HOP-Rec

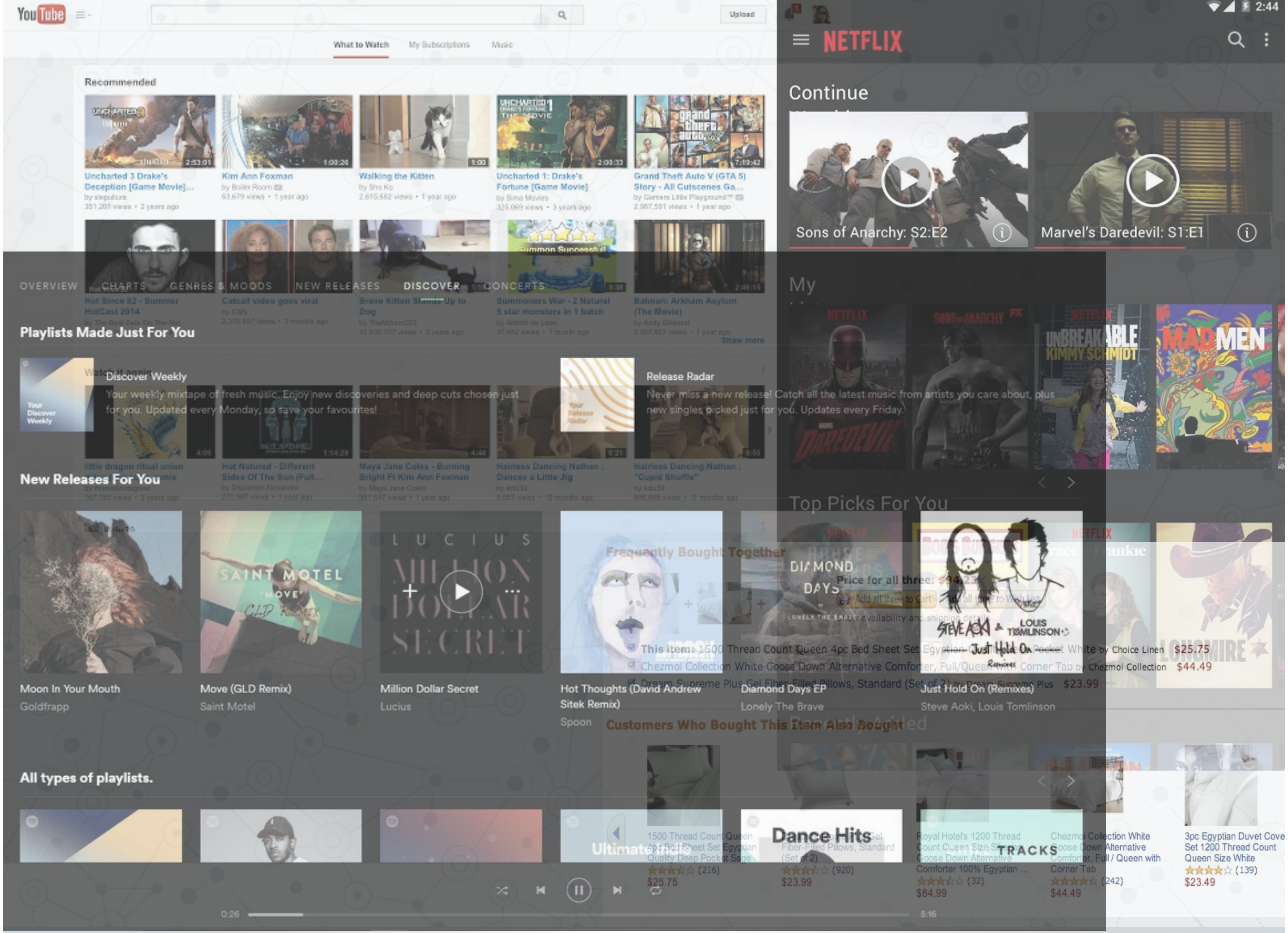
High-Order Proximity for Implicit Recommendation

Jheng-Hong Yang†
Chih-Ming Chen‡
Chuan-Ju Wang†
Ming Feng Tsai‡

†Academia Sinica, Taiwan
‡National Chengchi University, Taiwan

Rec*ommender* **Sys***tems are software tools and techniques providing suggestions for items* ***to be of use*** *to a user*

Francesco, R., R. Lior, and S. Bracha. "Introduction to Recommender Systems Handbook, RecommenderSystems Handbook." (2011).


What to Watch
My Subscriptions
Music
Recommended
NETFLIX
Continue
Sons of Anarchy: S2:E2
Marvel's Daredevil: S1:E1
Top Picks For You
OVERVIEW
CHARTS
GENRES & MOODS
NEW RELEASES
DISCOVER
CONCERTS
Playlists Made Just For You
Discover Weekly
Your weekly mixtape of fresh music. Enjoy new discoveries and deep cuts chosen just for you. Updated every Monday, so save your favourites!
Release Radar
Never miss a new release! Catch all the latest music from artists you care about, plus new singles picked just for you. Updates every Friday.
New Releases For You
Moon In Your Mouth
Goldfrapp
Move (GLD Remix)
Saint Motel
Million Dollar Secret
Lucius
Hot Thoughts (David Andrew Sitek Remix)
Spoon
Diamond Days EP
Lonely The Brave
Just Hold On (Remixes)
Steve Aoki, Louis Tomlinson
Frequently Bought Together
Customers Who Bought This Item Also Bought
All types of playlists.
Dance Hits

Collaborative Filtering

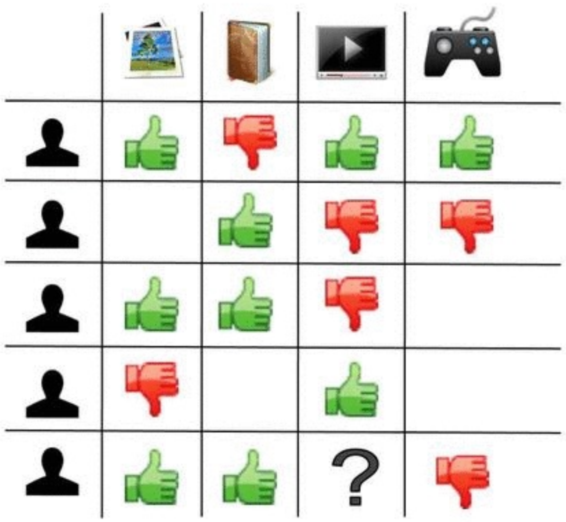
Make **predictions** by collecting preference **information from users**

Figure: https://www.youtube.com/watch?v=FSKIoPyIEkM

Collaborative Filtering

Latent Factor Model

Discover **shared latent factors** between users and items by decomposing user-item interaction matrix



Graph-based Model

Explore the **relationships** between users and items within a **graph structure**

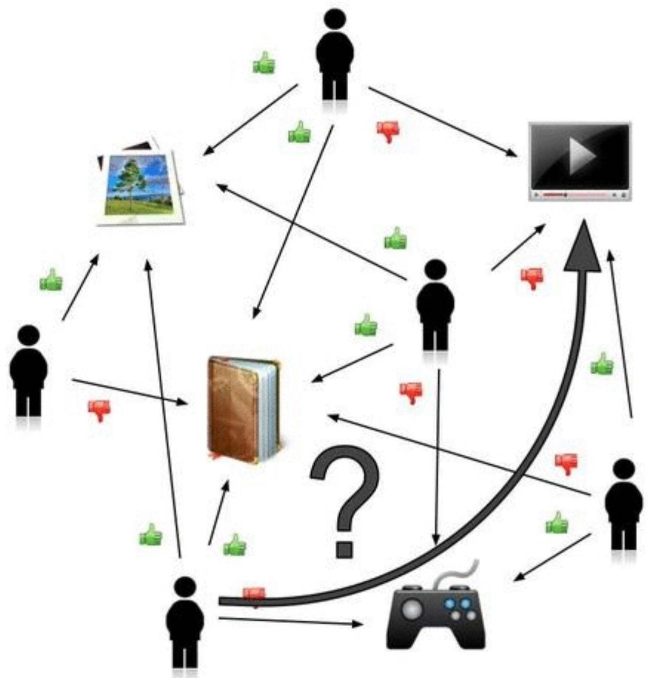


Figure: https://en.wikipedia.org/wiki/Collaborative_filtering

Related Works

Latent Factor Model

- Matrix Factorization (MF)
- Bayesian Personalized Ranking (BPR)
- Weighted Approximate-Rank Pairwise (WARP) loss
- K-Order Statistic (K-OS) loss

Pointwise

$$
\begin{array}{c} \\ \mathbf{U_1} \\ \\ \\ \mathbf{U_m} \end{array}
\begin{array}{c}
\begin{array}{cccc} \phantom{e_{11}} & \mathbf{I_2} & & \mathbf{I_n} \end{array} \\
\begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1n} \\ e_{21} & e_{22} & & e_{2n} \\ \vdots & & & \vdots \\ e_{m1} & e_{m2} & \cdots & e_{mn} \end{bmatrix}
\end{array}
\approx \begin{bmatrix} \\ U \\ \\ \end{bmatrix} \begin{bmatrix} \quad & I & \quad \end{bmatrix}
$$

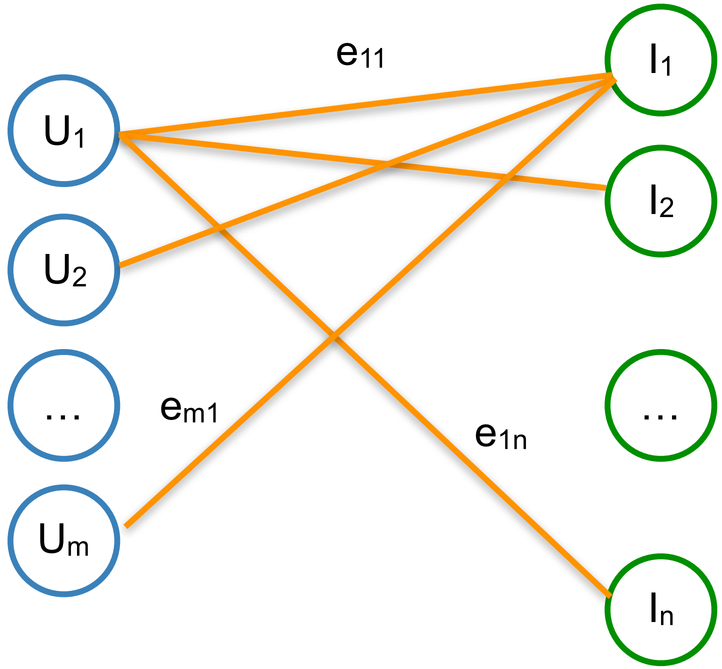
Pairwise

User-Item interaction matrix   User latent factors   Item latent factors

Related Works

Graph-based Model

- PageRank
- ItemRank
- S-step random walk distribution (hitting time)
- Popularity-based Re-ranking (RP³(β)) (transition probability)



Motivation

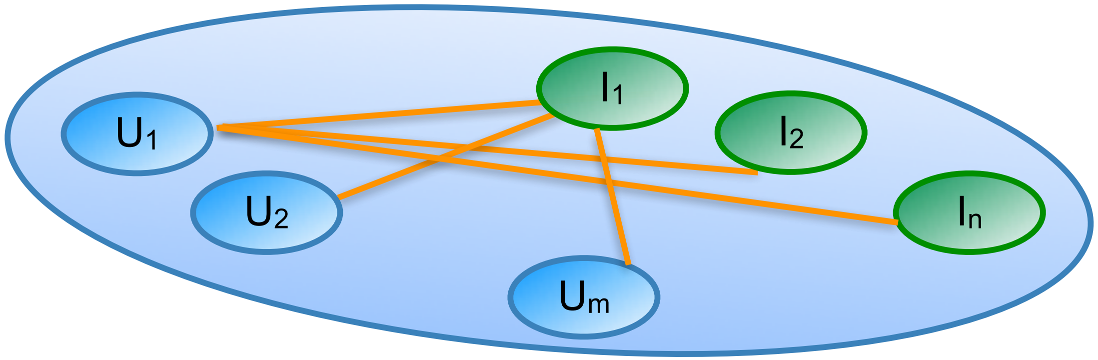
Motivation: Problem

Latent Factor Model

- Only discriminate shallow observations within user-item interaction

Graph-based Model

- Explore higher-order proximities within graph, but unreached item will not be affect remotely

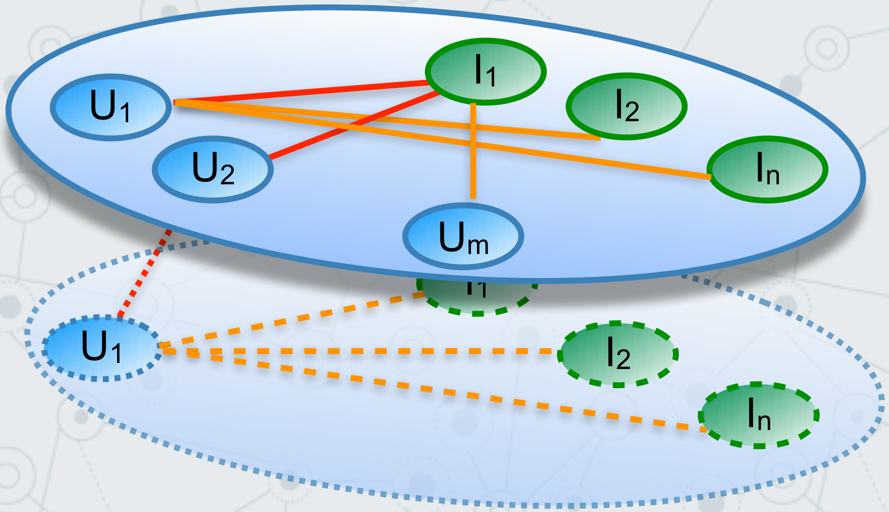


Park, Haekyu, Jinhong Jung, and U. Kang. "A comparative study of matrix factorization and random walk with restart in recommender systems." *Big Data (Big Data), 2017 IEEE International Conference on*. IEEE, 2017.

Motivation: Hybrid

HOP-Rec = Latent Factor Model + Graph-based model

- **Order-aware objectives**
- **Latent factor based**



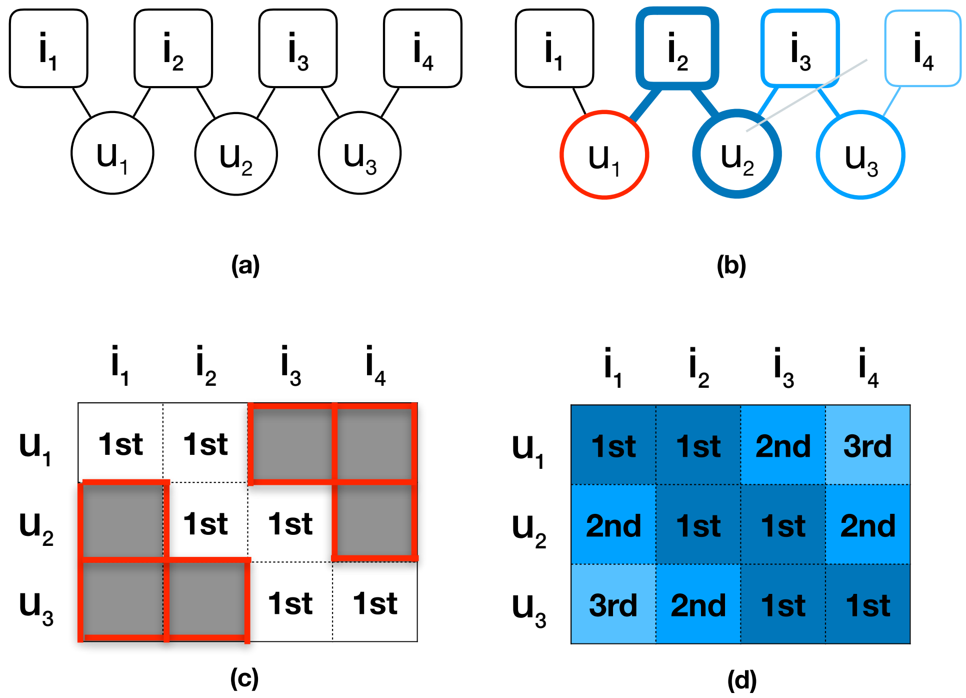
Order = 1

Order = 2 for U_2 through U_1, because they share the same taste on I_1

Motivation: Assumption

Assumptions:

- The preference of unknown items could be **estimated from indirect observations.**
- The estimation should be **ordered by the neighborhood proximities.**



Model: Methodology

Objective function

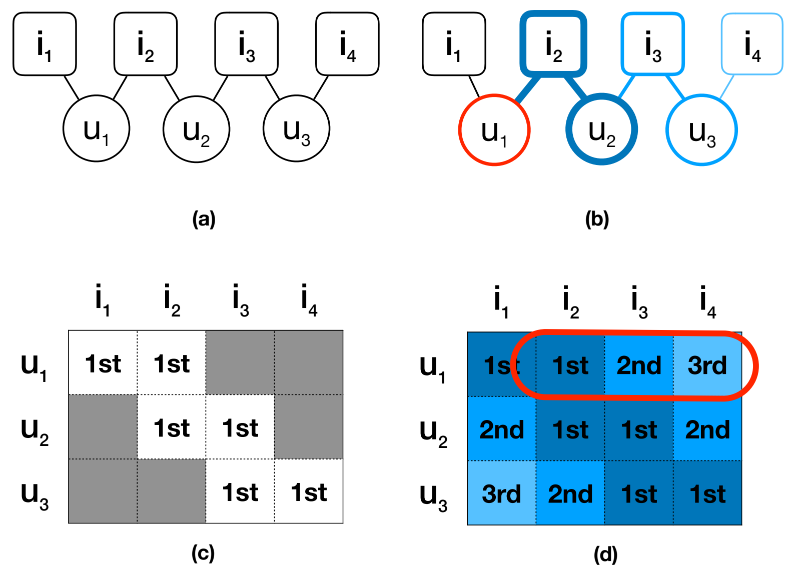
- **Hybrid**

$$\mathcal{L}_{HOP} = \sum_{\substack{1\leq k\leq K \\ u,(i,i')}} \overbrace{\mathcal{C}(k)\,\mathbb{E}_{\substack{i\sim P_u^k \\ i'\sim P_N}}}^{\text{graph model}} \overbrace{\left[\mathcal{F}\left(\theta_u^{\intercal}\theta_{i'}, \theta_u^{\intercal}\theta_i\right)\right]}^{\text{factorization model}} + \lambda_{\Theta}\left\|\Theta\right\|_2^2$$

- **Pairwise**

$$\mathcal{F}(\theta_u^{\intercal}\theta_{i'}, \theta_u^{\intercal}\theta_i) = \mathbb{1}_{\{\theta_u^{\intercal}\theta_{i'} - \theta_u^{\intercal}\theta_i > \epsilon_k\}} \log\left[\sigma\left(\theta_u^{\intercal}\theta_{i'} - \theta_u^{\intercal}\theta_i\right)\right]$$

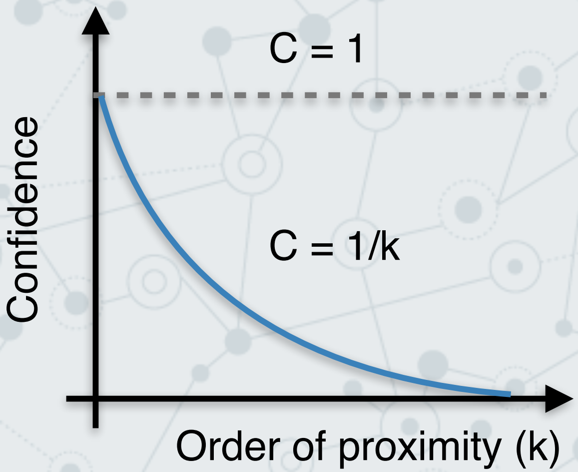
- **Optimization Flow**
  - ➡ Sample path on graph
  - ➡ Optimize pairwise relationship between positive and negative items of different orders



Three Stories

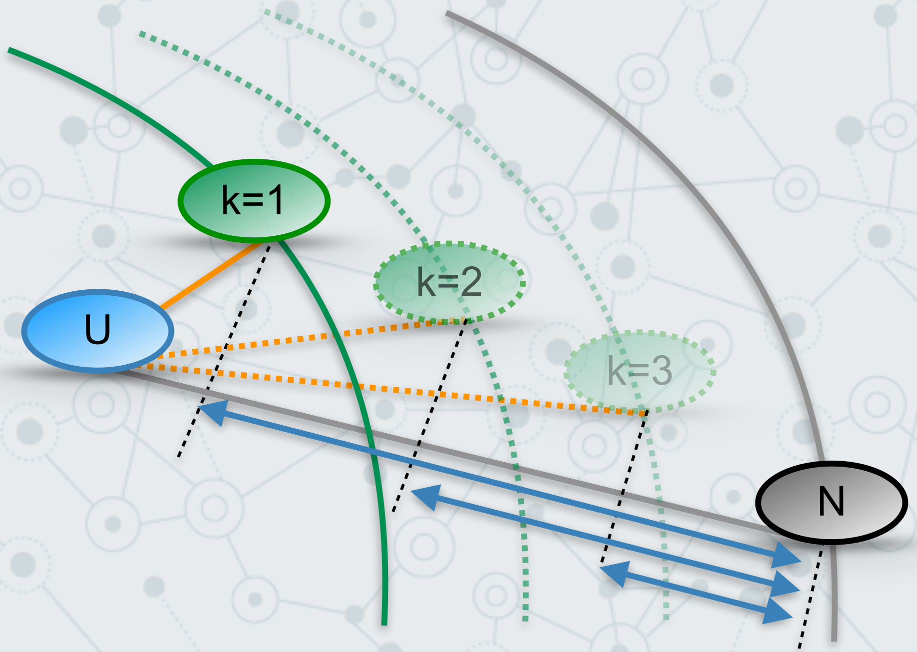
Model: Details

Confidence model



$$\mathcal{L}_{HOP} = \sum_{\substack{1 \le k \le K \\ u,(i,i')}} \overbrace{\mathcal{C}(k)\,\mathbb{E}_{\substack{i \sim P_u^k \\ i' \sim P_N}}}^{\text{graph model}} \overbrace{\left[\mathcal{F}\left(\theta_u^\intercal \theta_{i'}, \theta_u^\intercal \theta_i\right)\right]}^{\text{factorization model}} + \lambda_\Theta \|\Theta\|_2^2$$

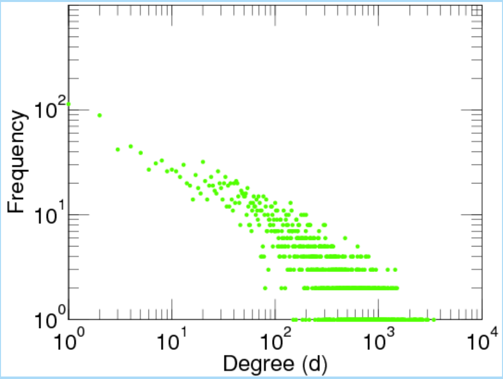
$$\mathcal{F}(\theta_u^\intercal \theta_{i'}, \theta_u^\intercal \theta_i) = \mathbb{1}_{\{\theta_u^\intercal \theta_{i'} - \theta_u^\intercal \theta_i > \epsilon_k\}} \log\left[\sigma\left(\theta_u^\intercal \theta_{i'} - \theta_u^\intercal \theta_i\right)\right]$$

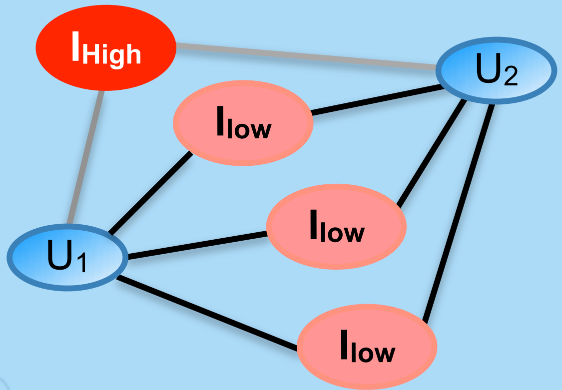


- Items discovered at different order should be **distinguished from other negative items**

Model: Details

Movielens-1M: Movie degree





Degree Sampling

- For each user, **probability** of sampling paths with **high degree** items **will be low**

- Balanced by increase probability of sampling high degree items

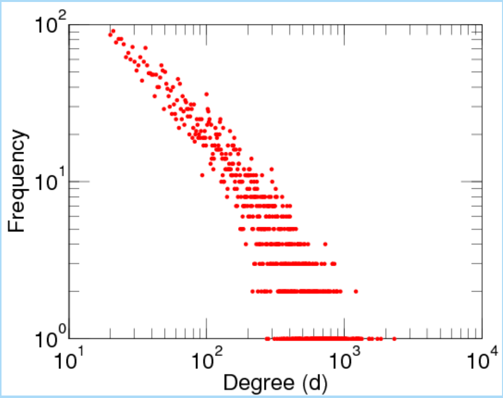
$$\mathcal{L}_{HOP} = \sum_{\substack{1 \le k \le K \\ u,(i,i')}} \overbrace{\mathcal{C}(k)\, \mathbb{E}_{\substack{i \sim P_u^k \\ i' \sim P_N}}}^{\text{graph model}} \overbrace{\left[\mathcal{F}\left(\theta_u^\intercal \theta_{i'}, \theta_u^\intercal \theta_i\right)\right]}^{\text{factorization model}} + \lambda_\Theta \left\|\Theta\right\|_2^2$$

$$p_x^k(y) = \begin{cases} \frac{a_{xy} deg(y)}{\sum_{y'} a_{xy'} deg(y')} & \text{if } k = 1 \text{ and } x \in U, \\ \frac{a_{yx} deg(y)}{\sum_{y'} a_{y'x} deg(y')} & \text{if } k = 1 \text{ and } x \in I, \\ p_x^1(\alpha) p_\alpha^{k-1}(\beta) p_\beta^1(y) & \text{if } k > 1, \end{cases}$$

Christoffel, Fabian, et al. "Blockbusters and wallflowers: Accurate, diverse, and scalable recommendations with random walks." *Proceedings of the 9th ACM Conference on Recommender Systems*. ACM, 2015.

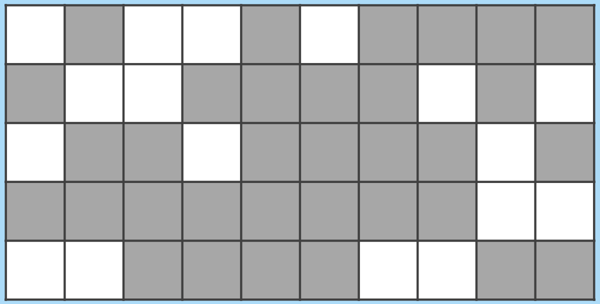
Model: Details

Movielens-1M: User degree



Item

User



Negative sampling

- For each user, **most** items are **not rated** (negative) rather rated (positive)
- We **uniformly** draw negative items from all items

$$\mathcal{L}_{HOP} = \sum_{\substack{1 \le k \le K \\ u,(i,i')}} \overbrace{\mathcal{C}(k)\, \mathbb{E}_{\substack{i \sim P_u^k \\ i' \sim P_N}}}^{\text{graph model}} \overbrace{\left[\mathcal{F}\left(\theta_u^\intercal \theta_{i'}, \theta_u^\intercal \theta_i\right)\right]}^{\text{factorization model}} + \lambda_\Theta \left\|\Theta\right\|_2^2$$

Experiment: Implicit Feedback

To predict whether user U_m interact with item I_n

$$\begin{bmatrix} 1 & 0 & \dots & 1 \\ 0 & 0 & & 1 \\ \vdots & & & \vdots \\ 1 & 1 & \dots & 0 \end{bmatrix} \approx \begin{bmatrix} \\ U \\ \\ \end{bmatrix} \begin{bmatrix} \quad\quad I \quad\quad \end{bmatrix}$$

Datasets

Table 1: Datasets

Dataset	CiteUlike[a]	MovieLens-1M[b]	MovieLens-20M[b]	Amazon-Book[c]
Users ($\|U\|$)	3,527	6,034	136,674	449,475
Items ($\|I\|$)	6,339	3,125	13,680	292,65
Feedback ($\|E\|$)	77,546	574,376	9,977,451	6,444,944
Density	0.347%	3.046%	0.534%	0.005%

[a] http://www.wanghao.in/data/ctrsr_datasets.rar
[b] https://grouplens.org/datasets/movielens
[c] http://jmcauley.ucsd.edu/data/amazon

Experiment: Evaluation

Performance comparison

Table 2: Performance comparison

	CiteUlike			MovieLens-1M			MovieLens-20M			Amazon-Book		
	P@10	R@10	MAP@10	P@10	R@10	MAP@10	P@10	R@10	MAP@10	P@10	R@10	MAP@10
MF	4.1%	13.1%	6.7%	17.7%	13.1%	11.7%	14.9%	14.0%	11.3%	0.7%	3.7%	1.4%
BPR	3.8%	14.2%	6.4%	18.1%	13.2%	12.5%	13.3%	14.3%	10.4%	1.0%	5.3%	2.5%
WARP	5.4%	18.3%	9.1%	24.8%	18.5%	18.5%	20.7%	21.4%	17.2%	1.4%	7.6%	3.2%
K-OS	5.6%	19.4%	9.5%	23.0%	17.3%	16.4%	19.6%	20.5%	15.7%	1.5%	7.9%	3.5%
$RP^3(\beta)$	5.9%	21.2%	3.2%	22.8%	17.2%	14.2%	17.3%	19.4%	10.3%	-	-	-
HOP	5.9%	**21.3%**	*__10.8%__	*__25.9%__	*__20.5%__	*__19.6%__	*__21.2%__	*__22.3%__	*__17.9%__	1.5%	7.9%	*__3.6%__
%Improv.	0.0%	0.5%	13.7%	4.4%	10.8%	5.9%	2.4%	4.2%	4.1%	0.0%	0.0%	2.9%

Experiment: Evaluation

K-order sensitivity

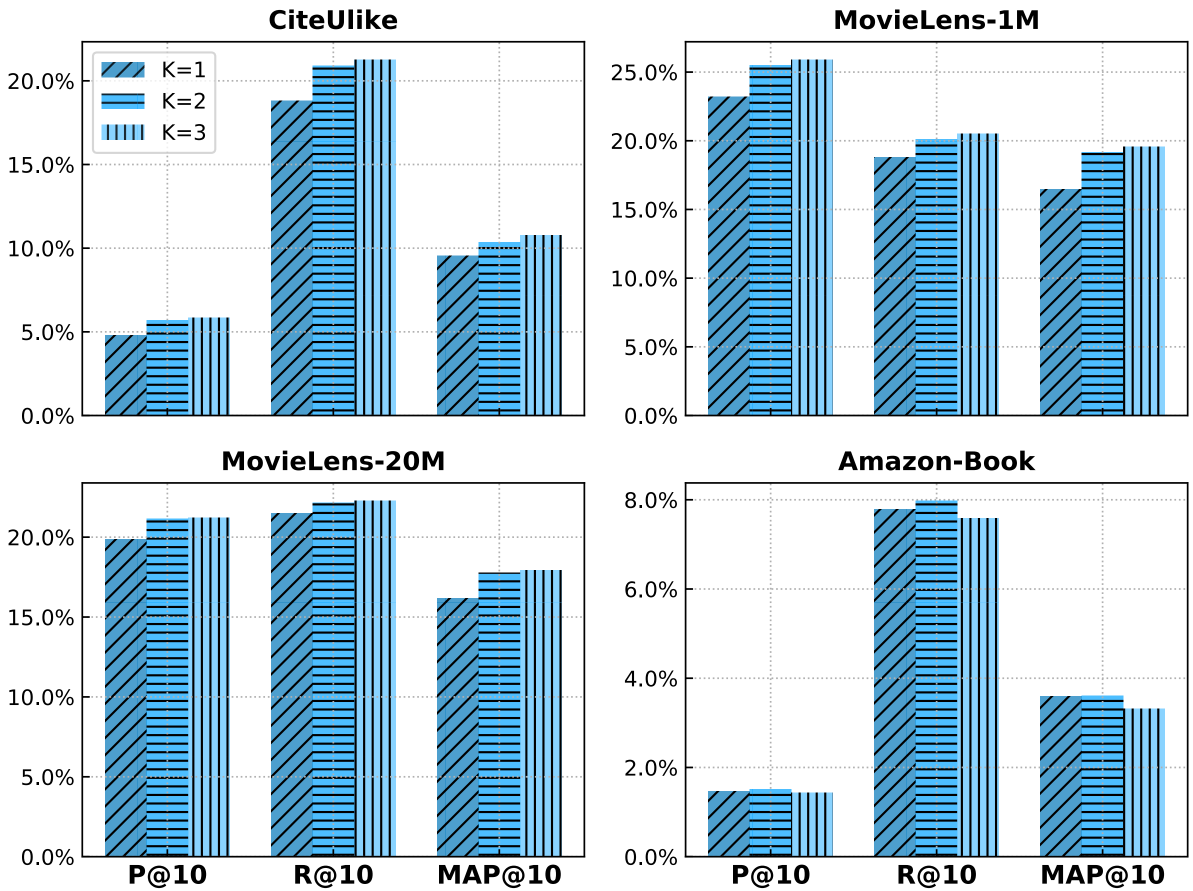


Figure 2: Sensitivity analysis with respect to K

Conclusion

Take-home Messages

- We demonstrate an approach that combines both representative CF-based methods
- High-order information within user-item interaction matrix is helpful for implicit recommendation

Conclusion

Future Works

- Conduct comprehensive study on walking strategies
- Extend HOP-Rec to explicit feedback problems
- Examine other metrics: diversity
- Develop more efficient negative sampling strategy
- Develop other forms of confidence model

Thanks!

I am Jheng-Hong Yang

You can find me at:

jhyang@citi.sinica.edu.tw

DOI

CFDA & CLIP Labs

Source Code: ProNet