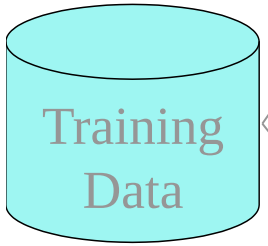
Transfer Learning in Social Recommendation Systems

Qiang Yang

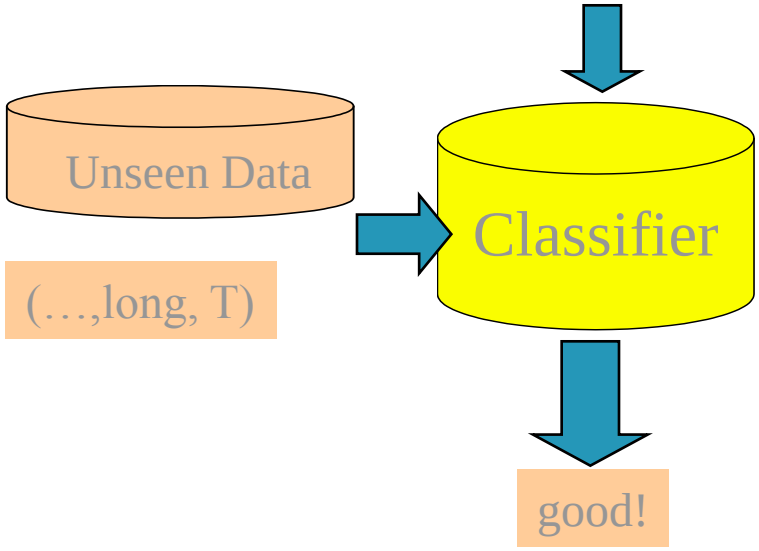
Hong Kong University of Science and Technology
Hong Kong, China

http://www.cse.ust.hk/~qyang

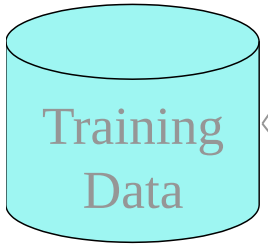
Traditional Machine Learning



Occ	Palm Lines	ACL	Fortune?
Prof	long	T	good
Lawyer	short	F	bad
PhD Stu	broken	T	good
Doc	long	F	bad

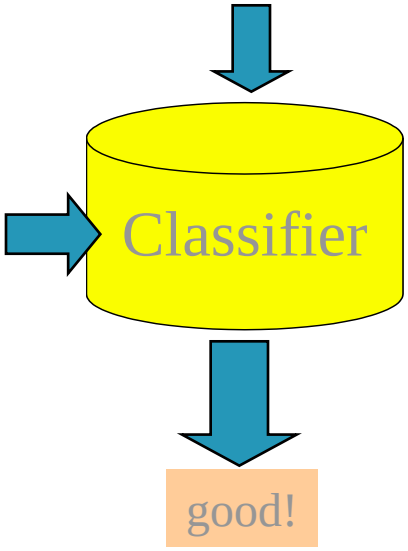


Traditional Machine Learning



Occ	Palm Lines	ACL	Fortune?
Prof	long	T	good
Lawyer	short	F	bad
PhD Stu	broken	T	good
Doc	long	F	bad

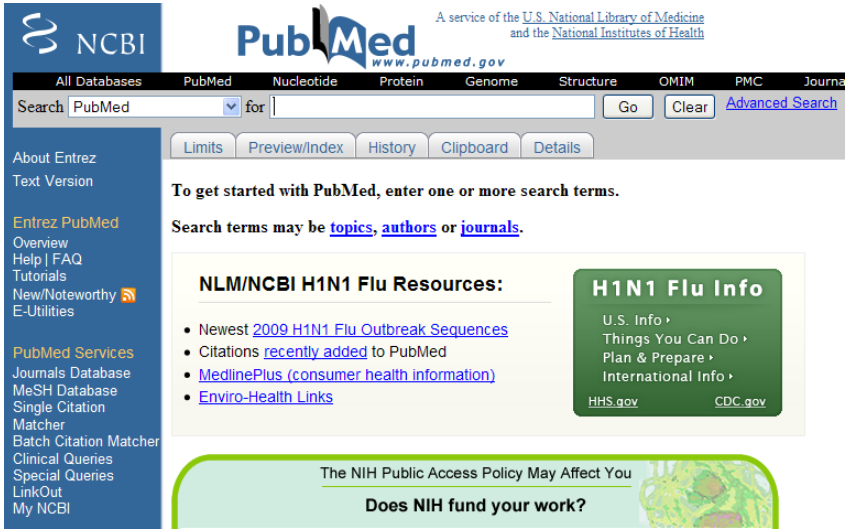
What if…



A Major Assumption

- Training and future (test) data
  - follow the same distribution, and
  - are in same feature space

When distributions are different



- Part-of-Speech tagging
- Named-Entity Recognition
- Classification

When Features are different

- Heterogeneous: different feature spaces

Training: Text

Future: Images

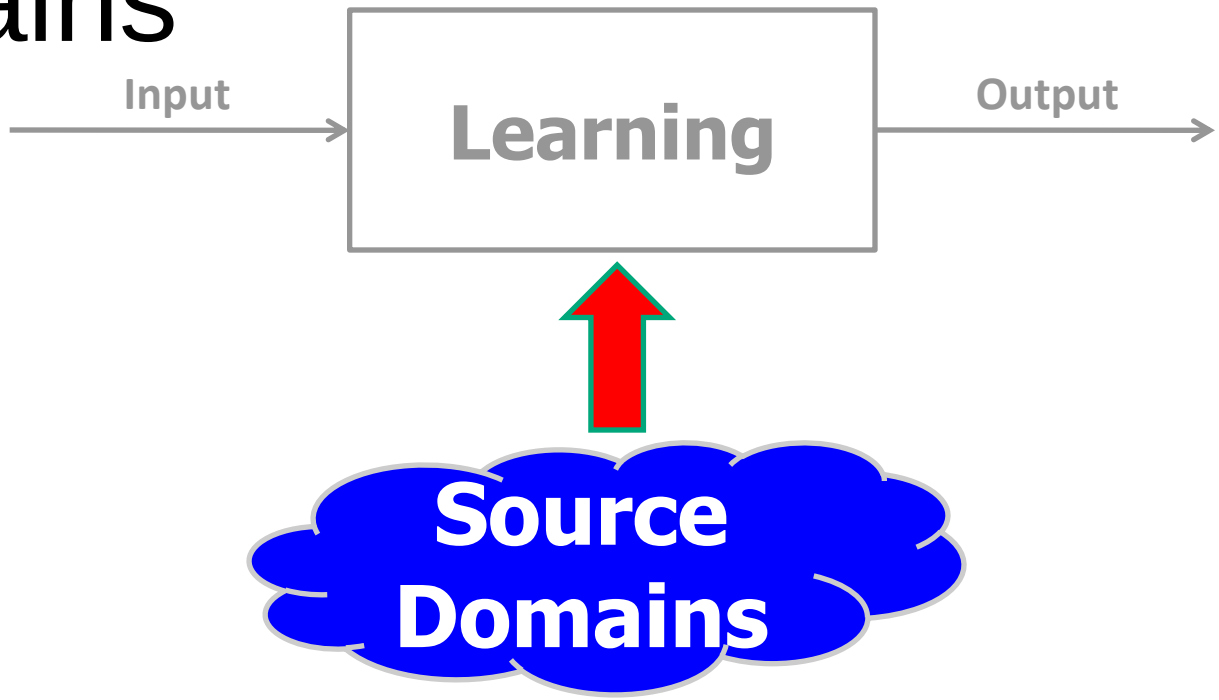
Apples

The apple is the pomaceous fruit of the apple tree, species Malus domestica in the rose family Rosaceae ...

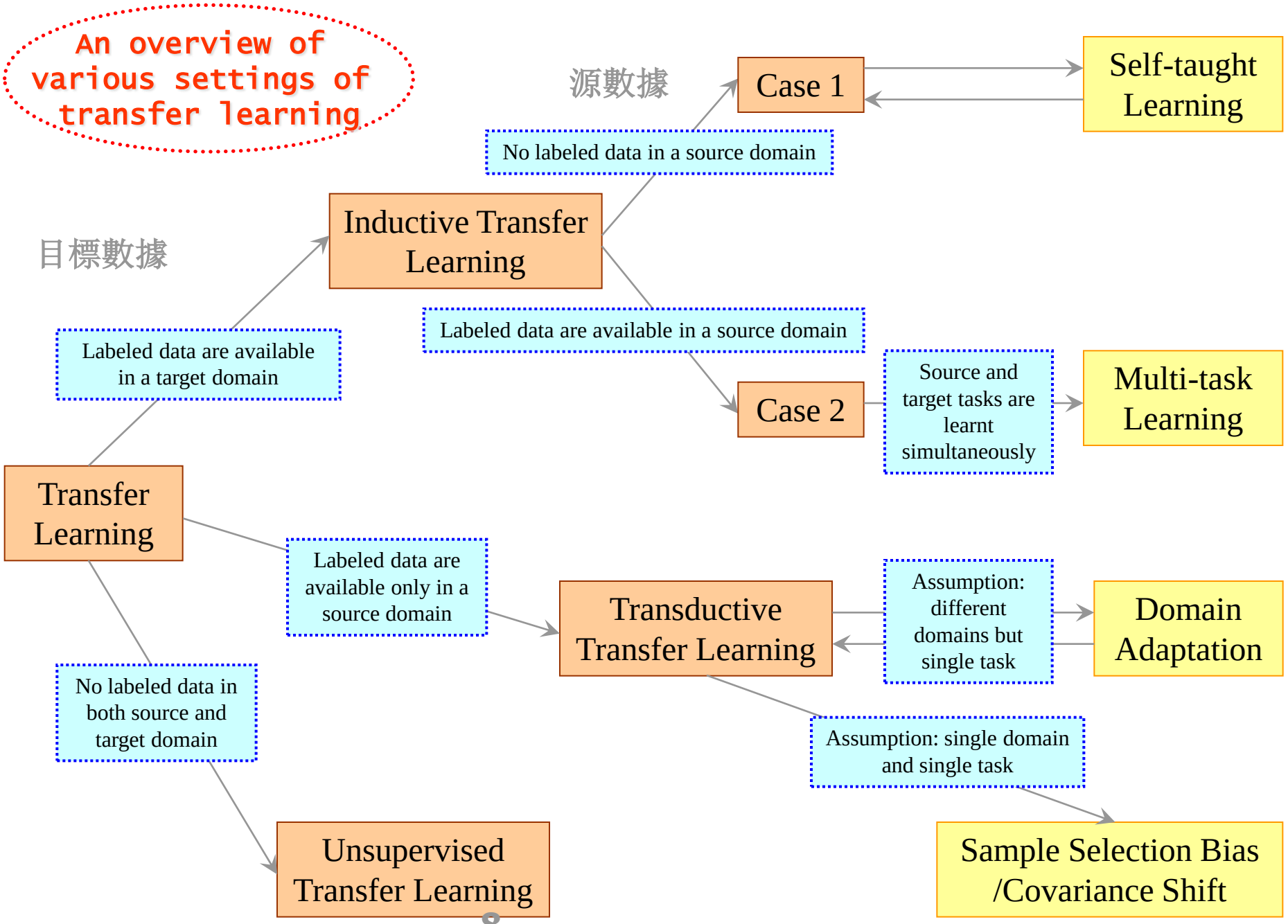
Bananas

Banana is the common name for a type of fruit and also the herbaceous plants of the genus Musa which produce this commonly eaten fruit ...

Transfer Learning: Source Domains



	Source Domain	Target Domain
Training Data	Labeled/Unlabeled	Labeled/Unlabeled
Test Data		Unlabeled


An overview of various settings of transfer learning
源數據
目標數據
Case 1
Self-taught Learning
No labeled data in a source domain
Inductive Transfer Learning
Labeled data are available in a target domain
Labeled data are available in a source domain
Case 2
Source and target tasks are learnt simultaneously
Multi-task Learning
Transfer Learning
Labeled data are available only in a source domain
Transductive Transfer Learning
Assumption: different domains but single task
Domain Adaptation
No labeled data in both source and target domain
Assumption: single domain and single task
Unsupervised Transfer Learning
Sample Selection Bias /Covariance Shift

An overview of various settings of transfer learning

源數據

目標數據

Transfer Learning

Labeled data are available in a target domain

Labeled data are available in a source domain

No labeled data in both source and target domain

Inductive Transfer Learning

No labeled data in a source domain

Case 1

Self-taught Learning

Labeled data are available in a source domain

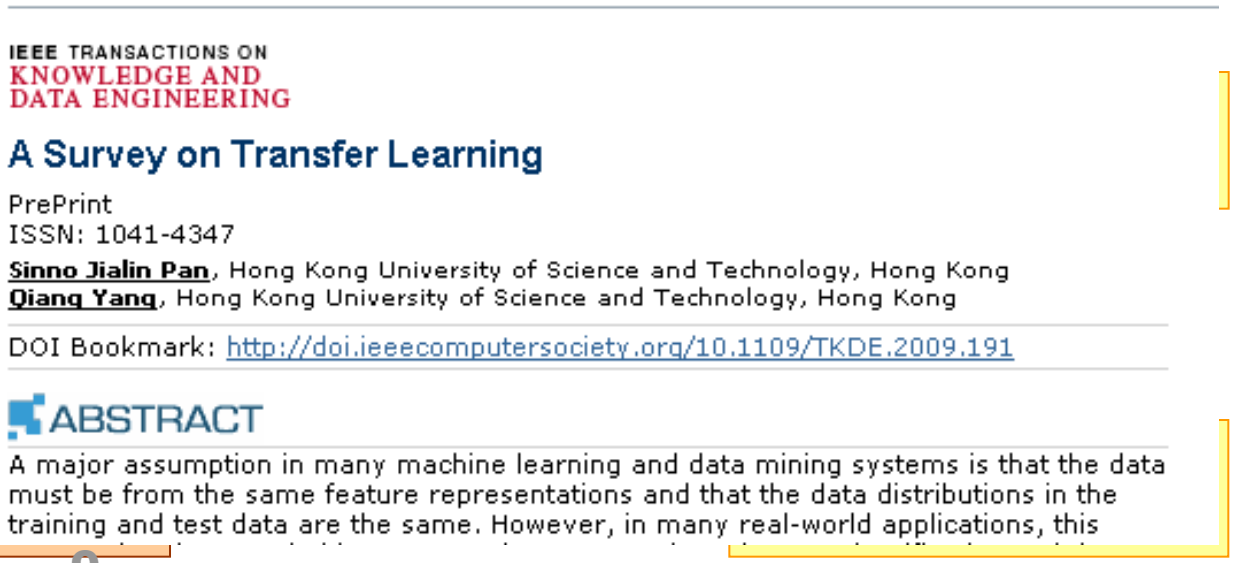
Case 2

Source and target tasks are learnt simultaneously

Multi-task Learning

Labeled available source

Unsupervised Transfer



IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING

A Survey on Transfer Learning

PrePrint
ISSN: 1041-4347

Sinno Jialin Pan, Hong Kong University of Science and Technology, Hong Kong
Qiang Yang, Hong Kong University of Science and Technology, Hong Kong

DOI Bookmark: http://doi.ieeecomputersociety.org/10.1109/TKDE.2009.191

ABSTRACT

A major assumption in many machine learning and data mining systems is that the data must be from the same feature representations and that the data distributions in the training and test data are the same. However, in many real-world applications, this

TL Resources

- **Sinno Jialin Pan's Homepage**
  - http://www1.i2r.a-star.edu.sg/~jspan/SurveyTL.htm
- Source Code, Data and References
  - http://www.cse.ust.hk/TL

Social Media Can Be Bridges in Transfer Learning

Social Web for Transfer Learning

- Source data: labeled or unlabeled
- Target training data: labeled
- ACL 2009

Training: Text

Testing: Images

Apple

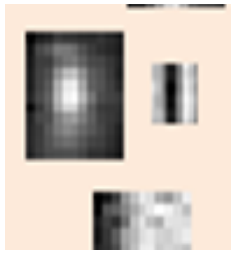
The apple is the pomaceous fruit of the apple tree, species Malus domestica in the rose family Rosaceae ...

Banana

Banana is the common name for a type of fruit and also the herbaceous plants of the genus Musa which produce this commonly eaten fruit ...

Annotated PLSA Model for Clustering Z

Words from Source Data



SIFT Features

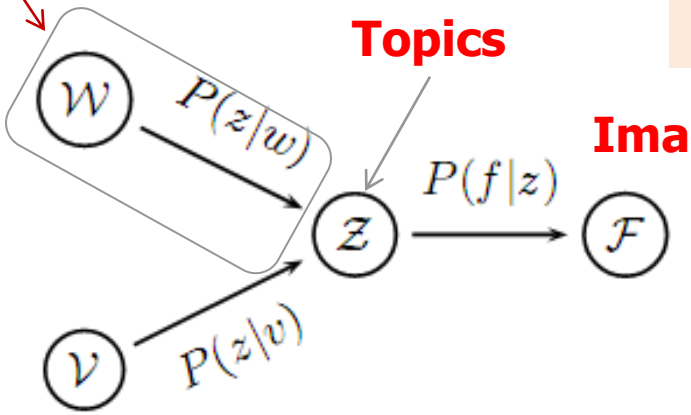


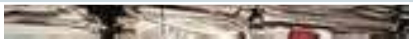
Image instances in **target** data

From Flickr.com

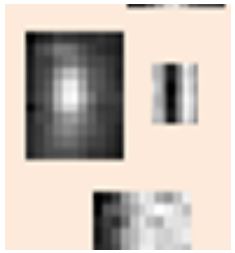
… Tags
Lion
Animal
Simba
Hakuna
Matata
FlickrBigCats
…

Annotated PLSA Model for Clustering Z

Caltech 256 Data	Heterogeneous Transfer Learning
Average Entropy Improvement	5.7%



Words from Source Data



Topics

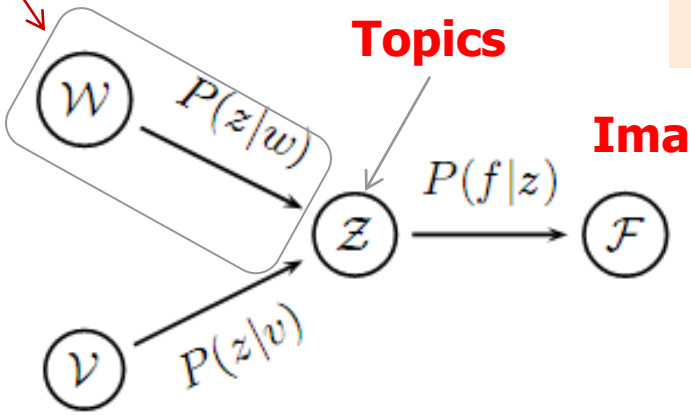


Image features

Image instances in target data

From Flickr.com

… Tags
Lion
Animal
Simba
Hakuna
Matata
FlickrBigCats
…

- “Heterogeneous transfer learning for image classification”
  - Y. Zhu, G. Xue, Q. Yang et al.
  - AAAI 2011

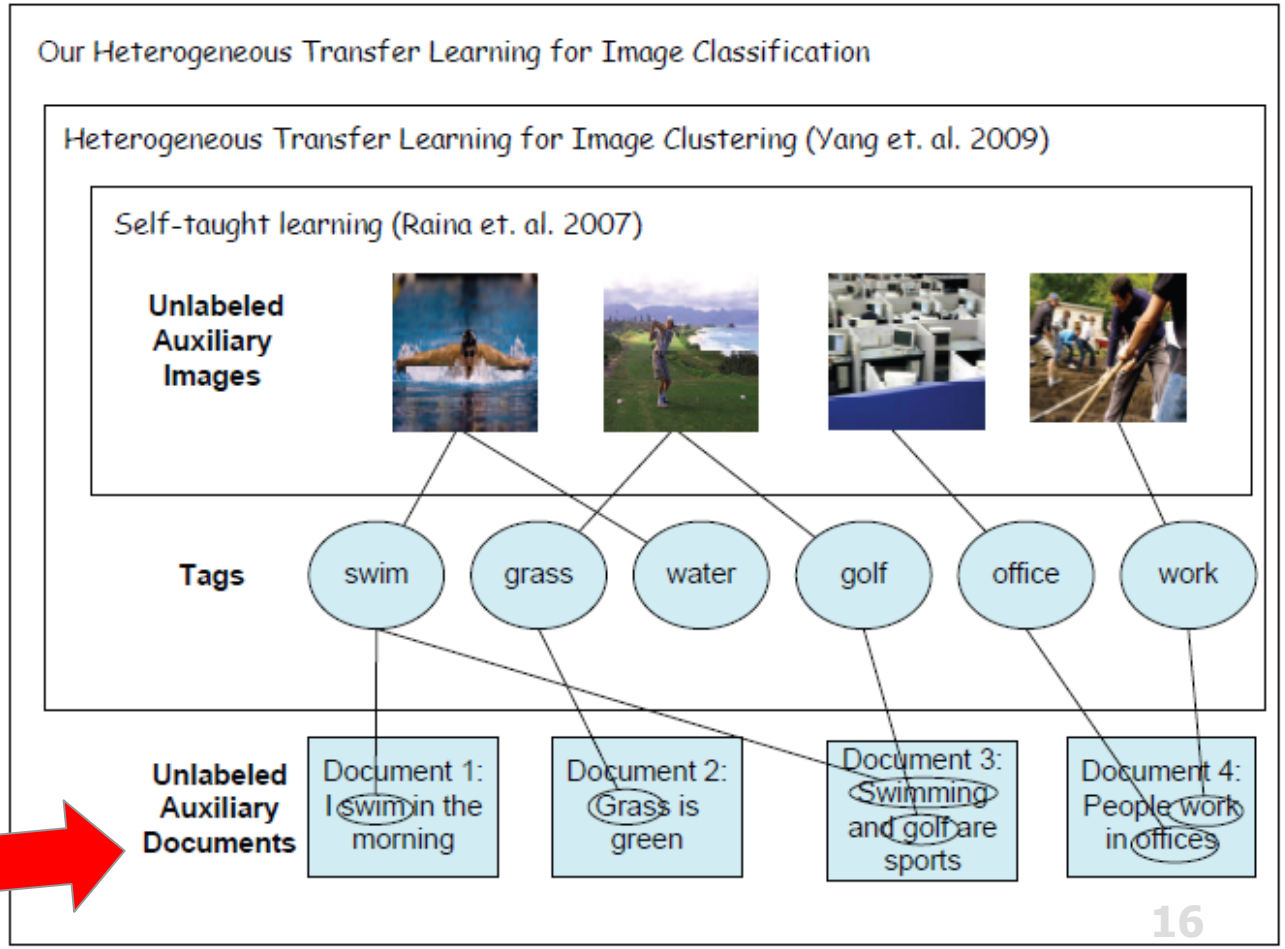
Source Data: Unlabeled Documents

Target data

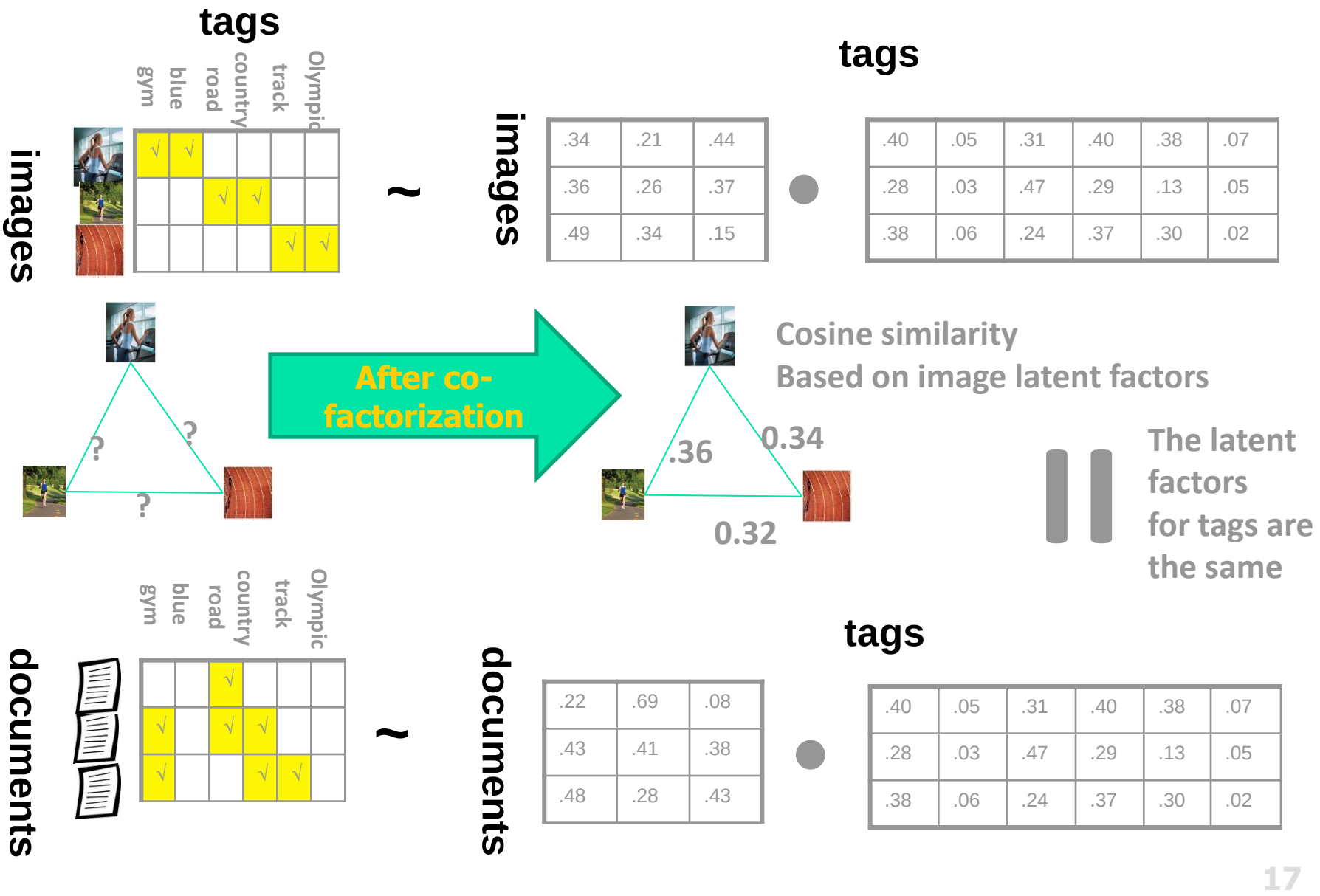
A few labeled images as training samples

Testing samples: not available during training.

Unlabeled Source data



Latent Feature Learning by Collective matrix factorization



Optimization:
Collective Matrix Factorization (CMF)

- G1 - `image-features'-tag matrix
- G2 – document-tag matrix
- W – words-latent matrix
- U – `image-features'-latent matrix
- V – tag-latent matrix
- R(U,V, W) - regularization to avoid over-fitting

$$\min_{\mathbf{U},\mathbf{V},\mathbf{W}} \lambda||\mathbf{G}_1-\mathbf{U}\mathbf{V}^T||+(1-\lambda)||\mathbf{G}_2-\mathbf{W}\mathbf{V}^T||+R(\mathbf{U},\mathbf{V},\mathbf{W})$$

The latent semantic view of images

The latent semantic view of tags

$$\mathbf{U} = \{u_j | j = 1..m\}$$

u_j is a row vector.

Heterogeneous Transfer Learning Algorithm

Algorithm 1 Image Semantic View Learning via CMF

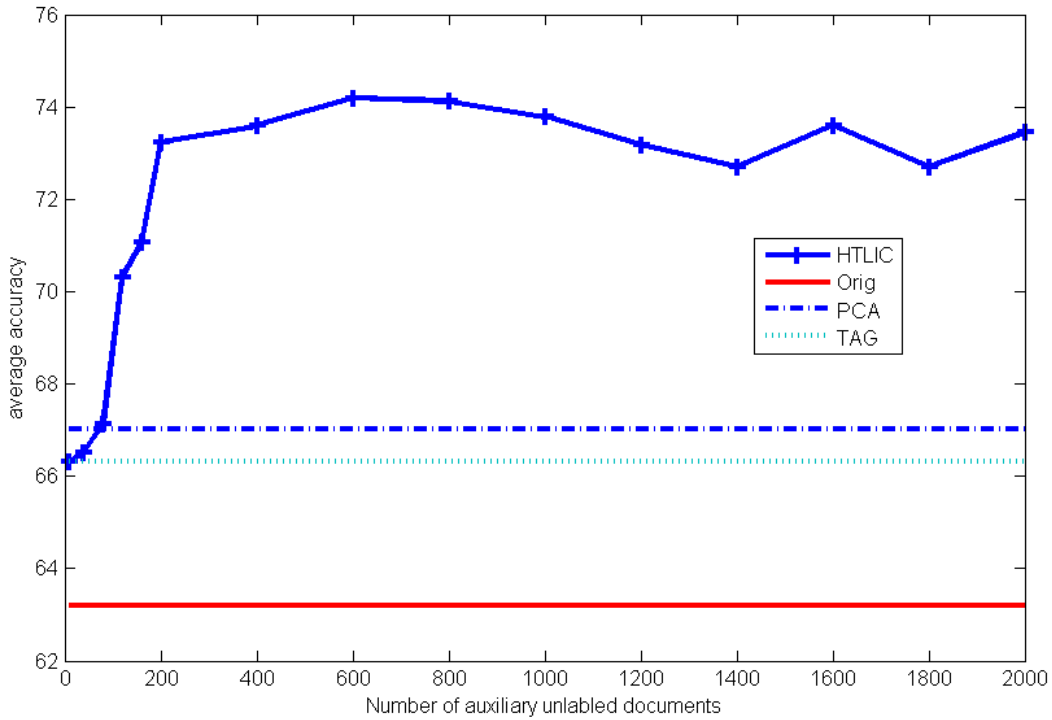
Input: A auxiliary image matrix $\mathbf{Z}$ with its corresponding annotation matrix $\mathbf{T}$, a document-tag relational matrix $\mathbf{F}$, a parameter λ, and the number of latent factors g.

Output: A new representation $\mathbf{U}$ for images $\mathbf{Z}$.

1: Compute $\mathbf{G} = \mathbf{Z}^\top \mathbf{T}$ and randomly initialize matrices $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{W}$.
2: **repeat**
3: Fix $\mathbf{U}$ and $\mathbf{V}$, apply conjugate gradient descent (CGD) (Shewchuk 1994) on (1) to update $\mathbf{W}$;
4: Fix $\mathbf{U}$ and $\mathbf{W}$, apply CGD on (1) to update $\mathbf{V}$;
5: Fix $\mathbf{W}$ and $\mathbf{V}$, apply CGD on (1) to update $\mathbf{U}$;
6: **until** $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{W}$ are convergent.

Experiment: # documents

Accuracy

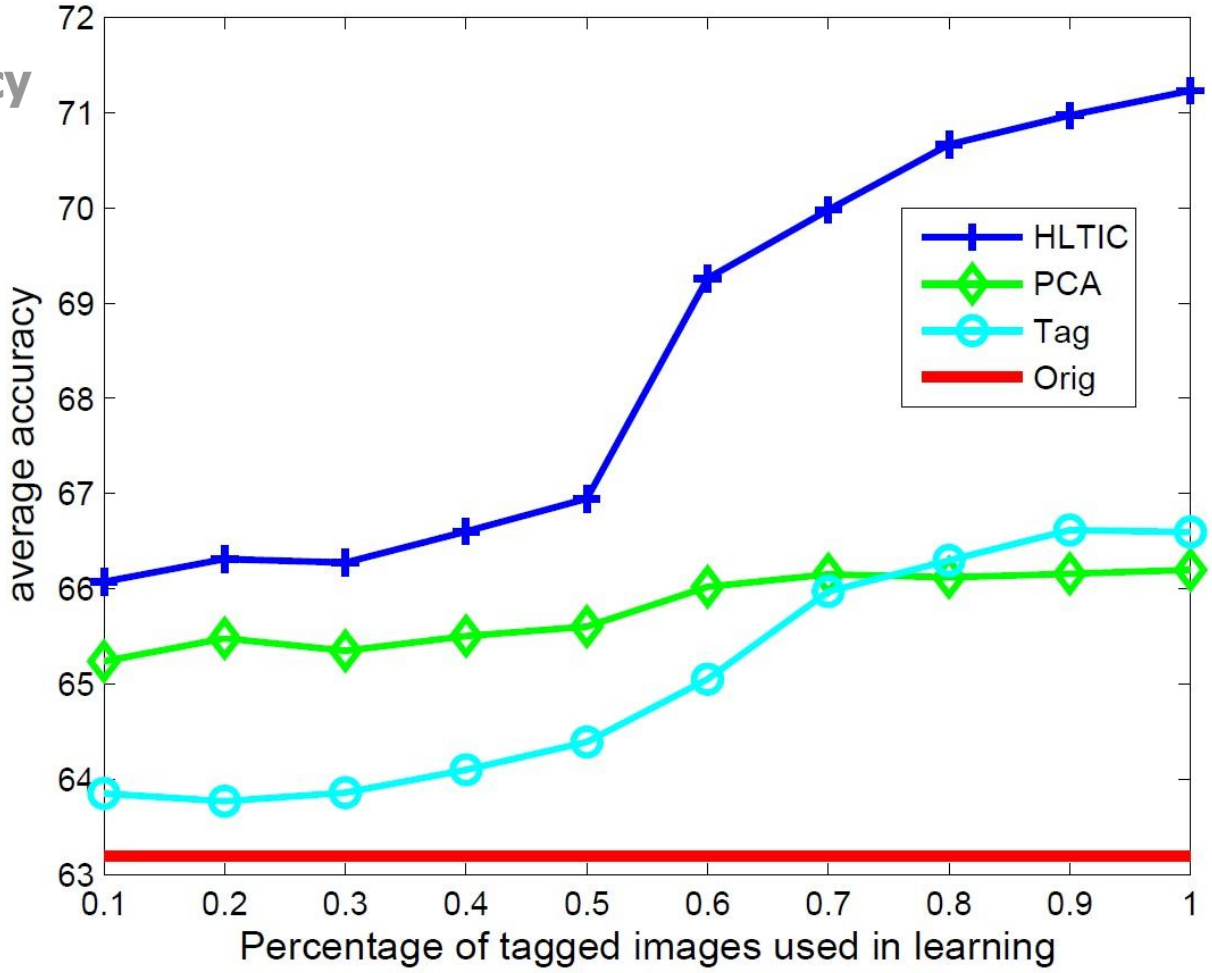


 # documents

When more text documents are used in learning, the accuracy increases.

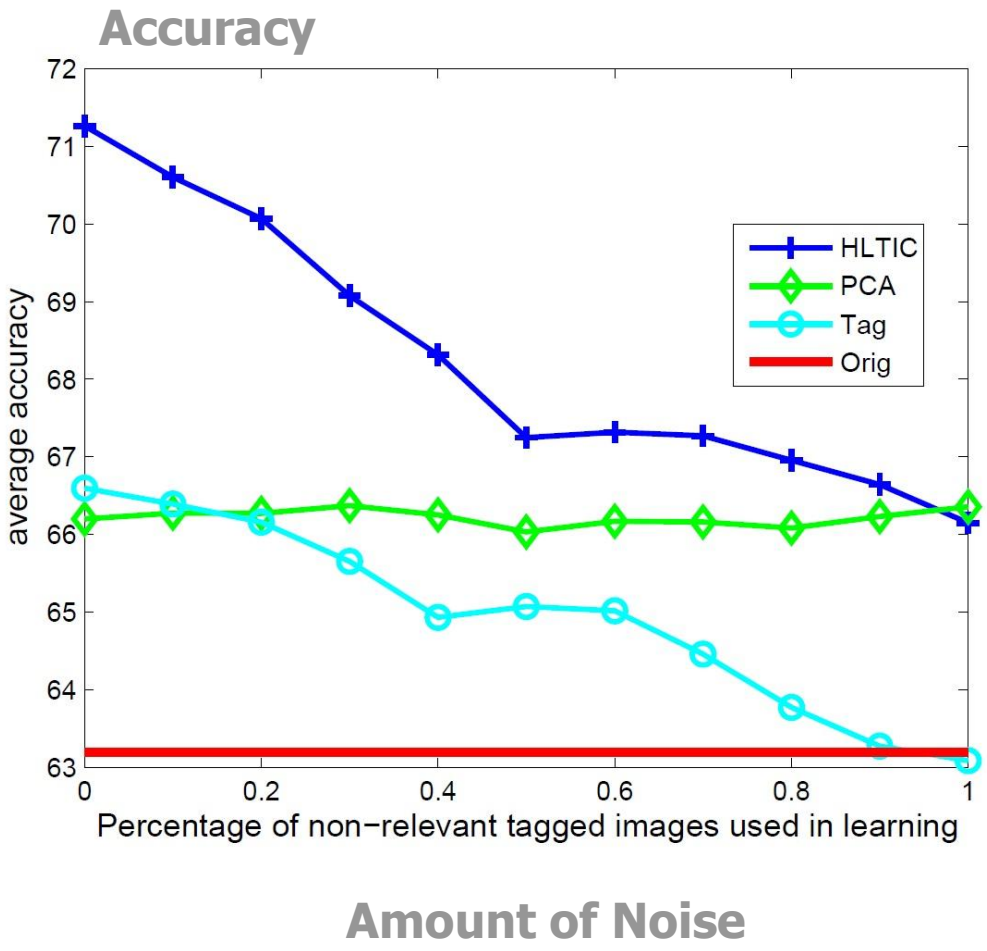
Experiment: # Tagged images

Accuracy



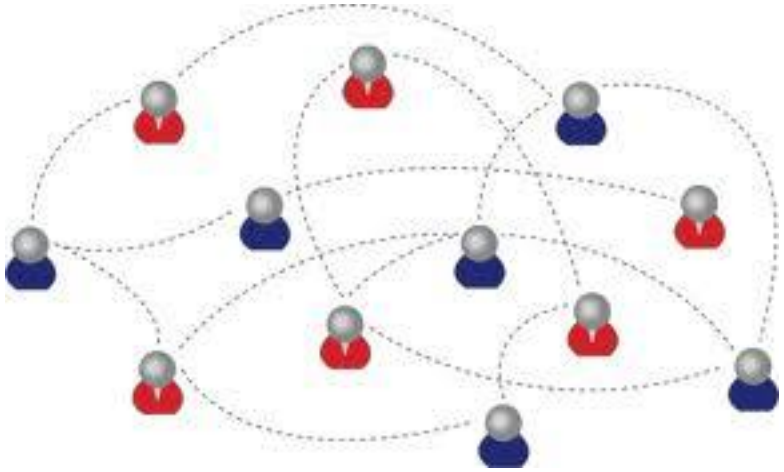
Tagged Images

Experiment: Noise



- We considered the "noise" of the tagged image.
- When the tagged images are totally irrelevant, our method reduced to PCA; and the Tag baseline, which depends on tagged images, reduced to a pure SVM.

Social Recommendations as Source Data



Recommendation Systems

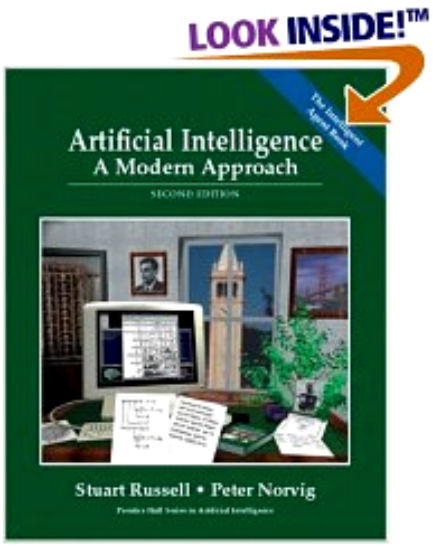
Recommendation Systems

amazon.com | Your Amazon.com | Books | See all 41 Product Categories | Your Account | Cart | Your Lists | Help |

Advanced Search | Browse Subjects | Bestsellers | The New York Times® Best Sellers | New & Future Releases | Libros En Español | Magazines | Sell Your Stuff |

Search Books GO Gift Certificates

Join Amazon Prime and ship Two-Day for free and Overnight for $3.99. Already a member? Sign in.



hare your own customer images
inside another edition of this book

Artificial Intelligence: A Modern Approach (2nd Edition) (Hardcover)

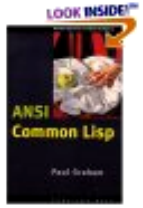
by Stuart J. Russell (Author), Peter Norvig (Author)

★★★★☆ (70 customer reviews)

Customers Who Bought This Item Also Bought

Machine Learning by Tom M. Mitchell

★★★★½ (28) $123.79



ANSI Common LISP by Paul Graham

★★★★☆ (25) $60.30

Paradigms of Artificial Intelligence Programming: Case Studies in Common Lisp by Peter Norvig

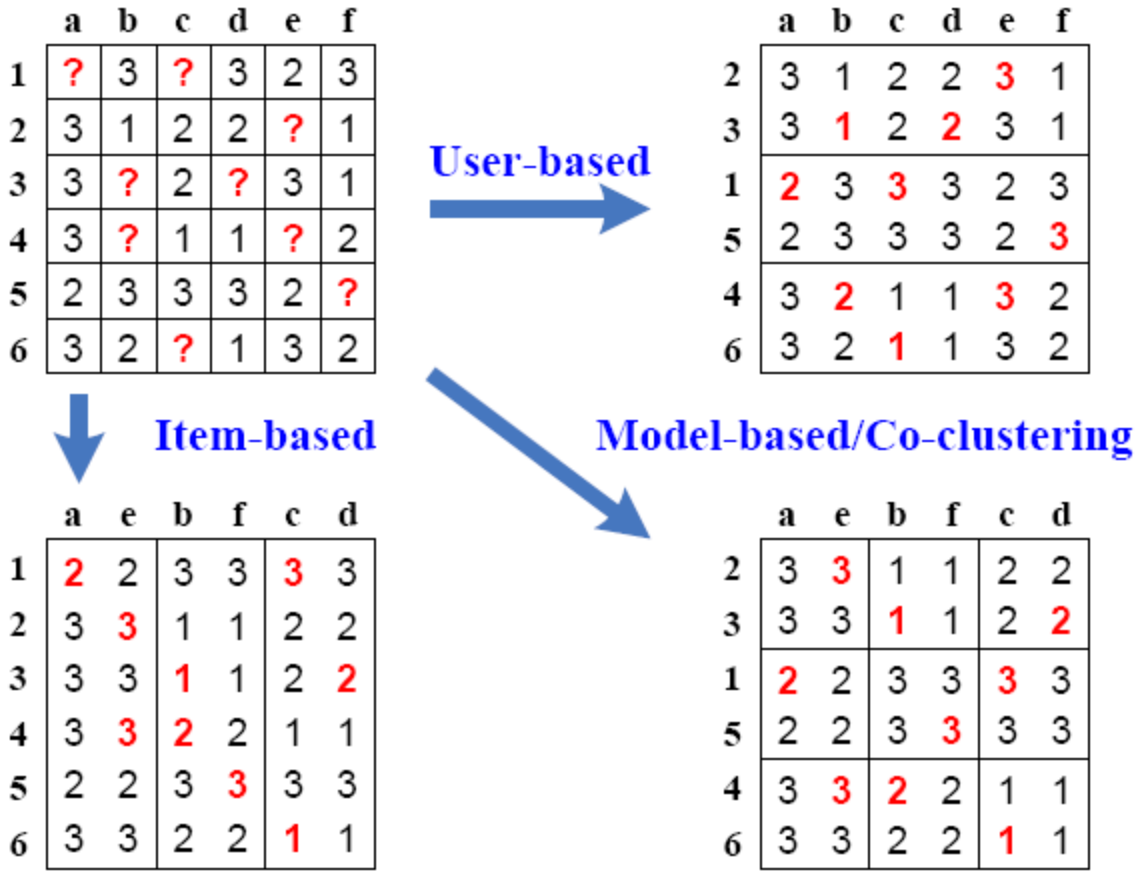
★★★★★ (7) $64.74

Product Recommendation as Link Prediction

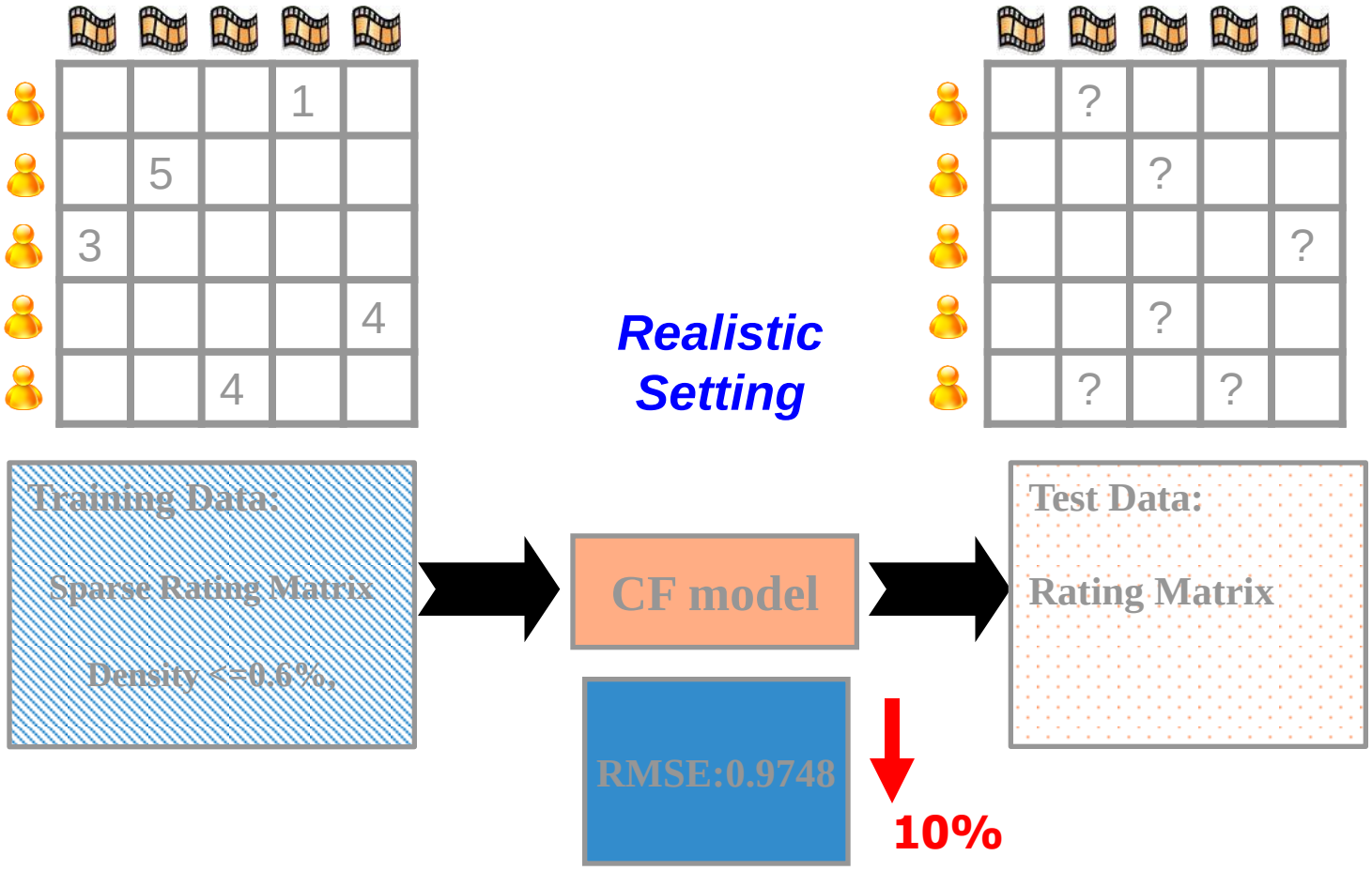
- Task: predict missing links in a network
- Focus:
  - bipartite graph of users and items

Essentials of Collaborative Filtering

- Discover latent user/item groups by (co)-clustering
- Share ratings within clusters to fill in missing values

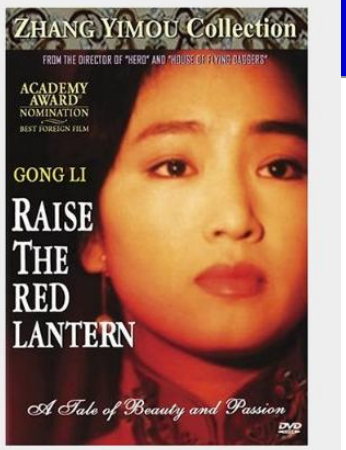


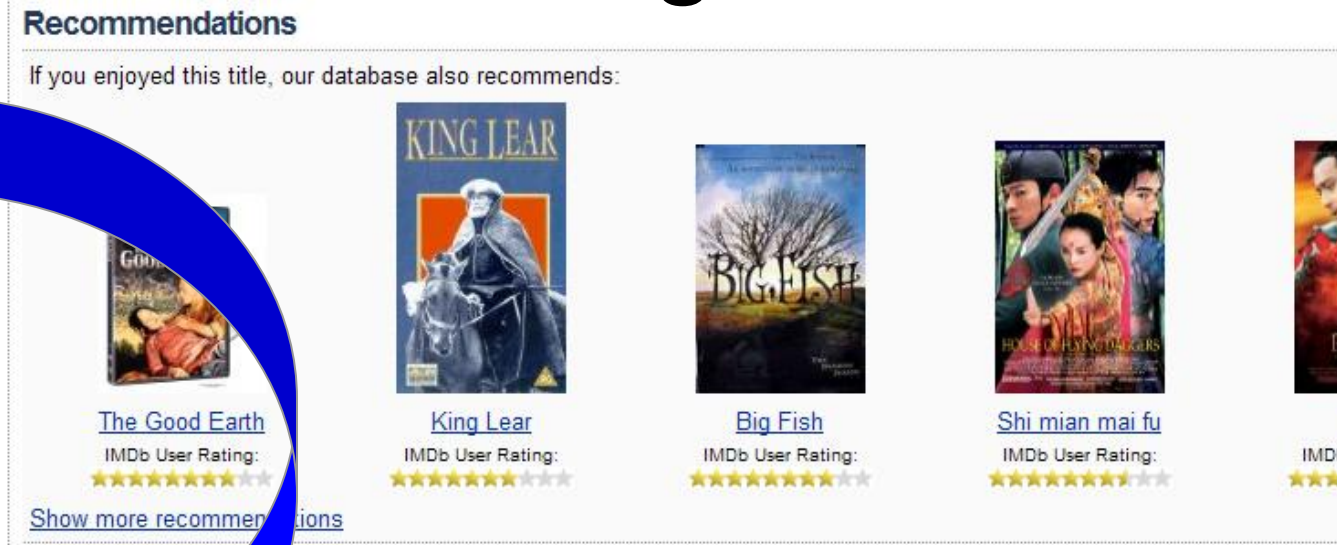
Data Sparsity in Collaborative Filtering



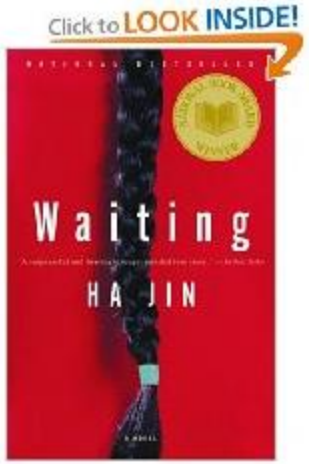
Transfer Learning for Collaborative Filtering?

IMDB Database





Amazon.com



29

Codebook Transfer

- Bin Li, Qiang Yang, Xiangyang Xue.
- Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction.
- In *Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence* (**IJCAI '09**), Pasadena, CA, USA, July 11-17, 2009.

Codebook Construction

- **Definition 2.1** (Codebook). A $k \times l$ matrix which compresses the cluster-level rating patterns of k user clusters and l item clusters.

	a	b	c	d	e	f
1	2	3	3	3	2	?
2	3	1	2	2	?	1
3	3	1	?	2	3	1
4	?	2	1	1	3	2
5	2	3	3	3	2	3
6	3	2	1	?	3	2

	a	e	b	f	c	d
2	3	?	1	1	2	2
3	3	3	1	1	?	2
1	2	2	3	?	3	3
5	2	2	3	3	3	3
4	?	3	2	2	1	1
6	3	3	2	2	1	?

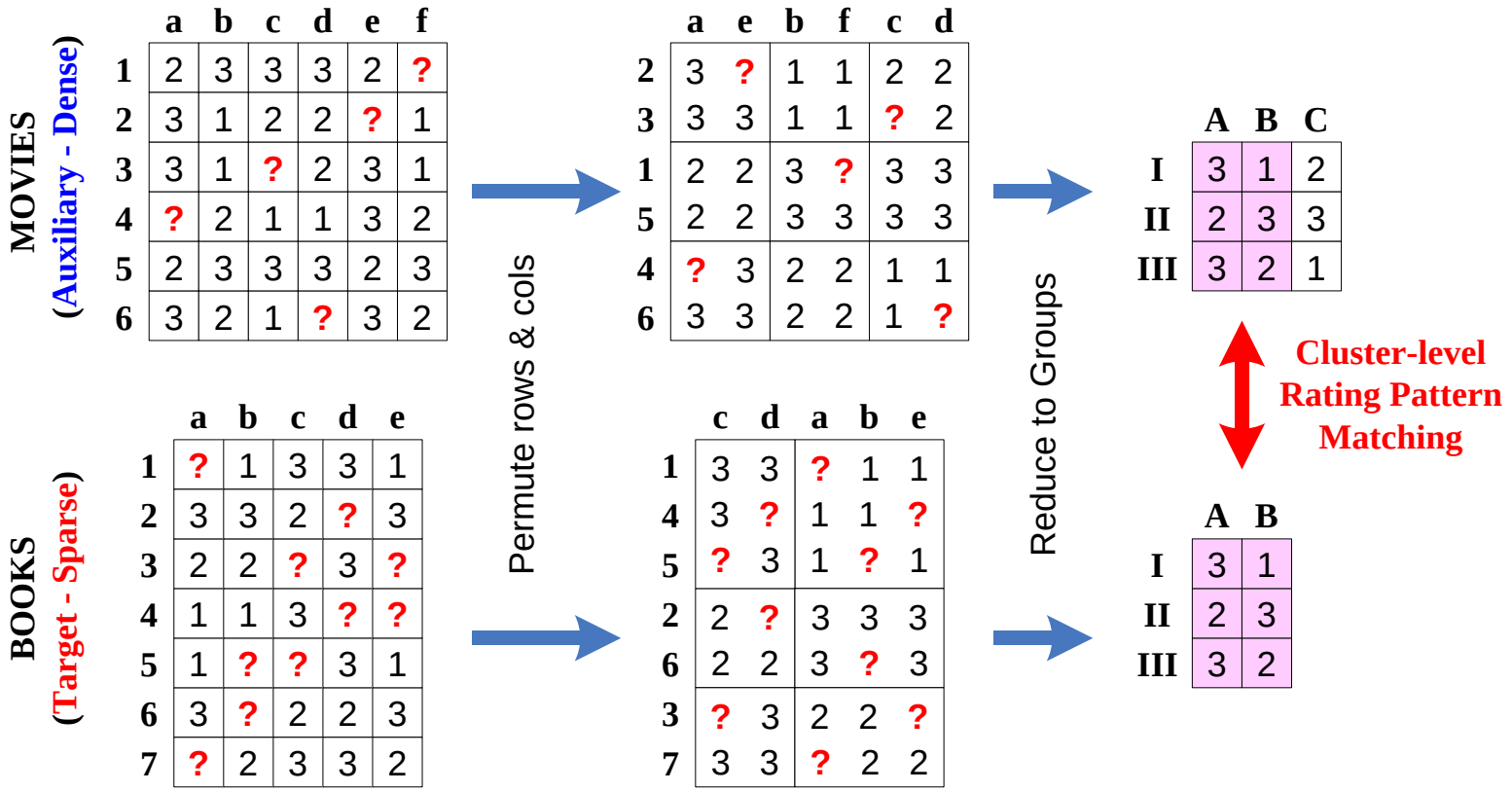
	A	B	C
I	3	1	2
II	2	3	3
III	3	2	1

Codebook

- Codebook: User prototypes rate on item prototypes
- Encoding: Find prototypes for users and items and get indices
- Decoding: Recover rating matrix based on codebook and indices

Knowledge Sharing via Cluster-Level Rating Matrix

- Source (Dense): Encode cluster-level rating patterns
- Target (Sparse): Map users/items to the encoded prototypes



Step 1: Codebook Construction

- Co-cluster rows (users) and columns (items) in $\mathbf{X}_{aux}$
- Get user/item cluster indicators $\mathbf{U}_{aux} \in \{0, 1\}^{n \times k}$, $\mathbf{V}_{aux} \in \{0, 1\}^{m \times l}$

$$\mathbf{B} = [\mathbf{U}_{aux}^{\top} \mathbf{X}_{aux} \mathbf{V}_{aux}] \oslash [\mathbf{U}_{aux}^{\top} \mathbf{1}\mathbf{1}^{\top} \mathbf{V}_{aux}]$$

	a	e	b	f	c	d
2	3	?	1	1	2	2
3	3	3	1	1	?	2
1	2	2	3	?	3	3
5	2	2	3	3	3	3
4	?	3	2	2	1	1
6	3	3	2	2	1	?

(3+3+3)/3 = 3

(1+1+1+1)/4 = 1

(2+2+2)/3 = 2

……

	A	B	C
I	3	1	2
II	2	3	3
III	3	2	1

Codebook B

Step 2: Codebook Transfer

- Objective

  Expand target matrix, while minimizing the difference between $\mathbf{X}_{tgt}$ and the reconstructed one

$$\min_{\substack{\mathbf{U}_{tgt}\in\{0,1\}^{p\times k}\\ \mathbf{V}_{tgt}\in\{0,1\}^{q\times l}}} \left\| [\mathbf{X}_{tgt} - \mathbf{U}_{tgt}\mathbf{B}\mathbf{V}_{tgt}^{\top}] \circ \mathbf{W} \right\|_F^2$$

$$\text{s.t.} \quad \mathbf{U}_{tgt}\mathbf{1} = \mathbf{1}, \mathbf{V}_{tgt}\mathbf{1} = \mathbf{1}$$

- User/item cluster indicators $\mathbf{U}_{tgt}$ and $\mathbf{V}_{tgt}$ for $\mathbf{X}_{tgt}$
- Binary weighting matrix $\mathbf{W}$ for observed ratings in $\mathbf{X}_{tgt}$
- Alternate greedy searches for $\mathbf{U}_{tgt}$ and $\mathbf{V}_{tgt}$ to a local minimum

Experimental Setup

- Data Sets
  - EachMovie (Auxiliary): 500 users × 500 movies
  - MovieLens (Target): 500 users × 1000 movies
  - Book-Crossing (Target): 500 users × 1000 books

- Compared Methods
  - Pearson Correlation Coefficients (PCC)
  - Scalable Cluster-based Smoothing (CBS)
  - Weighted Low-rank Approximation (WLR)
  - **Codebook Transfer (CBT)**

- Evaluation Protocol
  - First 100/200/300 users for training; last 200 users for testing
  - Given 5/10/15 observable ratings for each test user

Experimental Results (1): Books → Movies

- MAE Comparison on MovieLens
  - average over 10 sampled test sets
  - Lower is better

Training Set	Method	Given5	Given10	Given15
ML100	PCC	0.930	0.883	0.873
	CBS	0.874	0.845	0.839
	WLR	0.915	0.875	0.890
	CBT	**0.840**	**0.802**	**0.786**
ML200	PCC	0.905	0.878	0.878
	CBS	0.871	0.833	0.828
	WLR	0.941	0.903	0.883
	CBT	**0.839**	**0.800**	**0.784**
ML300	PCC	0.897	0.882	0.885
	CBS	0.870	0.834	0.819
	WLR	1.018	0.962	0.938
	CBT	**0.840**	**0.801**	**0.785**

Limitations of Codebook Transfer

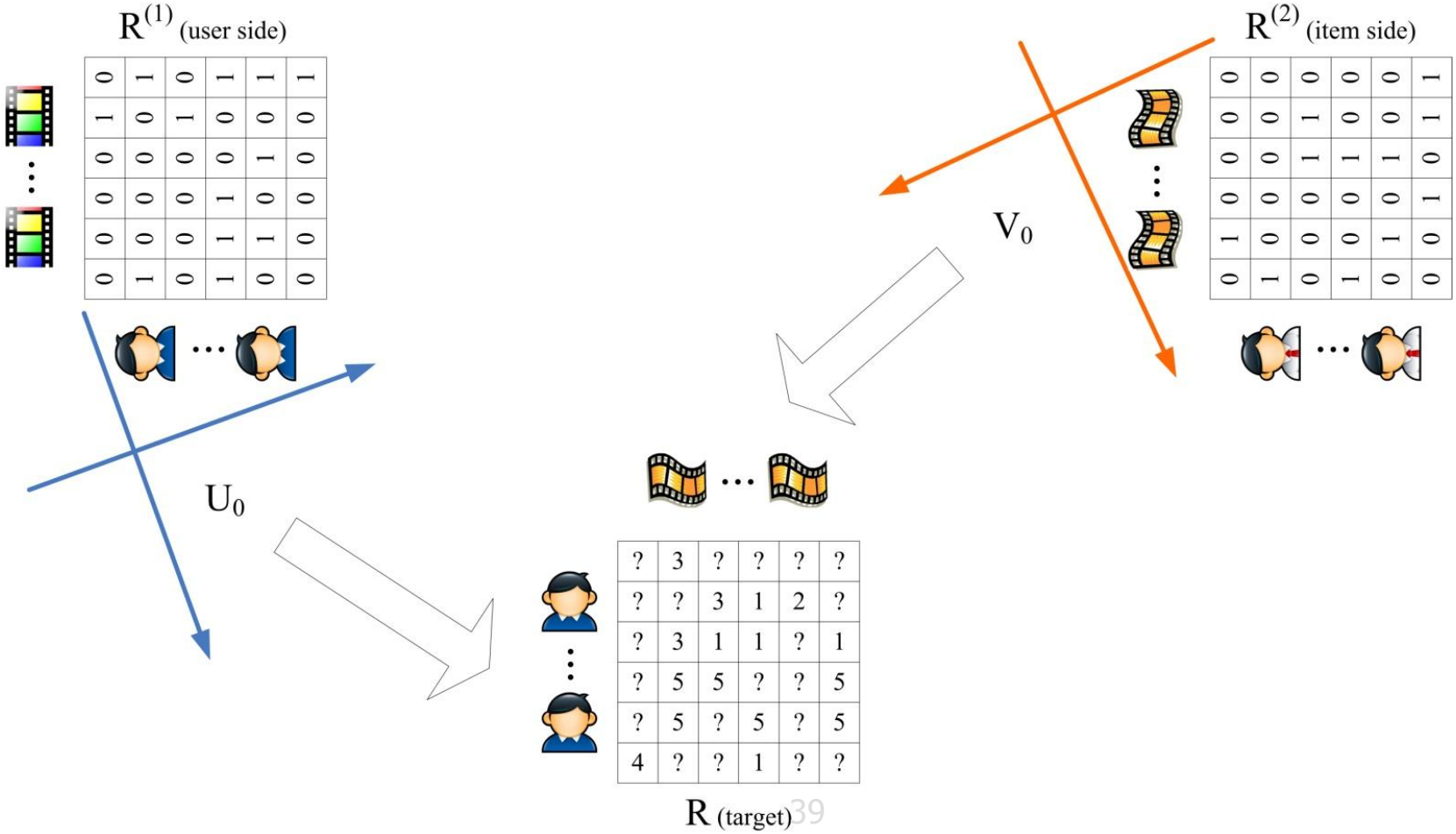
- Same rating range
  - Source and target data must have the same range of ratings [1, 5]
- In reality
  - Range of ratings can be 0/1 or [1,5]
  - Only works well when target domain is extremely sparse

Coordinate System Transfer

- Weike Pan, Evan Xiang, Nathan Liu and Qiang Yang.
- Transfer Learning in Collaborative Filtering for Sparsity Reduction.
- In Proceedings of the 24th AAAI Conference on Artificial Intelligence (**AAAI-10**). Atlanta, Georgia, USA. July 11-15, 2010.

Our Solution: Coordinate System Transfer

- Step 1: Coordinate System Construction ($\mathbf{U}_0, \mathbf{V}_0$)
- Step 2: Coordinate System Adaptation



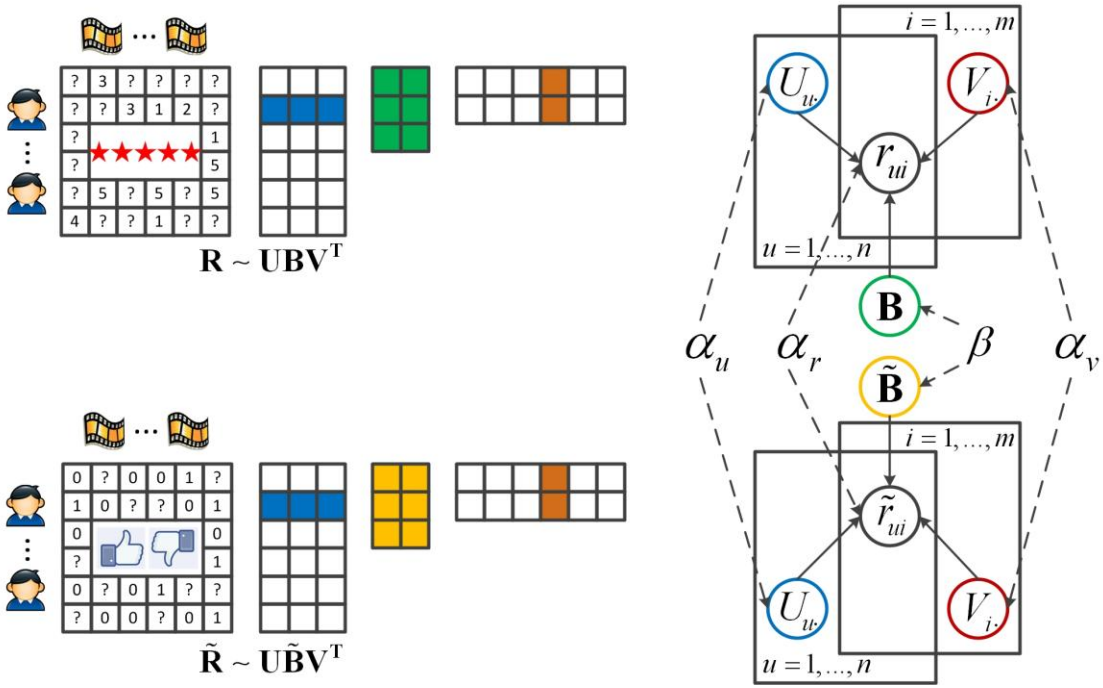
IJCAI 2011 Talk (Tuesday, 10:30am)
Transfer Learning to Predict Missing Ratings via Heterogeneous User Feedbacks

Weike Pan, Nathan N. Liu, Evan W. Xiang, Qiang Yang
{weikep, nliu, wxiang, qyang}@cse.ust.hk

Department of Computer Science and Engineering
Hong Kong University of Science and Technology
Hong Kong, China

IJCAI-11, Barcelona, Catalonia, Spain. July 16-22, 2011

When both users and items align



$$\min_{\mathbf{U},\mathbf{V},\mathbf{B},\tilde{\mathbf{B}}} \quad \mathcal{F}(\mathbf{R} \sim \mathbf{U}\mathbf{B}\mathbf{V}^T) + \lambda \mathcal{F}(\tilde{\mathbf{R}} \sim \mathbf{U}\tilde{\mathbf{B}}\mathbf{V}^T)$$
$$\text{s.t.} \quad \mathbf{U}, \mathbf{V} \in \mathfrak{D}$$

$\mathfrak{D}$: the range domain of user-specific and item-specific feature matrices

$\mathfrak{D}_{\mathbb{R}} = \{\mathbf{U} \in \mathbb{R}^{n \times d}, \mathbf{V} \in \mathbb{R}^{m \times d}\}$ CMTF (collective matrix tri-factorization)

$\mathfrak{D}_{\perp} = \mathfrak{D}_{\mathbb{R}} \cap \{\mathbf{U}^T\mathbf{U} = \mathbf{I}, \mathbf{V}^T\mathbf{V} = \mathbf{I}\}$ CSVD (collective SVD)

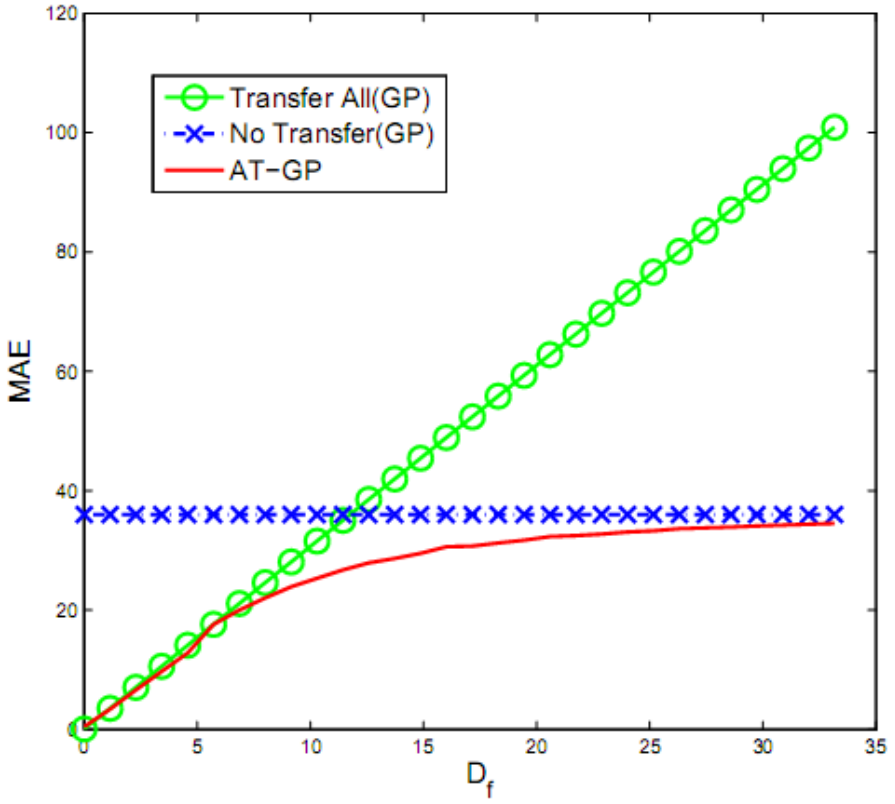
Limitation of CST and CBT

- Different source domains are related to the target domain in the same way.
- In reality,
  - Book to Movies: related
  - Food to Movies: not related
- Rating bias
  - Users tend to rate items that they like
    - Thus there are more rating = 5 than rating = 2

Adaptive Transfer Learning

- Bin Cao, Sinno Jialin Pan, Yu Zhang, Dit-Yan Yeung and Qiang Yang. Adaptive Transfer Learning. In Proceedings of the 24th AAAI Conference on Artificial Intelligence (**AAAI-10**). Atlanta, Georgia, USA. July 11-15, 2010.

Adaptive: transfer-all and transfer-none



Distance between the source and target tasks

- As good as Transfer All when the source and target tasks are very similar.
- Not worse than No Transfer when the source and target tasks are not related at all.

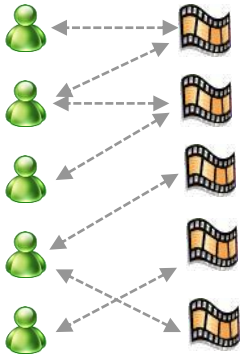
Social Media (Wiki) as Source Data

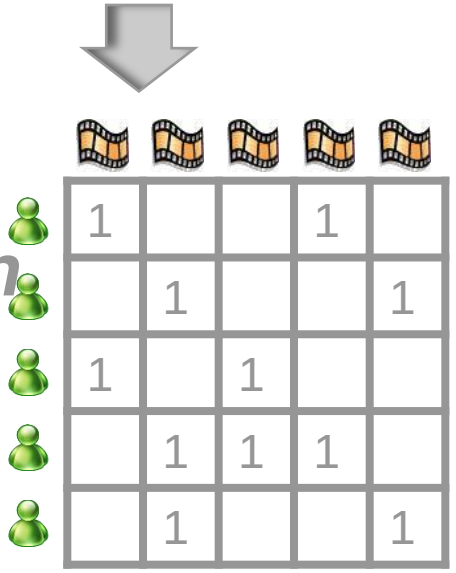
Social-behavior Transfer Learning for Recommendation Systems

Qian Xu, Evan Wei Xiang and Qiang Yang
Hong Kong University of Science and Technology
Hong Kong, China

Wikipedia as the source

- *Wikipedia is a mirror of Web*
- *Wikipedia can cover a large set of items*
- *Wikipedia editing log contains rich social preference information*





WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox

User contributions

From Wikipedia, the free encyclopedia

Help:User contributions

For 98.211.116.174 (talk | block log | uploads | logs | filter log)

Search for contributions

Show contributions of new accounts only

User: 98.211.116.174 Namespace: all Deleted o

Only show edits that are latest revisions

Tag filter:

From year (and earlier): From month (and earlier): all

- 23:53, 27 June 2011 (diff | hist) Madagascar (franchise) *(→Penguins film (TBA))* *(Tag: references removed)*
- 00:29, 25 June 2011 (diff | hist) List of Big Time Rush episodes *(→Season 2 (2010–2011))*
- 00:26, 25 June 2011 (diff | hist) List of Big Time Rush episodes *(→Season 2 (2010–2011))*
- 18:06, 14 June 2011 (diff | hist) Kung Fu Panda *(→Cast)*
- 17:56, 14 June 2011 (diff | hist) Kung Fu Panda *(→Plot)*
- 22:02, 8 May 2011 (diff | hist) Kung Fu Panda (film series) *(→Future)*
- 02:51, 28 April 2011 (diff | hist) The Fast and the Furious (film series) *(→Future)*
- 02:40, 20 April 2011 (diff | hist) Sonny with a Chance *(→Main characters)*
- 02:35, 20 April 2011 (diff | hist) List of Victorious episodes *(→References)* *(Tag: reference list removal)*
- 02:30, 20 April 2011 (diff | hist) List of Victorious episodes *(→Season 2: 2011)*
- 02:26, 20 April 2011 (diff | hist) Victorious *(→Characters)*
- 02:21, 20 April 2011 (diff | hist) Rio (film) *(→Plot)*

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox

User contributions

From Wikipedia, the free encyclopedia
Help:User contributions
For 98.211.116.174 (talk | block log | uploads | logs | filter log)

Search for contributions

Show contributions of new accounts only
User: 98.211.116.174 Namespace: all Deleted o
Only show edits that are latest revisions
Tag filter:
From year (and earlier): From month (and earlier): all

- 23:53, 27 June 2011 (diff | hist) Madagascar (franchise) (→*Penguins film (TBA)*) (*Tag: references removed*)
- 00:29, 25 June 2011 (diff | hist) List of Big Time Rush episodes (→*Season 2 (2010–2011)*)
- 00:26, 25 June 2011 (diff | hist) List of Big Time Rush episodes (→*Season 2 (2010–2011)*)
- 18:06, 14 June 2011 (diff | hist) Kung Fu Panda (→*Cast*)
- 17:56, 14 June 2011 (diff | hist) Kung Fu Panda (→*Plot*)
- 22:02, 8 May 2011 (diff | hist) Kung Fu Panda (film series) (→*Future*)
- 02:51, 28 April 2011 (diff | hist) The Fast and the Furious (film series) (→*Future*)
- 02:40, 20 April 2011 (diff | hist) Sonny with a Chance (→*Main characters*)
- 02:35, 20 April 2011 (diff | hist) List of Victorious episodes (→*References*) (*Tag: reference list removal*)
- 02:30, 20 April 2011 (diff | hist) List of Victorious episodes (→*Season 2: 2011*)
- 02:26, 20 April 2011 (diff | hist) Victorious (→*Characters*)
- 02:21, 20 April 2011 (diff | hist) Rio (film) (→*Plot*)

WIKIPEDIA
opedia

Article Discussion Read Edit Search

Kung Fu Panda

From Wikipedia, the free encyclopedia

Kung Fu Panda is a 2008 American computer-animated action comedy film produced by DreamWorks Animation and distributed by Paramount Pictures. It was directed by John Wayne Stevenson and Mark Osborne and produced by Melissa Cobb, and stars the voice of Jack Black along with Dustin Hoffman, Jackie Chan, Angelina Jolie, Ian McShane, Seth Rogen, Lucy Liu, David Cross, Randall Duk Kim, James Hong, Dan Fogler and Michael Clarke Duncan. Set in a version of old China populated by humanoid talking animals, the plot revolves around a bumbling humanoid panda named Po who aspires to be a kung fu master. When an evil humanoid kung fu warrior is foretold to escape from prison, Po is unwittingly named the chosen one destined to bring peace to the land, much to the chagrin of

Kung Fu Panda

Theatrical poster

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox

User contributions

From Wikipedia, the free encyclopedia
Help:User contributions
For 98.211.116.174 (talk | block log | uploads | logs | filter log)

Search for contributions

Show contributions of new accounts only
User: 98.211.116.174 Namespace: all Deleted o
Only show edits that are latest revisions
Tag filter:
From year (and earlier): From month (and earlier): all

- 23:53, 27 June 2011 (diff | hist) Madagascar (franchise) (→*Penguins film (TBA)*) (*Tag: references removed*)
- 00:29, 25 June 2011 (diff | hist) List of Big Time Rush episodes (→*Season 2 (2010–2011)*)
- 00:26, 25 June 2011 (diff | hist) List of Big Time Rush episodes (→*Season 2 (2010–2011)*)
- 18:06, 14 June 2011 (diff | hist) Kung Fu Panda (→*Cast*)
- 17:56, 14 June 2011 (diff | hist) Kung Fu Panda (→*Plot*)
- 22:02, 8 May 2011 (diff | hist) Kung Fu Panda (film series) (→*Future*)
- 02:51, 28 April 2011 (diff | hist) The Fast and the Furious (film series) (→*Future*)
- 02:40, 20 April 2011 (diff | hist) Sonny with a Chance (→*Main characters*)
- 02:35, 20 April 2011 (diff | hist) List of Victorious episodes (→*References*) (*Tag: reference list removal*)
- 02:30, 20 April 2011 (diff | hist) List of Victorious episodes (→*Season 2: 2011*)
- 02:26, 20 April 2011 (diff | hist) Victorious (→*Characters*)
- 02:21, 20 April 2011 (diff | hist) Rio (film) (→*Plot*)

Article Discussion Read Edit Search

Kung Fu Panda

From Wikipedia, the free encyclopedia

Kung Fu Panda is a 2008 American computer-animated action comedy film produced by DreamWorks Animation and distributed by Paramount Pictures. It was directed by John Wayne Stevenson and Mark Osborne and produced by Melissa Cobb, and stars the voice of Jack Black along with Dustin Hoffman, Jackie Chan, Angelina Jolie, Ian McShane, Seth Rogen, Lucy Liu, David Cross, Randall Duk Kim, James Hong, Dan Fogler and Michael

Kung Fu Panda

Article Discussion d Edit Search

Rio (film)

From Wikipedia, the free encyclopedia

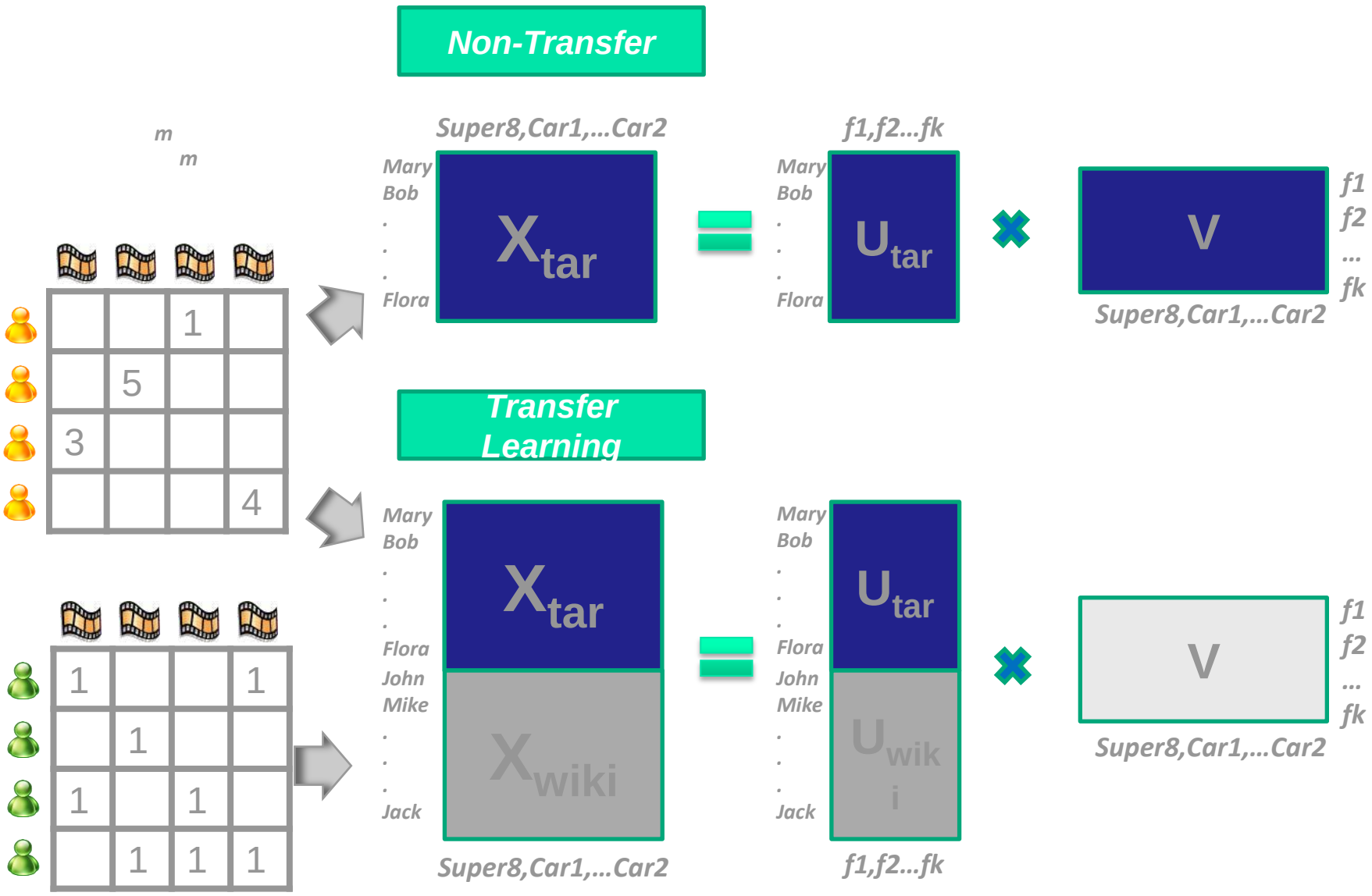
Rio, often promoted as ***Rio: The Movie***, is a 2011 American 3D computer-animated musical comedy film produced by Blue Sky Studios and directed by Carlos Saldanha. The title refers to the Brazilian city of Rio de Janeiro,[5] in which the film is set. The film features the voices of Jesse Eisenberg, Anne Hathaway, George Lopez, Jemaine Clement and Jake T. Austin.[6] The film tells the story of Blu (Eisenberg), a blue macaw who is taken to Rio de Janeiro to mate with a female. He eventually falls in love with Jewel (Hathaway), a

Rio

Theatrical release poster

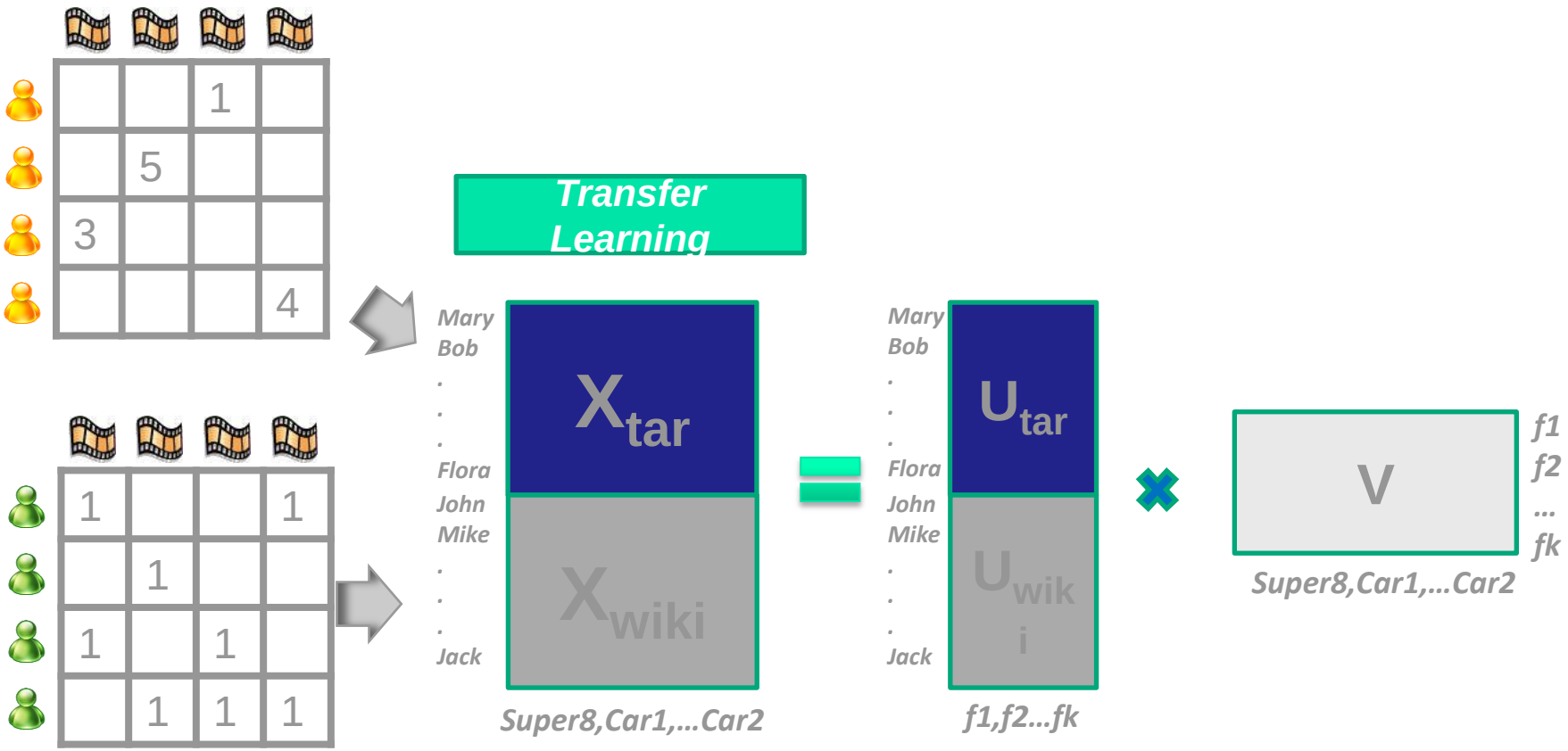
Languages
Български
Dansk

Transfer Learning via COEDIT



Transfer Learning via COEDIT

$$\mathcal{L} = \sum_{s \in \{S^m_{wiki}, S_{tar}\}} (\alpha_s \sum_{(i,j) \in \mathcal{R}_s} \left(\mathbf{u}_i^T \mathbf{v}_j - X_{s,ij}\right)^2) + \lambda(\| \mathbf{U} \|_F^2 + \| \mathbf{V} \|_F^2),$$



Experiments

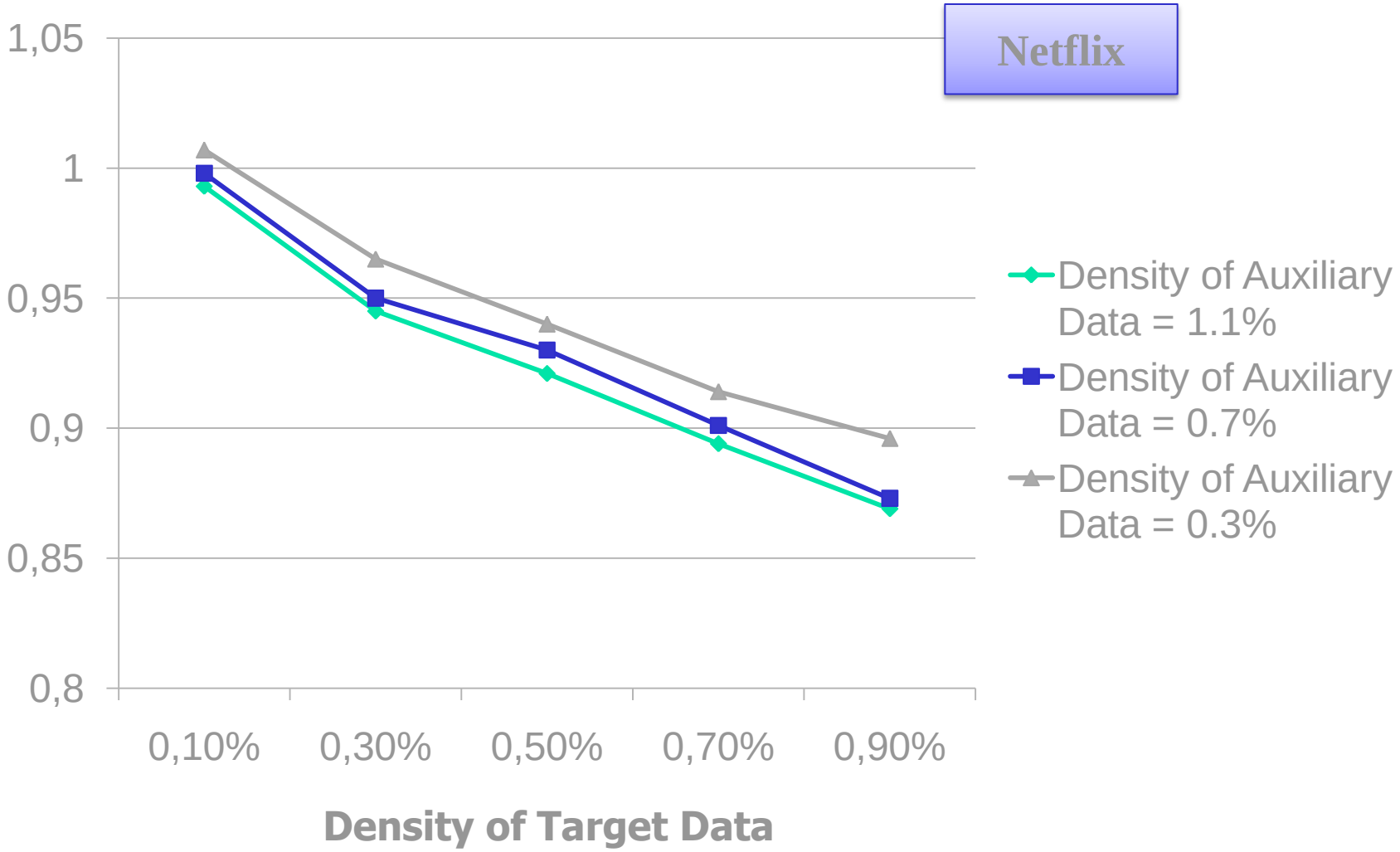
- Data
  - Netflix: 100 million ratings, 480 thousand users, 17,770 movies
  - MovieLens: 10 million ratings, 71,000 users, 10,681 movies
  - Wikipedia: align 11,000 movies to the articles
- Effectiveness test
  - 70% training, 30% testing
  - Density: 0.1%-0.9%
  - k={3,5,10,15,20}, λ={1,2,5,10,20}
  - Effectiveness on COEDIT for knowledge Transfer
  - Effectiveness on Wikipedia Data Selection
- Efficiency test
  - Parallel Learning for COEDIT

Can COEDIT Improve Recommendation Performance ?

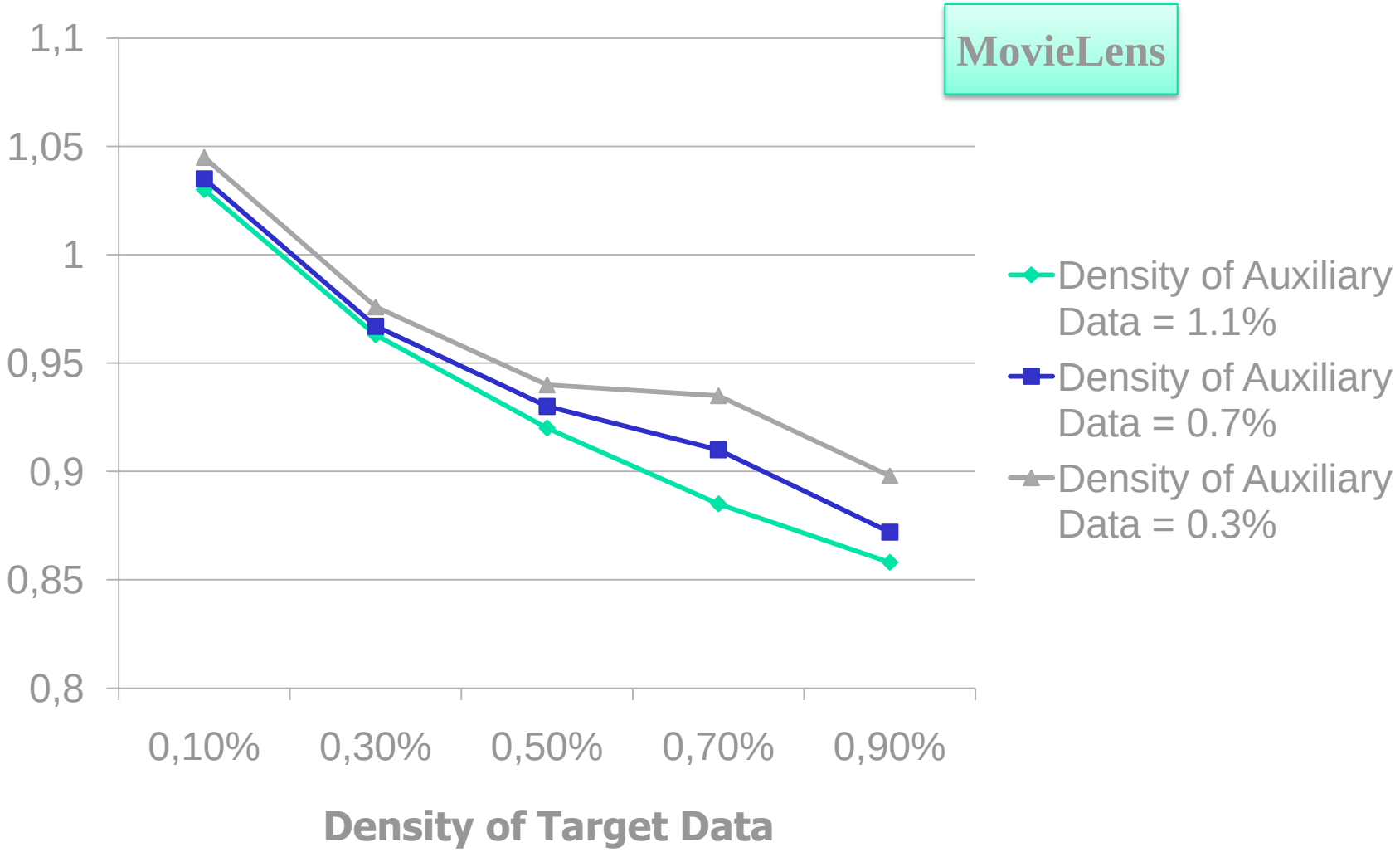
- Compare COEDIT with baseline methods
  - Without transfer
    - Average filling method (AF)
    - Latent factorization model (LFM)
  - With transfer
    - $T_{content}$: movie-word matrix as X_{wiki}
    - T_{link}: movie-neighborhood matrix as X_{wiki}
- Target density = 0.9%

	Netflix	MovieLens
AF	0.918	0.912
LFM	0.900	0.894
$T_{content}$	0.899	0.890
T_{link}	0.891	0.888
COEDIT	0.869	0.858

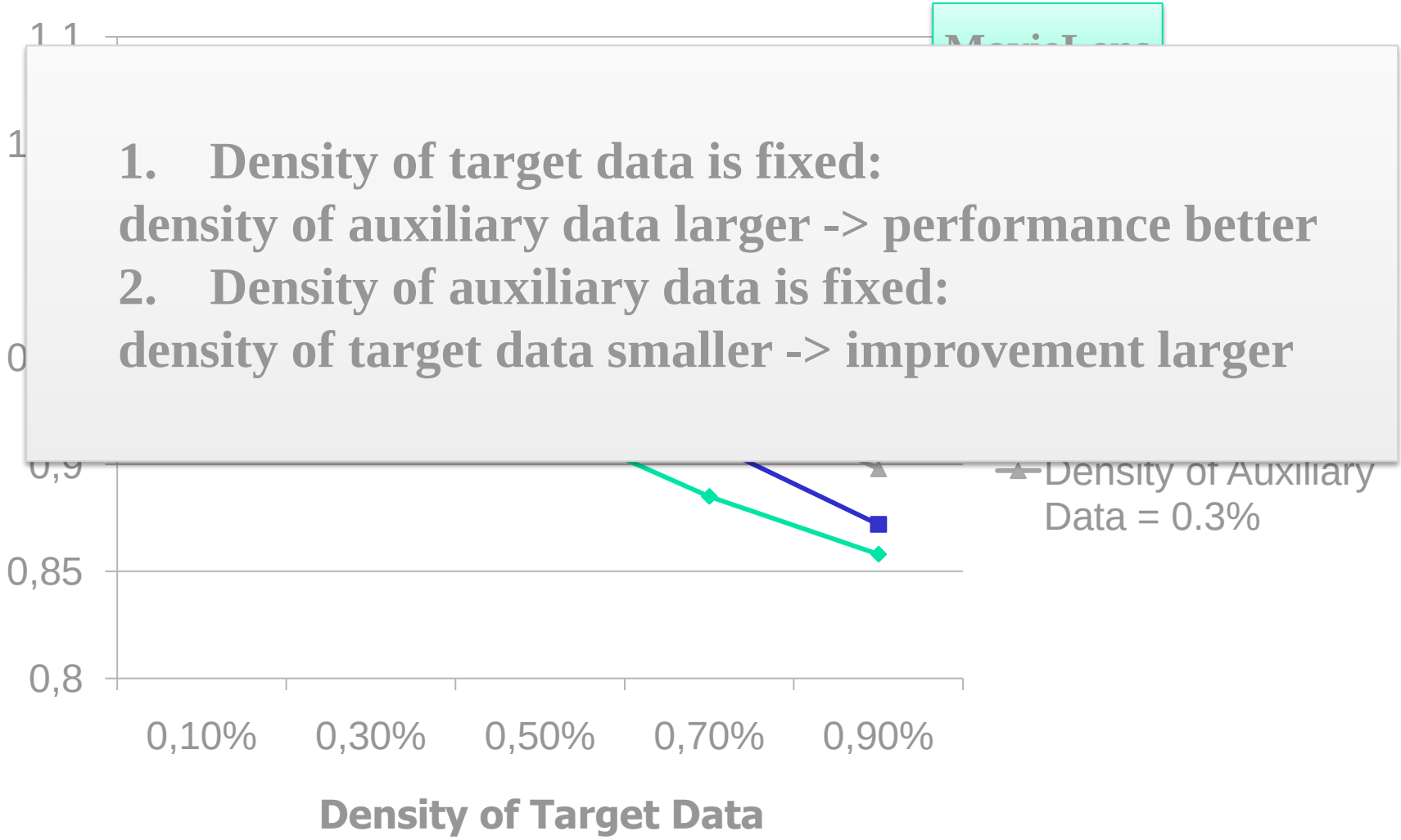
How Does the Density Affect Results?



How Does the Density Affect Results?

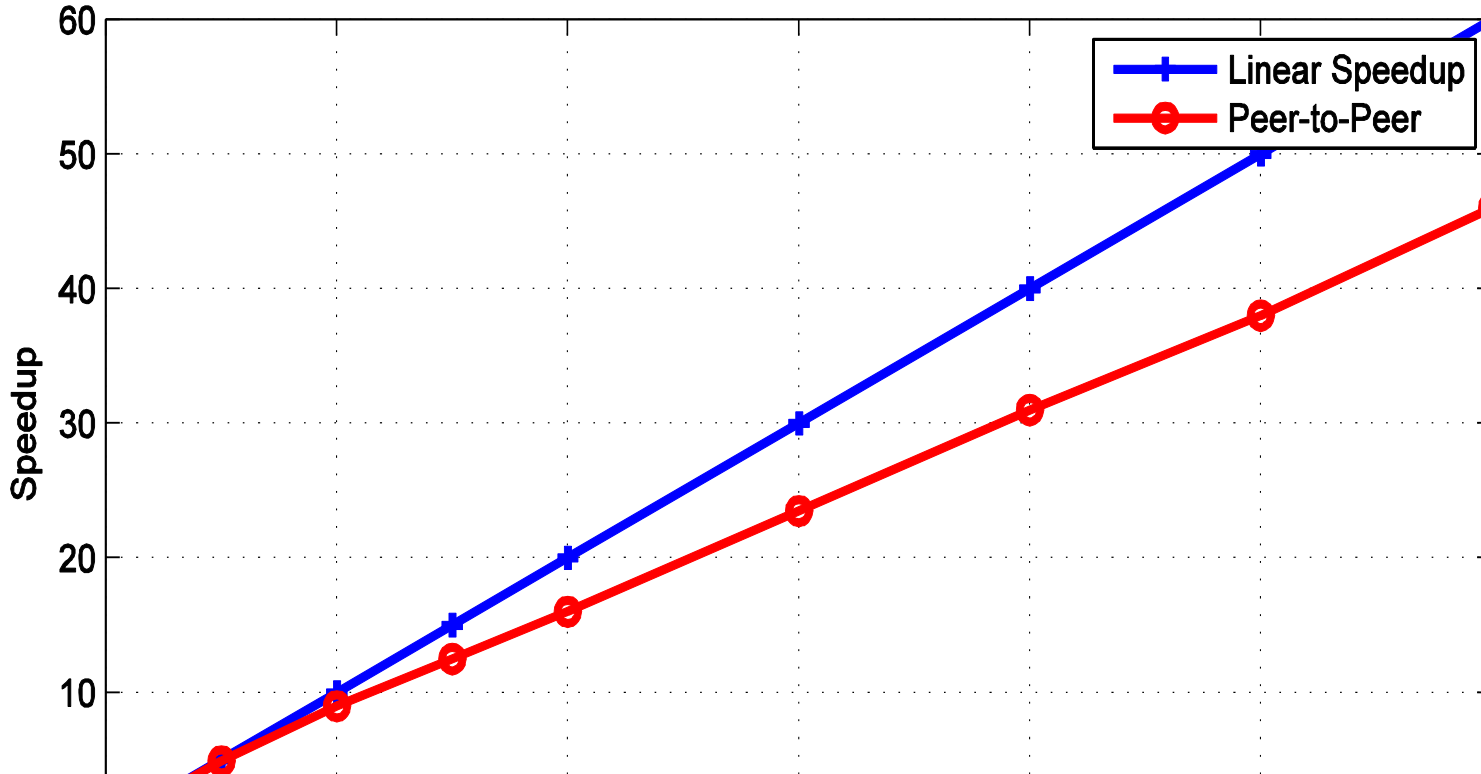


How Does the Density Affect Results?



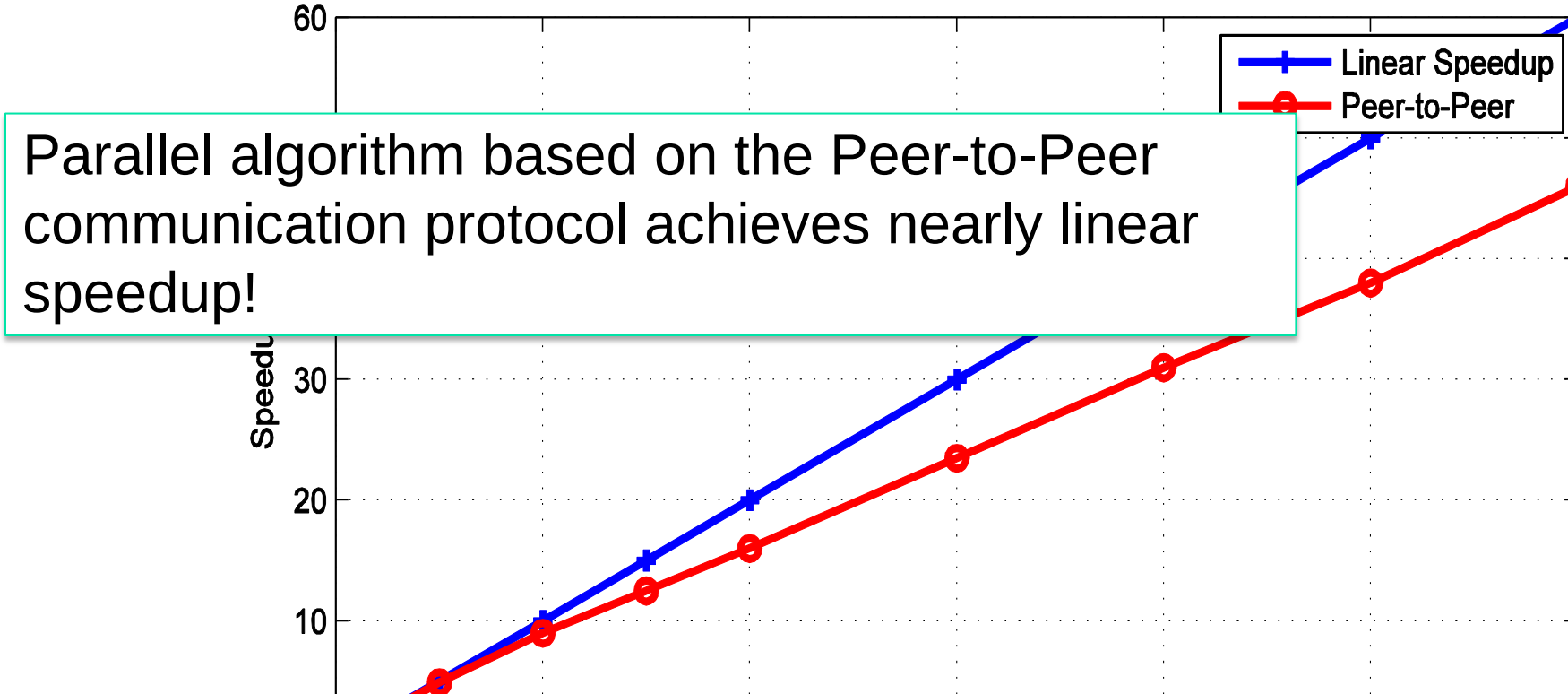
Efficiency Test

- 1 Gb/s LAN based cluster of 8 servers with Intel 8-core 2.93 GHz CPU and 24GB memory



Efficiency Test

- 1 Gb/s LAN based cluster of 8 servers with Intel 8-core 2.93 GHz CPU and 24GB memory



Parallel algorithm based on the Peer-to-Peer communication protocol achieves nearly linear speedup!

Summary

- COEDIT: transfers co-editing knowledge in Wikipedia to solve the data sparsity problem in collaborative filtering tasks.
  - Co-editing knowledge in Wikipedia can effectively help solve the data sparsity problem in other target domains
  - Parallel algorithm can be used to scale up the transfer learning efficiency

Conclusions and Future Work

- Transfer Learning with social media
  - Social media as translators
  - Transfer Learning for CF
  - Transferring knowledge from Wiki via COEDIT
- Challenges
  - How other social knowledge can be used to help with the tasks in other domains.
  - Investigate how to analyze the domain differences for source data selection.

Future Work

Acknowledgements

- HKUST:
  - Sinno J. Pan, Huayan Wang, Bin Cao, Evan Wei Xiang, Derek Hao Hu, Nathan Nan Liu, Qian Xu, Vincent Wenchen Zheng, Weike Pan, etc.
- Shanghai Jiaotong University:
  - Wenyuan Dai, Guirong Xue, Yuqiang Chen, Prof. Yong Yu, Xiao Ling, Ou Jin.
- Visiting Students
  - Bin Li (Fudan U.), Xiaoxiao Shi (Zhong Shan U.), etc.

TL Resources

- http://www.cse.ust.hk/TL