Empirical Comparison of Graph Embeddings for Trust-Based Collaborative Filtering

Tomislav Duricic, Hussain Hussain, Emanuel Lacic, Dominik Kowald, Denis Helic, and Elisabeth Lex

SOCIAL COMPUTING LAB

TU Graz

ISDS INSTITUTE OF INTERACTIVE SYSTEMS AND DATA SCIENCE

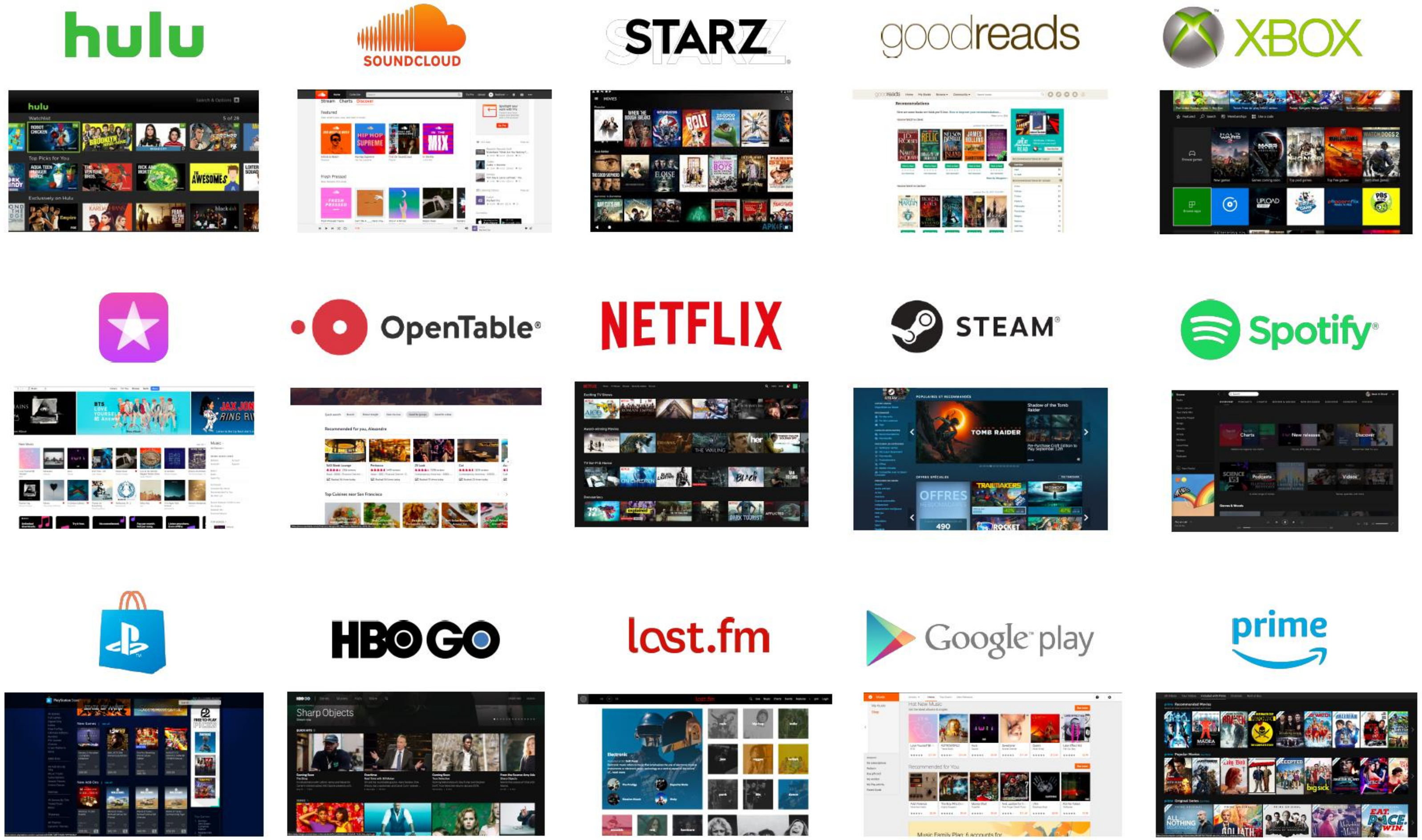
KNOW Center

Content

Motivation & Background

*„**Recommender systems** are crucial tools to overcome the **information overload** brought about by the Internet” [1]*

[1] Blattner, M., Hunziker, A., & Laureti, P. (2007). When are recommender systems useful?. *arXiv preprint arXiv:0709.2562*.



Recommender systems examples [2]

[2] https://www.slideshare.net/CrossingMinds/recommendation-system-explained/4

User-based collaborative filtering

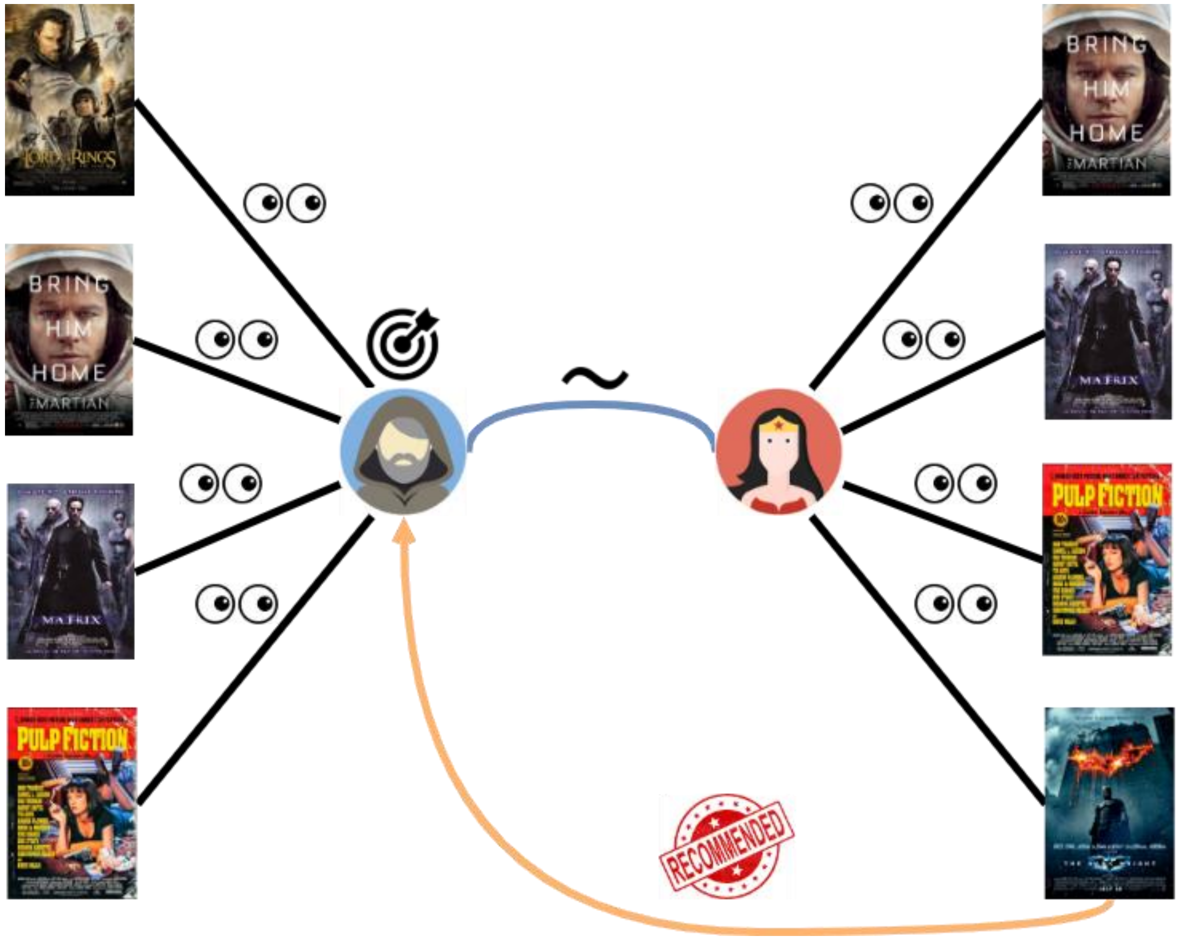
How does it generate recommendations for the target user?

1. Find ***k*** most similar users, i.e., ***k-nearest neighbors***

   ‼ *User similarity can be calculated in many different ways*

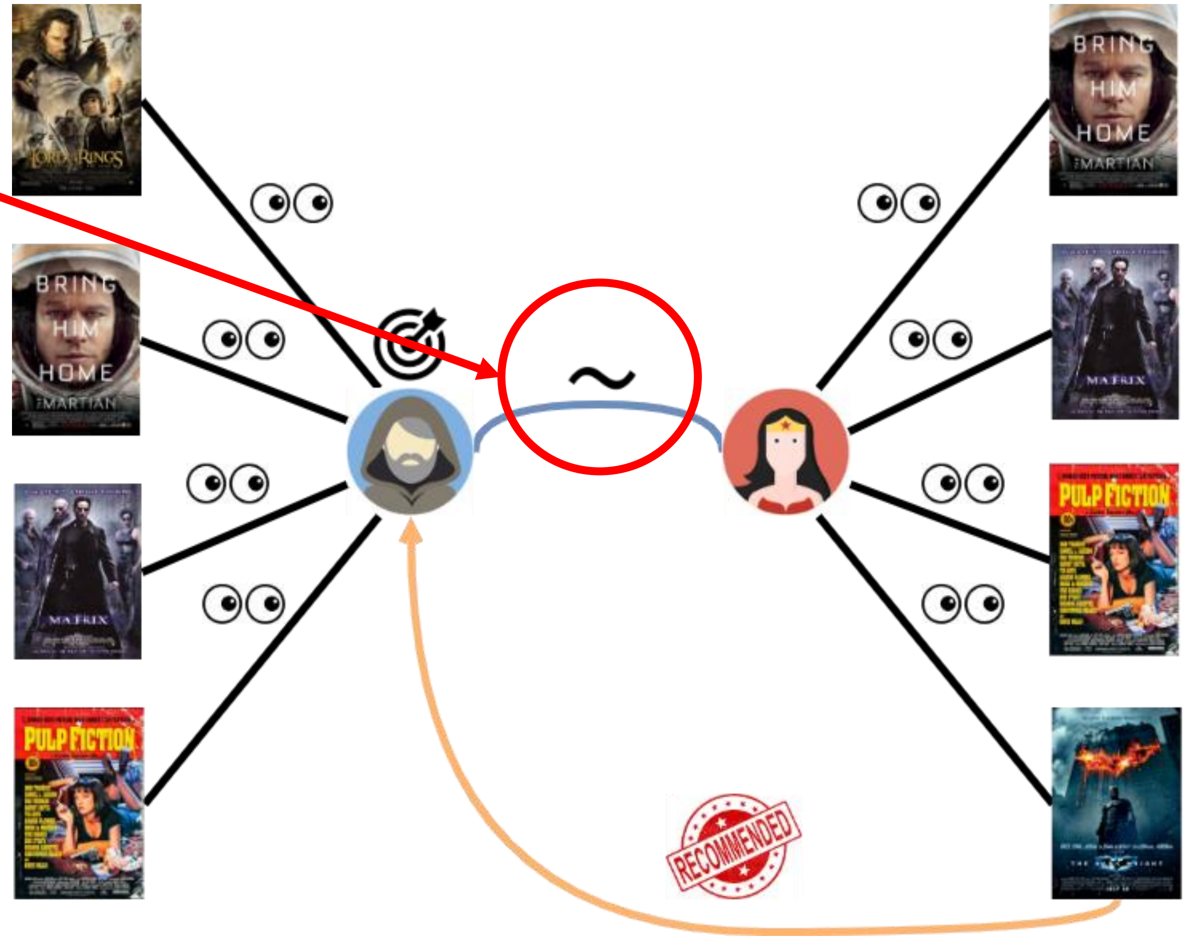
2. Look at the items they like and combine them to create a **ranked list** of recommendations

User-based collaborative filtering



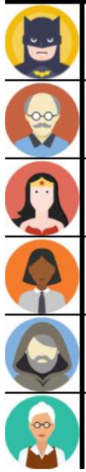
User-based collaborative filtering

How is user similarity calculated?



User-based collaborative filtering (from item ratings)

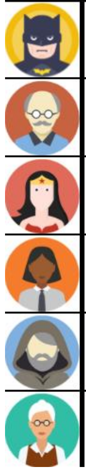
Ratings Matrix ***R***



		5	4		4	
	4	3		1		3
	5		2	2		4
	4	1		4		3
		3	4	4		5
			1		5	5

User-based collaborative filtering (from item ratings)

Ratings Matrix ***R***



	5	4		4	
4	3		1		3
5		2	2		4
4	1		4		3
	3	4	4		5
		1		5	5

Pairwise similarity calculator

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

User-based collaborative filtering (from item ratings)

Ratings Matrix ***R***

	5	4		4	
4	3		1		3
5		2	2		4
4	1		4		3
	3	4	4		5
		1		5	5

Pairwise similarity calculator

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Similarity Matrix ***S***

1			-0.19		1
	1		-0.5		0.5
		1	1	-1	0.19
-0.19	-0.5	1	1		0.41
		-1		1	
1	0.5	0.19	0.41		1

User-based collaborative filtering (from item ratings)

Used for finding *k*-nearest neighbors

Ratings Matrix *R*

	5	4		4	
4	3		1		3
5		2	2		4
4	1		4		3
	3	4	4		5
		1		5	5

Pairwise similarity calculator

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Similarity Matrix *S*

1			-0.19		1
	1		-0.5		0.5
		1	1	-1	0.19
-0.19	-0.5	1	1		0.41
		-1		1	
1	0.5	0.19	0.41		1

User-based collaborative filtering (from item ratings)

CONS:

- **Data sparsity –** ratings matrix is usually very sparse
  - *Often tackled by creating latent user representations from user rating vectors*

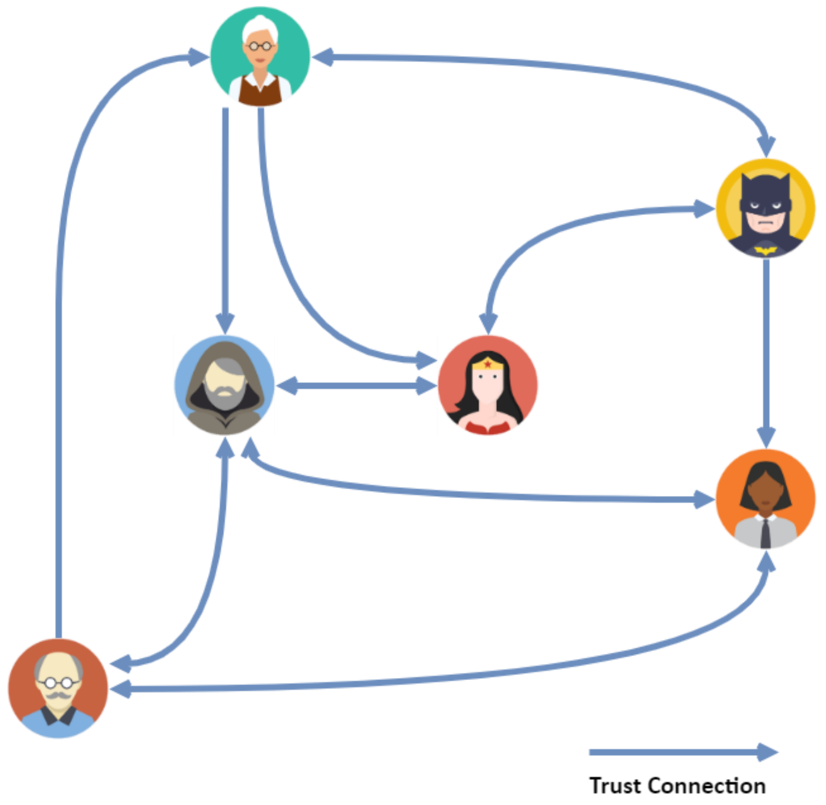
- **Cold-start user problem -** No or poor recommendations for users with none or very little item interactions
  - *One proposed solution is to use* ***additional user data*** *such as* ***social connections*** *between users (if they exist), e.g.,* ***trust connections***

Trust-based collaborative filtering [3]

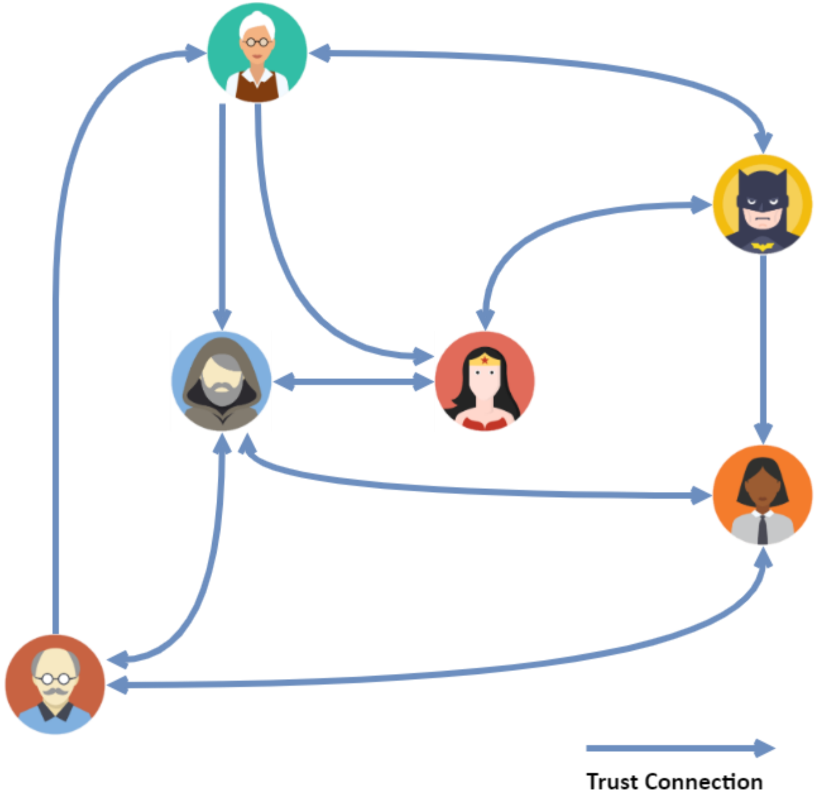
- **User-based Collaborative Filtering using user trust network as input**
  - *In evaluated datasets, trust network* $G = (V, E)$ *is an undirected unweighted graph with no self-loops consisting of a set of nodes* V *(users) and a set of edges* E *(trust connections).*

[3] Lathia, N., Hailes, S., & Capra, L. (2008, June). Trust-based collaborative filtering. In IFIP international conference on trust management (pp. 119-134). Springer, Boston, MA.

Trust network example

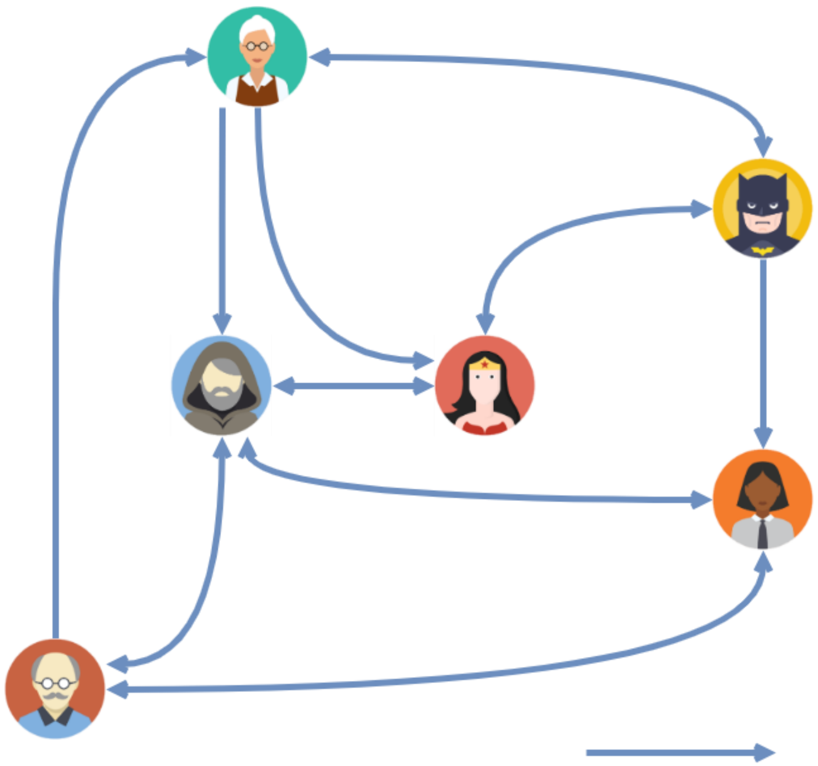


Trust network example



Can be represented as an adjacency matrix ***A***

Trust network example



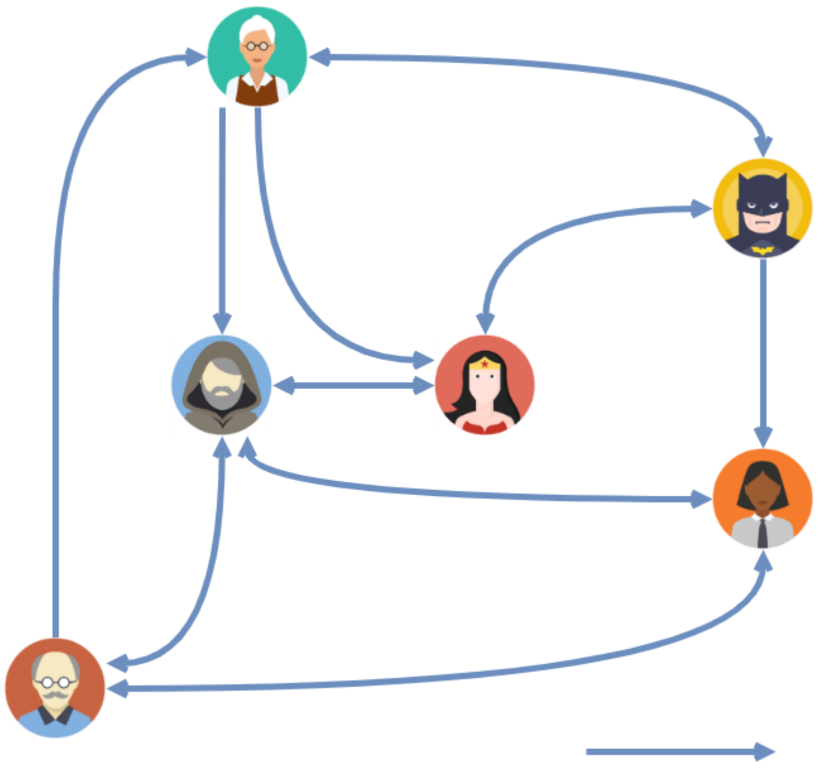
Can be represented as an adjacency matrix ***A***

Adjacency matrix ***A***

	0	0	1	1	0	1
	0	0	0	1	1	1
	1	0	0	0	1	0
	0	1	0	0	1	0
	0	1	1	0	0	0
	1	0	1	0	1	0

Trust network example

Can be used directly for finding *k*-nearest neighbors. However...



Can be represented as an adjacency matrix ***A***

Adjacency matrix ***A***

	0	0	1	1	0	1
	0	0	0	1	1	1
	1	0	0	0	1	0
	0	1	0	0	1	0
	0	1	1	0	0	0
	1	0	1	0	1	0

Trust network adjacency matrix

- Tipically very **sparse**, i.e., network density is low
  - *Total number of edges is much lower than the total number of possible edges*
- **Cold-start user problem -** No or poor recommendations for users with none or very little **trust connections**
- No measure of how much one user trusts another
- **Solution:**
  a) *By directly employing a similarity metric of choice to* ***A***
  b) *Converting user vectors from* ***A*** *into latent feature space and then employing a similarity metric of choice to the embeddings matrix* $\boldsymbol{Z} \in \mathbb{R}^{|V| \times d}$

a) Direct similarity metric approach

Adjacency matrix *A*

	0	0	1	1	0	1
	0	0	0	1	1	1
	1	0	0	0	1	0
	0	1	0	0	1	0
	0	1	1	0	0	0
	1	0	1	0	1	0

a) Direct similarity metric approach

Adjacency matrix *A*

0	0	1	1	0	1
0	0	0	1	1	1
1	0	0	0	1	0
0	1	0	0	1	0
0	1	1	0	0	0
1	0	1	0	1	0

Pairwise similarity calculator

$$J(\mathbf{A}_{a,*},\mathbf{A}_{b,*}) = \frac{|\mathbf{A}_{a,*} \cap \mathbf{A}_{b,*}|}{|\mathbf{A}_{a,*} \cup \mathbf{A}_{b,*}|}$$

a) Direct similarity metric approach

Adjacency matrix *A*

	0	0	1	1	0	1
	0	0	0	1	1	1
	1	0	0	0	1	0
	0	1	0	0	1	0
	0	1	1	0	0	0
	1	0	1	0	1	0

Pairwise similarity calculator

$$J(\mathbf{A}_{a,*},\mathbf{A}_{b,*}) = \frac{|\mathbf{A}_{a,*} \cap \mathbf{A}_{b,*}|}{|\mathbf{A}_{a,*} \cup \mathbf{A}_{b,*}|}$$

Similarity Matrix *S*

	1	0	0.25	0	0.5	0
	0	1	0.25	0	0.2	0
	0.25	0.25	1	0.25	0.17	0.25
	0	0	0.25	1	0.2	1
	0.5	0.2	0.17	0.2	1	0.2
	0	0	0.25	1	0.2	1

b) Latent user representations approach

Adjacency matrix ***A***

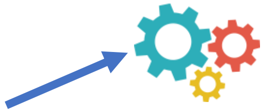
	0	0	1	1	0	1
	0	0	0	1	1	1
	1	0	0	0	1	0
	0	1	0	0	1	0
	0	1	1	0	0	0
	1	0	1	0	1	0

b) Latent user representations approach

Adjacency matrix ***A***

	0	0	1	1	0	1
	0	0	0	1	1	1
	1	0	0	0	1	0
	0	1	0	0	1	0
	0	1	1	0	0	0
	1	0	1	0	1	0

Graph embeddings approach



$\boldsymbol{d} = \mathbf{5}$

b) Latent user representations approach

Adjacency matrix ***A***

	0	0	1	1	0	1
	0	0	0	1	1	1
	1	0	0	0	1	0
	0	1	0	0	1	0
	0	1	1	0	0	0
	1	0	1	0	1	0

Graph embeddings approach

$\boldsymbol{d} = \mathbf{5}$

Embeddings matrix $\boldsymbol{Z} \in \mathbb{R}^{|V| \times d}$

	φ_1	φ_2	φ_3	φ_4	φ_5
	0.003	0.978	0.174	0.061	-0.062
	0.06	0.923	0.224	0.083	-0.029
	0.24	0.854	0.29	0.069	-0.118
	0.368	0.808	0.344	0.208	-0.031
	0.066	0.904	0.175	0.325	-0.15
	0.443	0.763	0.391	0.205	-0.214

b) Latent user representations approach

Adjacency matrix *A*

	0	0	1	1	0	1
	0	0	0	1	1	1
	1	0	0	0	1	0
	0	1	0	0	1	0
	0	1	1	0	0	0
	1	0	1	0	1	0

Graph embeddings approach

$d = 5$

Embeddings matrix $Z \in \mathbb{R}^{|V| \times d}$

	φ_1	φ_2	φ_3	φ_4	φ_5
	0.003	0.978	0.174	0.061	-0.062
	0.06	0.923	0.224	0.083	-0.029
	0.24	0.854	0.29	0.069	-0.118
	0.368	0.808	0.344	0.208	-0.031
	0.066	0.904	0.175	0.325	-0.15
	0.443	0.763	0.391	0.205	-0.214

Pairwise similarity calculator

$$cos(\theta) = \frac{\mathrm{A}_{a,*} \cdot \mathrm{A}_{b,*}}{\|\mathrm{A}_{a,*}\| \, \|\mathrm{A}_{b,*}\|}$$

b) Latent user representations approach

Adjacency matrix ***A***

0	0	1	1	0	1
0	0	0	1	1	1
1	0	0	0	1	0
0	1	0	0	1	0
0	1	1	0	0	0
1	0	1	0	1	0

Graph embeddings approach

$d = 5$

Embeddings matrix $\boldsymbol{Z} \in \mathbb{R}^{|V| \times d}$

φ_1	φ_2	φ_3	φ_4	φ_5
0.003	0.978	0.174	0.061	-0.062
0.06	0.923	0.224	0.083	-0.029
0.24	0.854	0.29	0.069	-0.118
0.368	0.808	0.344	0.208	-0.031
0.066	0.904	0.175	0.325	-0.15
0.443	0.763	0.391	0.205	-0.214

Pairwise similarity calculator

$$cos(\theta) = \frac{\mathrm{A}_{a,*} \cdot \mathrm{A}_{b,*}}{\|\mathrm{A}_{a,*}\| \, \|\mathrm{A}_{b,*}\|}$$

Similarity matrix ***S***

1	0.99	0.95	0.89	0.96	0.84
0.99	1	0.97	0.92	0.96	0.87
0.95	0.97	1	0.97	0.94	0.95
0.89	0.93	0.97	1	0.92	0.98
0.96	0.96	0.94	0.92	1	0.89
0.84	0.87	0.96	0.98	0.89	1

Formulating the problem

We explore the **utility** of **graph embeddings** for finding ***k*-nearest neighbors** in **trust-based** **collaborative filtering** for **cold-start** users.

Experimental Setup

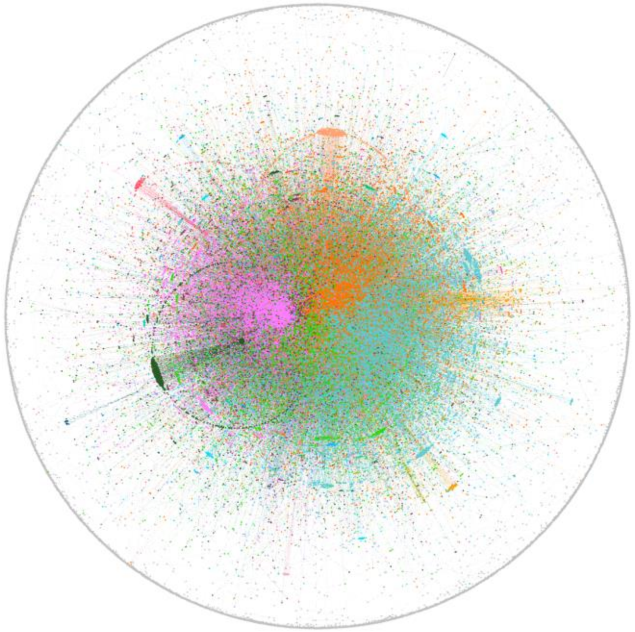
Datasets

Dataset	#Users	#Items	#Edges	#Ratings	Graph density
Epinions [4]	49,288	139,738	487,183	664,824	2×10^{-4}
Ciao [5]	19,533	16,121	40,133	72,665	1.85×10^{-3}
Filmtrust [6]	1,642	2,071	1,853	35,497	2.43×10^{-3}

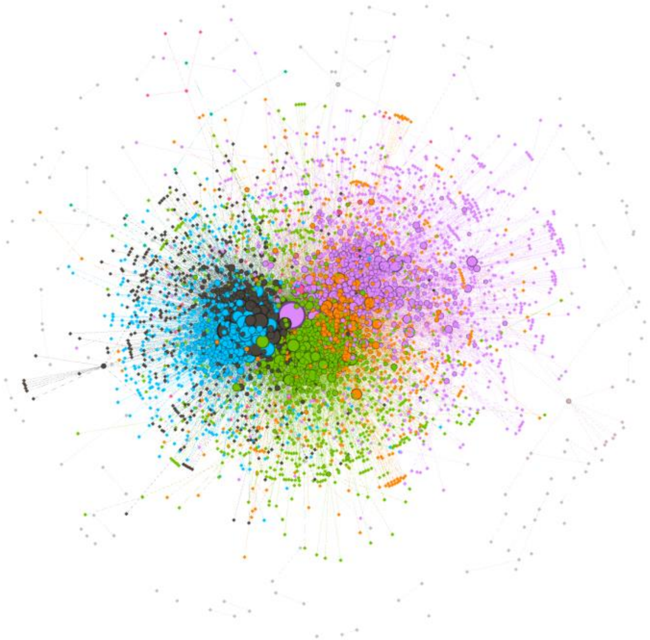
* This table shows the number of direct edges, however, in our experiments we convert each directed network to an undirected network by removing edge direction

[4] P. Massa and P. Avesani, "Trust-aware collaborative filtering for recommender systems," in OTM Confederated International Conferences" On the Move to Meaningful Internet Systems", pp. 492–508, Springer, 2004.
[5] J. Tang, H. Gao, and H. Liu, "mTrust: Discerning multi-faceted trust in a connected world," in Proceedings of the fifth ACM international conference on Web search and data mining, pp. 93–102, ACM, 2012.
[6] G. Guo, J. Zhang, and N. Yorke-Smith, "A novel bayesian similarity measure for recommender systems," 2013.

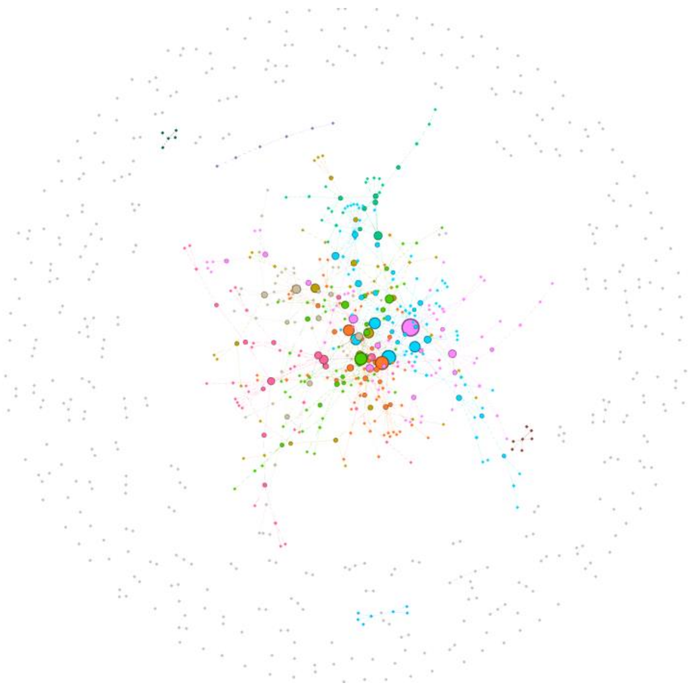
Datasets



Epinions



Ciao



FilmTrust

Dataset splits

- Train/test/validate
  - Train (*learn embeddings*) on the **undirected trust network**
    - We remove edge direction before learning embeddings on the input network
  - Validate (*select hyperparameters*) on **warm-start** users (>10 ratings)
  - Test (*evaluate model*) on **cold-start** users (≤10 ratings)

	Users with ratings		**Users with ratings & trust**	
Dataset	Warm-start	Cold-start	Warm-start (Validation set)	Cold-start (Test set)
Epinions	14,769	25,393	14,769	25,393
Ciao	1,020	16,591	571	2,124
Filmtrust	963	545	499	241

Recommender evaluation strategy

- Simulate ***unseen*** item recommendation ($n = 10$) in a ***kNN*** manner:

1. Remove n random items from the target user's (u_t) item history.
2. Find k-nearest neighbors N_k ($k = 40$) for the target user u_t from ***S***.
3. Assign a score for each item i users in N_k have interacted with:

$$score(i, u_t) = \sum_{v \in N_k} S_{u_t,v} \cdot R_v(i)$$

   where $R_v(i)$ corresponds to the rating assigned by user v to item i and $S_{u_t,v}$ represents the similarity score between the target user u_t and the neighbor user v.

4. Recommend n items ranked according to the above formula and compare them with the removed items.

Evaluation metrics

- **Accuracy**:
  - *nDCG@n - a ranking-dependent metric measuring recommendation accuracy based on the Discounted Cumulative Gain (DCG) measure* [7]
- ***Beyond accuracy***:
  - *novelty@n - corresponds to a recommender's ability to recommend long-tail items that the target user has probably not yet seen. We compute novelty using the Expected Popularity Complement (EPC) metric* [8]
  - *diversity@n* - describes how dissimilar items are in the recommendation list. We calculate it as the average dissimilarity of all pairs of items in the recommendation list [9]
  - *User Coverage* - defined as the number of users for whom at least one item recommendation could have been generated divided by the total number of users in the evaluation set [10]

[7] K. Järvelin, S. L. Price, L. M. Delcambre, and M. L. Nielsen, "Discounted cumulated gain based evaluation of multiple-query ir sessions," in Proceedings of ECIR'2008, pp. 4–15, Springer, Springer, 2008.
[8] S. Vargas and P. Castells, "Rank and relevance in novelty and diversity metrics for recommender systems," in Proceedings of the fifth ACM conference on Recommender systems, pp. 109–116, 2011.
[9] B. Smyth and P. McClave, "Similarity vs. diversity," in International conference on case-based reasoning, pp. 347–361, Springer, 2001.
[10] P. Massa and P. Avesani, "Trust-aware collaborative filtering for recommender systems," in OTM Confederated International Conferences" On the Move to Meaningful Internet Systems", pp. 492–508, Springer, 2004.

Evaluated approaches – latent user representations with graph embeddings

- <u>**Four**</u> distinct method families [11], i.e., **factorization-based** methods, **random-walk-based** approaches, **deep-learning-based** approaches, and the **LINE** approach which falls in neither of the first three families

[11] P. Goyal and E. Ferrara, “Graph embedding techniques, applications, and performance: A survey,” Knowledge-Based Systems, vol. 151, pp. 78 – 94, 2018.

Factorization-based approaches

Produce node embeddings using **matrix factorization**. The inner product between the resulting node embedding vectors approximates a deterministic graph proximity measure.

- *Graph Factorization* (**GF**) [12]
- *Laplacian Eigenmaps* (**LE**) [13]
- *Locally Linear Embedding* (**LLE**) [14]
- *High-Order Proximity preserved Embedding* (**HOPE**) [15]
- *Graph Representations with Global Structural Information* (**GraRep**) [16]

[12] A. Ahmed, N. Shervashidze, S. Narayanamurthy, V. Josifovski, and A. J. Smola, "Distributed large-scale natural graph factorization," in Proceedings of the 22nd international conference on World Wide Web, pp. 37–48, ACM, 2013.
[13] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," in Advances in neural information processing systems, pp. 585–591, 2002.
[14] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," science, vol. 290, no. 5500, pp. 2323–2326, 2000.
[15] M. Ou, P. Cui, J. Pei, Z. Zhang, and W. Zhu, "Asymmetric transitivity preserving graph embedding," in Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 1105–1114, ACM, 2016.
[16] S. Cao, W. Lu, and Q. Xu, "Grarep: Learning graph representations with global structural information," in Proceedings of the 24th ACM international on conference on information and knowledge management, pp. 891–900, ACM, 2015.

Random walk-based approaches

First identify the context of a node with a **random walk** and then learn the embeddings typically using a **skip-gram** model.

- *DeepWalk* [17]
- *Node2vec* [18]
- *Role2vec* [19]

[17] B. Perozzi, R. Al-Rfou, and S. Skiena, “Deepwalk: Online learning of social representations,” in Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’14, pp. 701–710, ACM, 2014.
[18] A. Grover and J. Leskovec, “node2vec: Scalable feature learning for networks,” in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016.
[19] N. K. Ahmed, R. Rossi, J. Boaz Lee, T. L. Willke, R. Zhou, X. Kong, and H. Eldardiry, “Learning Role-based Graph Embeddings,” arXiv e-prints, p. arXiv:1802.02896, Feb. 2018.

Deep Learning-based Approaches

Such approaches use **deep neural network** models to generate node embeddings.

- *Deep Neural Networks for Graph Representations* (**DNGR**) [20]
- *Structural Deep Network Embedding* (**SDNE**) [21]
- *Graph sample and aggregate* (**GraphSAGE**) [22]

[20] S. Cao, W. Lu, and Q. Xu, "Deep neural networks for learning graph representations," 2016.
[21] D. Wang, P. Cui, and W. Zhu, "Structural deep network embedding," in Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 1225–1234, ACM, 2016.
[22] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in Advances in Neural Information Processing Systems, pp. 1024–1034, 2017.

LINE

Large-Scale Information Network Embedding (**LINE**) [23] – creates embeddings that preserve **1st-order** and **2nd-order proximities** which are represented as joint and conditional probability distributions respectively.

[23] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, and Q. Mei, "Line: Large-scale information network embedding," in Proceedings of the 24th international conference on world wide web, pp. 1067–1077, International World Wide Web Conferences Steering Committee, 2015.

Evaluated approaches (baselines)

- Direct similarity metric approaches:
  - *Explicit directed trust* ($\mathbf{Trust}_{\mathbf{dir}}$)
  - *Explicit undirected trust* ($\mathbf{Trust}_{\mathbf{undir}}$)
  - *Explicit trust with Jaccard* ($\mathbf{Trust}_{\mathbf{jac}}$)
  - *Explicit trust with Katz similarity* ($\mathbf{Trust}_{\mathbf{Katz}}$) [24]
- *Most Popular* (**MP**)

[24] T. Duricic, E. Lacic, D. Kowald, and E. Lex, "Trust-based collaborative filtering: Tackling the cold start problem using regular equivalence," in Proceedings of the 12th ACM Conference on Recommender Systems, RecSys '18, pp. 446–450, ACM, 2018.

Results

Evaluation results

Cat.	Approach	Rank nDCG	Epinions nDCG	Epinions Nov.	Epinions Div.	Epinions UC	Ciao nDCG	Ciao Nov.	Ciao Div.	Ciao UC	Filmtrust nDCG	Filmtrust Nov.	Filmtrust Div.	Filmtrust UC
Baseline	$Trust_{dir}$	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	$Trust_{undir}$	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	$Trust_{jac}$	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	$Trust_{Katz}$	12	.0290	.0046	.6979	100 %	.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013	13.1 %	.3649	.0159	.1926	44.2 %
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581		.0228	.0036	.4042		**.3904**	.0151	.2235	
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, p < 0:01).

Evaluation results

Cat.	Approach	Rank nDCG	Epinions				Ciao				Filmtrust			
			nDCG	Nov.	Div.	UC	nDCG	Nov.	Div.	UC	nDCG	Nov.	Div.	UC
Baseline	Trust_{dir}	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	Trust_{undir}	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	Trust_{jac}	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	Trust_{Katz}	12	.0290	.0046	.6979	100 %	.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013	13.1 %	.3649	.0159	.1926	44.2 %
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581		.0228	.0036	.4042		**.3904**	.0151	.2235	
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, $p < 0:01$).

Evaluation results

Cat.	Approach	Rank nDCG	Epinions nDCG	Epinions Nov.	Epinions Div.	Epinions UC	Ciao nDCG	Ciao Nov.	Ciao Div.	Ciao UC	Filmtrust nDCG	Filmtrust Nov.	Filmtrust Div.	Filmtrust UC
Baseline	Trust_{dir}	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	Trust_{undir}	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	Trust_{jac}	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	Trust_{Katz}	12	.0290	.0046	.6979	100 %	.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013	13.1 %	.3649	.0159	.1926	44.2 %
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581		.0228	.0036	.4042		**.3904**	.0151	.2235	
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, p < 0:01).

Evaluation results

Cat.	Approach	Rank nDCG	Epinions nDCG	Epinions Nov.	Epinions Div.	Epinions UC	Ciao nDCG	Ciao Nov.	Ciao Div.	Ciao UC	Filmtrust nDCG	Filmtrust Nov.	Filmtrust Div.	Filmtrust UC
Baseline	Trust$_{dir}$	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	Trust$_{undir}$	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	Trust$_{jac}$	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	Trust$_{Katz}$	12	.0290	.0046	.6979		.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013		.3649	.0159	.1926	
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704	100 %	.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581		.0228	.0036	.4042	13.1 %	**.3904**	.0151	.2235	44.2 %
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, p < 0:01).

Evaluation results

Cat.	Approach	Rank *nDCG*	Epinions nDCG	Epinions Nov.	Epinions Div.	Epinions UC	Ciao nDCG	Ciao Nov.	Ciao Div.	Ciao UC	Filmtrust nDCG	Filmtrust Nov.	Filmtrust Div.	Filmtrust UC
Baseline	Trust_{dir}	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	Trust_{undir}	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	Trust_{jac}	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	Trust_{Katz}	12	.0290	.0046	.6979	100 %	.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013	13.1 %	.3649	.0159	.1926	44.2 %
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581		.0228	.0036	.4042		**.3904**	.0151	.2235	
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, p < 0:01).

Evaluation results

Cat.	Approach	Rank $nDCG$	Epinions				Ciao				Filmtrust			
			nDCG	Nov.	Div.	UC	nDCG	Nov.	Div.	UC	nDCG	Nov.	Div.	UC
Baseline	Trust_{dir}	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	Trust_{undir}	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	Trust_{jac}	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	Trust_{Katz}	12	.0290	.0046	.6979		.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013		.3649	.0159	.1926	
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581	100 %	.0228	.0036	.4042	13.1 %	**.3904**	.0151	.2235	44.2 %
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, p < 0:01).

Evaluation results

Cat.	Approach	Rank $nDCG$	Epinions nDCG	Epinions Nov.	Epinions Div.	Epinions UC	Ciao nDCG	Ciao Nov.	Ciao Div.	Ciao UC	Filmtrust nDCG	Filmtrust Nov.	Filmtrust Div.	Filmtrust UC
Baseline	Trust_{dir}	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	Trust_{undir}	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	Trust_{jac}	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	Trust_{Katz}	12	.0290	.0046	.6979	100 %	.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013	13.1 %	.3649	.0159	.1926	44.2 %
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581		.0228	.0036	.4042		**.3904**	.0151	.2235	
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, p < 0:01).

Evaluation results

Cat.	Approach	Rank $nDCG$	Epinions nDCG	Epinions Nov.	Epinions Div.	Epinions UC	Ciao nDCG	Ciao Nov.	Ciao Div.	Ciao UC	Filmtrust nDCG	Filmtrust Nov.	Filmtrust Div.	Filmtrust UC
Baseline	Trust$_{dir}$	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	Trust$_{undir}$	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	Trust$_{jac}$	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	Trust$_{Katz}$	12	.0290	.0046	.6979		.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013		.3649	.0159	.1926	
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581	100 %	.0228	.0036	.4042	13.1 %	**.3904**	.0151	.2235	44.2 %
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, p < 0:01).

Evaluation results

Cat.	Approach	Rank $_{nDCG}$	Epinions				Ciao				Filmtrust			
			nDCG	Nov.	Div.	UC	nDCG	Nov.	Div.	UC	nDCG	Nov.	Div.	UC
Baseline	Trust_{dir}	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	Trust_{undir}	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	Trust_{jac}	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	Trust_{Katz}	12	.0290	.0046	.6979	100 %	.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013	13.1 %	.3649	.0159	.1926	44.2 %
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581		.0228	.0036	.4042		**.3904**	.0151	.2235	
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, p < 0:01).

Evaluation results

Cat.	Approach	Rank nDCG	Epinions nDCG	Epinions Nov.	Epinions Div.	Epinions UC	Ciao nDCG	Ciao Nov.	Ciao Div.	Ciao UC	Filmtrust nDCG	Filmtrust Nov.	Filmtrust Div.	Filmtrust UC
Baseline	Trust_{dir}	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	Trust_{undir}	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	Trust_{jac}	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	Trust_{Katz}	12	.0290	.0046	.6979	100 %	.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013	13.1 %	.3649	.0159	.1926	44.2 %
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581		.0228	.0036	.4042		**.3904**	.0151	.2235	
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, p < 0:01).

Evaluation results

Cat.	Approach	Rank nDCG	Epinions nDCG	Epinions Nov.	Epinions Div.	Epinions UC	Ciao nDCG	Ciao Nov.	Ciao Div.	Ciao UC	Filmtrust nDCG	Filmtrust Nov.	Filmtrust Div.	Filmtrust UC
Baseline	Trust$_{dir}$	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	Trust$_{undir}$	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	Trust$_{jac}$	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	Trust$_{Katz}$	12	.0290	.0046	.6979	100 %	.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013	13.1 %	.3649	.0159	.1926	44.2 %
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581		.0228	.0036	.4042		**.3904**	.0151	.2235	
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, p < 0:01).

Evaluation results

Cat.	Approach	Rank $nDCG$	Epinions				Ciao				Filmtrust			
			nDCG	Nov.	Div.	UC	nDCG	Nov.	Div.	UC	nDCG	Nov.	Div.	UC
Baseline	Trust$_{dir}$	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	Trust$_{undir}$	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	Trust$_{jac}$	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	Trust$_{Katz}$	12	.0290	.0046	.6979	100 %	.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013	13.1 %	.3649	.0159	.1926	44.2 %
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581		.0228	.0036	.4042		**.3904**	.0151	.2235	
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, p < 0:01).

Evaluation results

Cat.	Approach	Rank (nDCG)	Epinions nDCG	Epinions Nov.	Epinions Div.	Epinions UC	Ciao nDCG	Ciao Nov.	Ciao Div.	Ciao UC	Filmtrust nDCG	Filmtrust Nov.	Filmtrust Div.	Filmtrust UC
Baseline	$Trust_{dir}$	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	$Trust_{undir}$	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	$Trust_{jac}$	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	$Trust_{Katz}$	12	.0290	.0046	.6979		.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977	100 %	.0239	.0036	.4013	13.1 %	.3649	.0159	.1926	44.2 %
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581		.0228	.0036	.4042		**.3904**	.0151	.2235	
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, p < 0:01).

Evaluation results

Cat.	Approach	Rank $nDCG$	Epinions nDCG	Epinions Nov.	Epinions Div.	Epinions UC	Ciao nDCG	Ciao Nov.	Ciao Div.	Ciao UC	Filmtrust nDCG	Filmtrust Nov.	Filmtrust Div.	Filmtrust UC
Baseline	Trust_{dir}	15	.0245	.0060	.6006	59.2%	.0140	.0028	.3700	3.9%	.2655	.0313	**.2784**	30.3%
	Trust_{undir}	15	.0260	.0063	.5960	97.0%	.0127	**.0045**	.3632	11.4%	.2739	.0284	.2731	42.0%
	Trust_{jac}	11	.0373	.0056	.6548	99.9%	.0176	.0027	.3996	12.8%	.3387	**.0369**	.2266	36.1%
	Trust_{Katz}	12	.0290	.0046	.6979	100 %	.0158	.0026	.3842	13.0%	.3681	.0322	.2185	42.9%
	MP	17	.0134	.0015	**.7621***		.0135	.0012	**.5666**	100%	.3551	.0137	.1672	100%
Factorization	LLE	7	.0309	.0044	.6977		.0239	.0036	.4013	13.1 %	.3649	.0159	.1926	44.2 %
	LE	3	.0318	.0045	.6961		.0231	.0034	.3962		.3715	.0161	.1853	
	GF	14	.0138	.0023	.7024		.0154	.0022	.3970		.3686	.0154	.1945	
	HOPE	3	.0331	.0047	.6728		.0220	.0033	.3956		.3718	.0158	.1827	
	GraRep	7	.0298	.0042	.6704		.0209	.0030	.3974		.3694	.0147	.1859	
RW	Node2vec	**1**	.0413	.0064	.6581		.0228	.0036	.4042		**.3904**	.0151	.2235	
	DeepWalk	2	**.0435***	**.0067***	.6707		**.0247***	.0037	.3992		.3654	.0152	.1950	
	Role2vec	6	.0363	.0054	.6910		.0149	.0024	.3933		.3695	.0151	.1919	
DL	DNGR	10	.0353	.0051	.6869		.0197	.0031	.4023		.3583	.0142	.1959	
	SDNE	12	.0184	.0022	.7412		.0175	.0028	.3921		.3687	.0152	.2003	
	GS	7	.0325	.0047	.6810		.0216	.0031	.3963		.3678	.0151	.1883	
LINE	LINE	5	.0407	.0063	.6566		.0222	.0033	.3992		.3667	.0150	.1947	

Values marked with * denote that the corresponding approach was significantly better than every other approach with respect to the appropriate metric according to a Wilcoxon signed-rank test (Bonferroni corrected, $p < 0:01$).

Evaluation metrics and user preferences

- Finally, we compute the **Kendall rank correlation** coefficient (Bonferroni corrected, $p < 0.01$) and observe the following:
  - Statistically significant **positive** mean correlation across all three datasets between nDCG and novelty, ranging from **0.43** on Epinions to **0.36** on Filmtrust
  - We also observe a statistically significant **negative** mean correlation between diversity and novelty on Epinions (**-0.15**)

Conclusion

Limitations and future work

- In our experiments, we treated the trust networks as **undirected** while, in reality, they are **directed**
  - *We aim to further explore how to preserve different properties of trust networks (e.g., asymmetry)*
- It is possible that we did not examine an ample enough space of **hyperparameters**[1]
- We used only ***A*** as input for the recommendation algorithms, we also plan to incorporate user features from ***R***
- We are also going to focus on **interpretability** of our results by studying node properties of **user neighborhoods**

[1] Details on the hyperparameter optimization can be found at: https://github.com/tduricic/trust-recommender/blob/master/docs/hyperparameter-optimization.md

Key takeaways

- We explored the **utility of graph embedding** approaches from **four** method families to generate latent user representations for **trust-based recommender systems** in a **cold-start setting**:
    - Random-walk-based approaches *Node2vec* and *DeepWalk* consistently achieve the best accuracy
    - *Node2vec* and *DeepWalk* scored high on novelty and diversity as well
    - Graph embeddings increase user coverage
    - In all three evaluated datasets, users tend to prefer novel recommendations

Thank you! Questions? Contact/follow us.

Get the paper:

https://arxiv.org/pdf/2003.13345.pdf

Social Computing

know-center.tugraz.at/research/areas/social-computing/

s-computing@know-center.at

@socialcomplab

Tomislav Duricic

tduricic@tugraz.at

@djurazzi

scholar.google.com/citations?user=_mY5j2UAAAAJ

researchgate.net/profile/Tomislav_Duricic

linkedin.com/in/tomislav-đuričić-954b54172