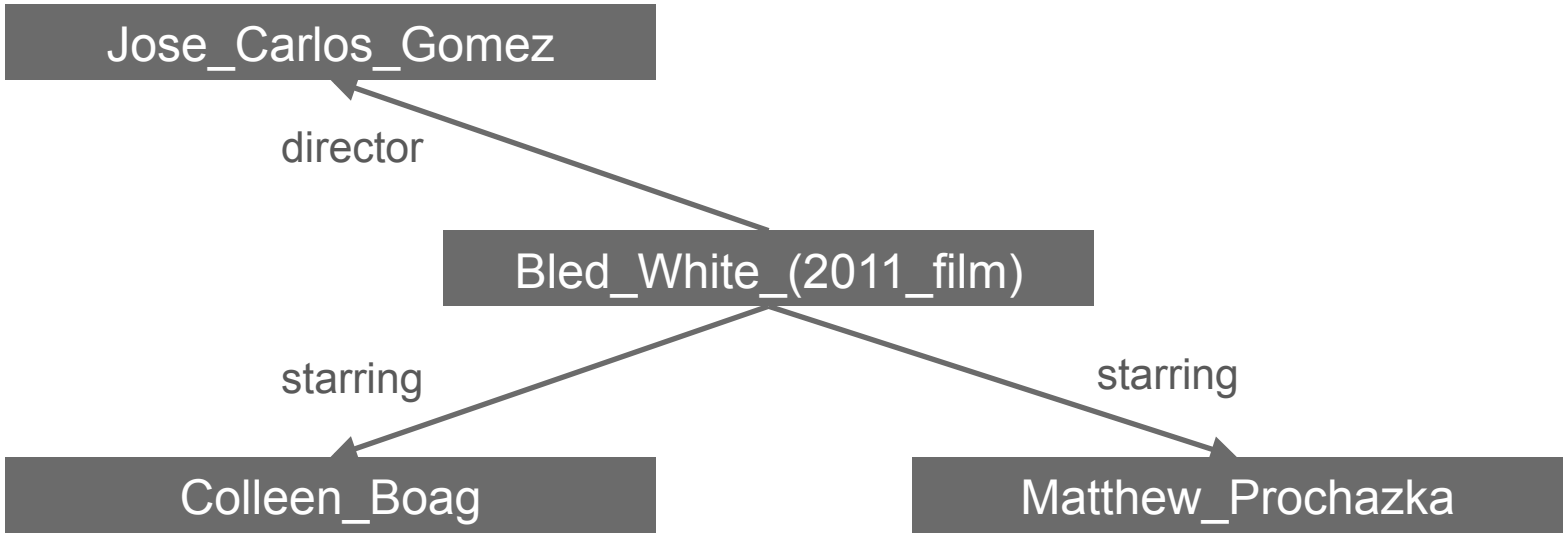
(some) related work

- semantic similarity/distance measures [ISWC'10, AAAI'10, SAC'16]
- graph-based algorithms such as PageRank [UMAP'16, WWW'15]
- machine learning approaches [RecSys'12, TIST'16, WISE'17]

Semantics-Aware RecSys

main focus

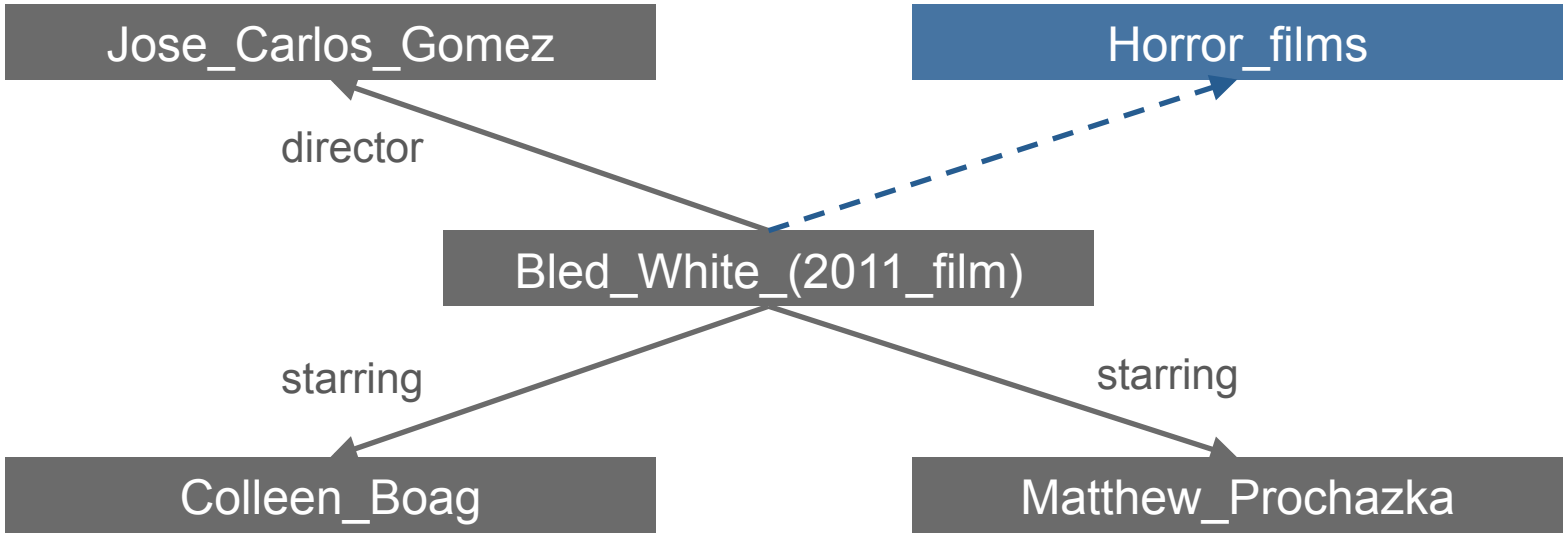
- how to consume **existing** knowledge about items/entities for RecSys?



Semantics-Aware RecSys

main focus

- how to consume **existing** knowledge about items/entities for RecSys?

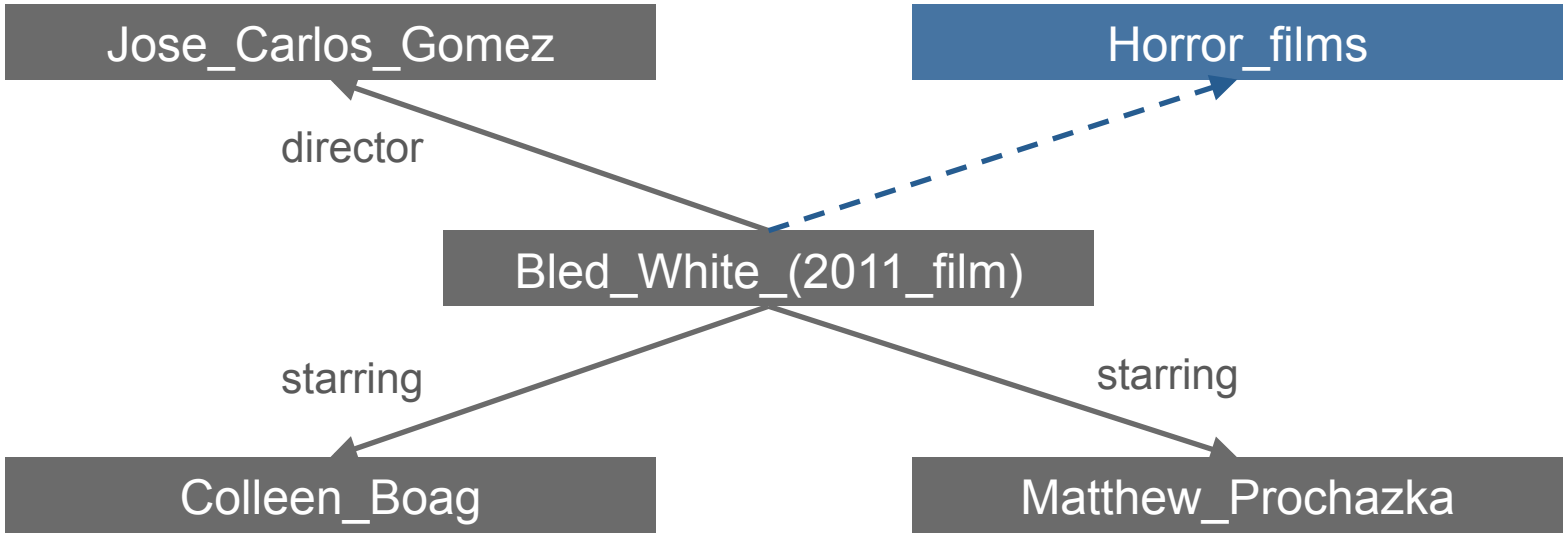


- previous work does not incorporate the **incompleteness** of KGs
  - a dedicated line of research: **KG completion**

Semantics-Aware RecSys

main focus

- how to consume **existing** knowledge about items/entities for RecSys?



- previous work does not incorporate the **incompleteness** of KGs
  - a dedicated line of research: **KG completion**
- transferring knowledge KG ➔ item rec.
  - how about transferring knowledge (user-item interactions) item rec. ➔ KG?

Item rec. and KG completion tasks

item rec.

- given user-item interactions (e.g., likes and dislikes)
  + background knowledge from knowledge graph (KG)
- provide the top-N item recommendations

KG completion

- given (*subject*, *predicate*)
- provide the top-N *object* recommendations [SAC'12, AAAI'16]
- candidate *objects*
  - all *objects* in the range of a given *predicate* in the DBpedia ontology

- (some) related work
  - factorization approaches [NIPS'13, SAC'12, WWW'12]
  - neural network-based models [ACL'15, AAAI'14, AAAI'16]
  - …

Item rec. and KG completion approaches

item rec.

- factorization machines (FMs) [TIST'12]

$$s_{d_{ui}} = \beta_0 + \beta_u + \beta_i + < \theta_u, \theta_i >$$

$$\ell(a_1, a_2) = \sum_{a_1 \in \mathcal{D}_{ui}^+} \sum_{a_2 \in \mathcal{D}_{ui}^-} -\log[\delta(s_{a_1} - s_{a_2})]$$

β_0 : bias
β_i : weight of item features
θ_i : latent factors
$\mathcal{D}_{ui}^+$: set of positive (u, i)

KG completion

- TransE to satisfy $\varphi_s + \varphi_p \approx \varphi_o$ for a valid triple (*s, p, o*) [NIPS'13]

$$s'_{d_{spo}} = \sqrt{\sum_{j=1}^{m} (\varphi_{s_j} + \varphi_{p_j} - \varphi_{o_j})^2}$$

$$\ell(b_1, b_2) = \sum_{b_1 \in \mathcal{D}_{spo}^+} \sum_{b_2 \in \mathcal{D}_{spo}^-} [\gamma + s'_{b_1} - s'_{b_2}]_+$$

φ_s : subject embedding
φ_p : predicate embedding
φ_o : object embedding

γ : margin (1.0)

Item rec. and KG completion approaches

item rec.

- factorization machines (FMs) [TIST'12]

$$s_{d_{ui}} = \beta_0 + \beta_u + \beta_i + < \theta_u, \theta_i >$$

$$\ell(a_1, a_2) = \sum_{a_1 \in \mathcal{D}^+} \sum_{a_2 \in \mathcal{D}^-} -\log[\delta(s_{a_1} - s_{a_2})]$$

two representations for the same item
in two different tasks

KG completion

- TransE to satisfy $\varphi_s + \varphi_p \approx \varphi_o$ for a valid triple (*s, p, o*) [NIPS'13]

$$s'_{d_{spo}} = \sqrt{\sum_{j=1}^{m} (\varphi_{s_j} + \varphi_{p_j} - \varphi_{o_j})^2}$$

$$\ell(b_1, b_2) = \sum_{b_1 \in \mathcal{D}^+_{spo}} \sum_{b_2 \in \mathcal{D}^-_{spo}} [\gamma + s'_{b_1} - s'_{b_2}]_+$$

Transfer Learning via Co-Factorization Model

transfer learning [TKDE'10]

- using one task as a **source** task and the other as a **target** task
- optimize the objective function of **target** task by transferring knowledge from **source** task

modeling the relationship between the two representations of items

- via shared latent space ($CoFM_A$)

$$\theta_i = \varphi_s = \rho_{is}$$

- via regularization ($CoFM_R$)

$$\lambda_{\varphi,\theta} \|\varphi_s - \theta_i\|_F^2$$

Transfer Learning via Co-Factorization Model

CoFM: putting everything together

$$Opt(CoFM) : Opt(T) + \varepsilon \times Opt(S)$$

ε transfer value: controls the importance of knowledge transfer from source task (S) to target task (T)

$$Opt(T) = \arg\min \sum_{d_T \in \mathcal{D}_T} \ell_T(\cdot)$$

training instances in T

$$Opt(S) = \arg\min \sum_{d'_S \in \mathcal{D}'_S} \ell_S(\cdot)$$

randomly sampled training instances in S

$$d_T(i) = d'_S(s)$$

$$|\mathcal{D}_T| = |\mathcal{D}'_S|$$

for SGD

Experiment - Datasets

Table 1: Statistics of Movielens and DBbook datasets.

		Movielens	**DBbook**
statistics of user-item interactions	# of users	3,997	6,181
	# of items	3,082	6,733
	# of ratings	827,042	72,372
	avg. # of ratings	206	12
	sparsity	93.27%	99.38%
	% of positive ratings	56%	45.85%
statistics of RDF triples	# of subjects	2,952 (3,082)	6,211 (6,733)
	# of predicates	21	36
	# of objects	18,550	16,476
	# of triples	81,835	72,911

Experiment - Training & Test Sets

item rec.

- 80% (20% for validation) for training, 20% for test

KG completion

- test set: given *s*, randomly choose (*s, p*) pair for *s*
- training set: other triples containing *s*

Experiment - Training & Test Sets

item rec.

- 80% (20% for validation) for training, 20% for test

KG completion

- test set: given *s*, randomly choose (*s, p*) pair for *s*
- training set: other triples containing *s*

repeat 5 times sampling new training & test sets, average results

Experiment - Evaluation Metrics

nDCG@N (normalized Discounted Cumulative Gain)

- takes into account the relevant items as well as their rank positions

MRR (mean reciprocal rank)

- the 1st relevant item occurs on average in recommendations

P@N (precision)

- mean probability of a retrieved item in the top-N list are relevant

R@N (recall)

- mean probability of relevant items retrieved in the top-N list

Experiment - Compared Methods

item rec.

- kNN (k-nearist neighbors algorithm)
- BPRMF [UAI'09]
  - CoFM without transferring knowledge from the KG completion task
- FM_{LOD} [WISE'17]
  - exploits KG-enabled features (existing knowledge) with FMs

KG completion

- MFPP (Most Frequent Per Predicate)
- PITF [SAC'12]
  - capturing interactions between *subject*, *predicate*, and *object* using FMs
- TransE [NIPS'13]
  - CoFM without transferring knowledge from item rec.

Results: Movielens

$CoFM_R$ > $CoFM_A$

	S: KG completion *T*: item recommendations					*S*: item recommendations *T*: KG completion				
	kNN	BPRMF	FM_{LOD}	$CoFM_A$	$CoFM_R$	MFPP	PITF	TransE	$CoFM_A$	$CoFM_R$
MRR	0.510	0.594	0.609	0.602	**0.622**	0.183	0.266	0.317	0.302	**0.322**
nDCG@5	0.358	0.425	0.436	0.429	**0.445**	0.149	0.248	0.292	0.279	**0.297**
P@5	0.291	0.355	0.366	0.360	**0.372**	0.070	0.096	0.123	**0.126**	**0.126**
R@5	0.075	0.097	0.100	0.098	**0.102**	0.103	0.230	0.241	0.240	**0.249**
nDCG@10	0.440	0.500	0.510	0.504	**0.518**	0.171	0.273	0.311	0.299	**0.316**
P@10	0.258	0.307	0.314	0.310	**0.318**	0.046	0.064	0.077	**0.081**	0.079
R@10	0.129	0.161	0.165	0.164	**0.170**	0.149	0.271	0.277	0.280	**0.283**
nDCG@20	0.583	0.645	0.653	0.648	**0.660**	0.194	0.297	0.331	0.321	**0.336**
P@20	0.218	0.252	0.257	0.254	**0.259**	0.031	0.042	0.047	**0.051**	0.048
R@20	0.213	0.257	0.261	0.260	**0.265**	0.199	0.311	0.313	0.318	**0.319**

Results: Movielens

incorporating incompleteness of KG > leveraging **existing** knowledge of KG

	S: KG completion T: item recommendations					S: item recommendations T: KG completion				
	kNN	BPRMF	FM_{LOD}	$CoFM_A$	$CoFM_R$	MFPP	PITF	TransE	$CoFM_A$	$CoFM_R$
MRR	0.510	0.594	0.609	0.602	**0.622**	0.183	0.266	0.317	0.302	**0.322**
nDCG@5	0.358	0.425	0.436	0.429	**0.445**	0.149	0.248	0.292	0.279	**0.297**
P@5	0.291	0.355	0.366	0.360	**0.372**	0.070	0.096	0.123	**0.126**	**0.126**
R@5	0.075	0.097	0.100	0.098	**0.102**	0.103	0.230	0.241	0.240	**0.249**
nDCG@10	0.440	0.500	0.510	0.504	**0.518**	0.171	0.273	0.311	0.299	**0.316**
P@10	0.258	0.307	0.314	0.310	**0.318**	0.046	0.064	0.077	**0.081**	0.079
R@10	0.129	0.161	0.165	0.164	**0.170**	0.149	0.271	0.277	0.280	**0.283**
nDCG@20	0.583	0.645	0.653	0.648	**0.660**	0.194	0.297	0.331	0.321	**0.336**
P@20	0.218	0.252	0.257	0.254	**0.259**	0.031	0.042	0.047	**0.051**	0.048
R@20	0.213	0.257	0.261	0.260	**0.265**	0.199	0.311	0.313	0.318	**0.319**

Results: Movielens

with knowledge transfer > **without** knowledge transfer

	S: KG completion T: item recommendations					S: item recommendations T: KG completion				
	kNN	BPRMF	FM_{LOD}	$CoFM_A$	$CoFM_R$	MFPP	PITF	TransE	$CoFM_A$	$CoFM_R$
MRR	0.510	0.594	0.609	0.602	**0.622**	0.183	0.266	0.317	0.302	**0.322**
nDCG@5	0.358	0.425	0.436	0.429	**0.445**	0.149	0.248	0.292	0.279	**0.297**
P@5	0.291	0.355	0.366	0.360	**0.372**	0.070	0.096	0.123	**0.126**	**0.126**
R@5	0.075	0.097	0.100	0.098	**0.102**	0.103	0.230	0.241	0.240	**0.249**
nDCG@10	0.440	0.500	0.510	0.504	**0.518**	0.171	0.273	0.311	0.299	**0.316**
P@10	0.258	0.307	0.314	0.310	**0.318**	0.046	0.064	0.077	**0.081**	0.079
R@10	0.129	0.161	0.165	0.164	**0.170**	0.149	0.271	0.277	0.280	**0.283**
nDCG@20	0.583	0.645	0.653	0.648	**0.660**	0.194	0.297	0.331	0.321	**0.336**
P@20	0.218	0.252	0.257	0.254	**0.259**	0.031	0.042	0.047	**0.051**	0.048
R@20	0.213	0.257	0.261	0.260	**0.265**	0.199	0.311	0.313	0.318	**0.319**

Summary

investigate knowledge transfer between the two tasks

- via a proposed co-factorization model, showed that
- incorporating the incompleteness of KG ↑ rec. performance
- knowledge from item rec. can be transferred to the KG completion task and ↑ the performance, which has not been explored yet

future work:

- extraction of background knowledge for items
  - ✗ only considered item/subject ➔ predicate ➔ object
- modeling the relationship between the two item representations
  - ✗ have the same dimensionality
- other state-of-the-art approaches for both tasks
- evaluation on datasets in other domains

funded by SAP sfi Science Foundation Ireland Insight

thank you for your attention!

Guangyuan Piao

homepage: http://parklize.github.io
e-mail: guangyuan.piao@insight-centre.org
scholar: https://goo.gl/tgK9bk
slideshare: http://www.slideshare.net/parklize