Recommendation algorithms in the macro-evolving network

Xiao Hu

UESTC

huxiao64@gmail.com

Collaborator: Dr. An Zeng

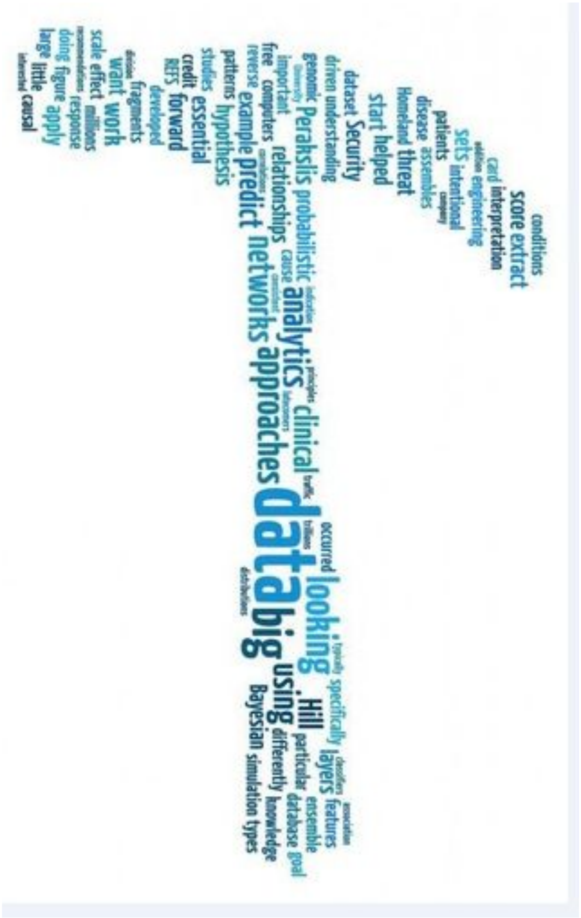
Fribourg University

Outline

- Introduction
- Recommendation algorithms
- The macro-evolution of recommendation algorithms
- Some questions

Introduction

- Big data age
- Information overload
- Recommendation systems
- However, most algorithms only evaluate their performance in single recommendation step
- The recommendation performance in the evolving network is overlooked



Metrics

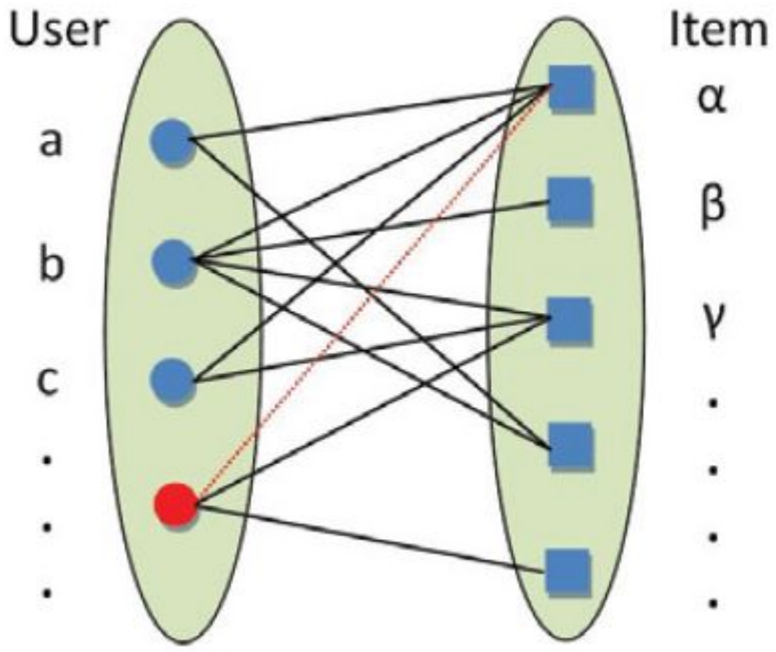
- Ranking score(RS)
- RS measures whether the ordering of the items in the recommendation list matches the users' real preference.
- $< RS >= \frac{1}{|E^P|} \sum_{i\alpha \in} RS_{i\alpha}$

Metrics

- Recall
- Precision
- Novelty
    - $N(L) = \frac{1}{UL} \sum_{i=1}^{U} \sum_{\alpha \in O_L^i} k_\alpha$

Recommendation algorithms

- Bipartite Graph(user-item network)
- $G = (U, O, E)$,
  - $U = (u_1, u_2, \dots, u_m)$
  - $O = \left(o_{1,} o_2, \dots, o_n\right)$
  - $E = (e_1, e_2, \dots, e_k)$

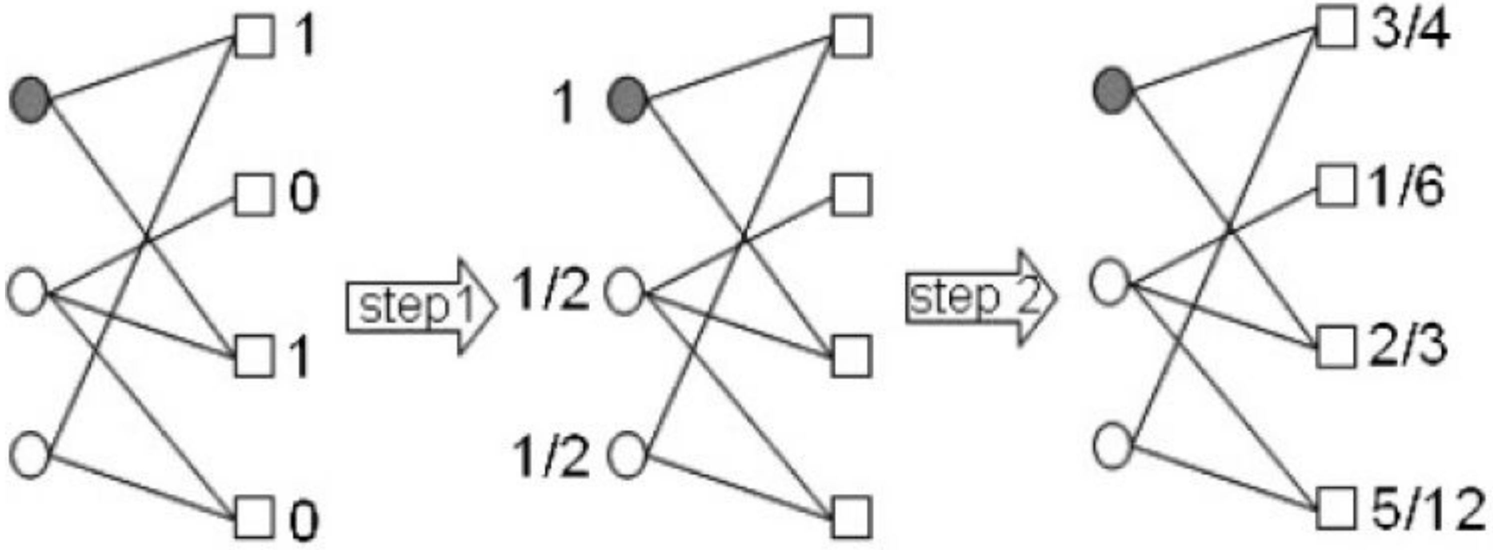


Mass Diffusion (MD)

Zhou, Tao, et al. "Bipartite network projection and personal recommendation." *Physical Review E* 76.4 (2007): 046115.

- Assign one unit of resource to each object
- Redistribute the resource through the user-item network

- $W_{\alpha\beta} = \frac{1}{k_\beta} \sum_{j=1}^{N} \frac{a_{j\alpha} a_{j\beta}}{k_j}$

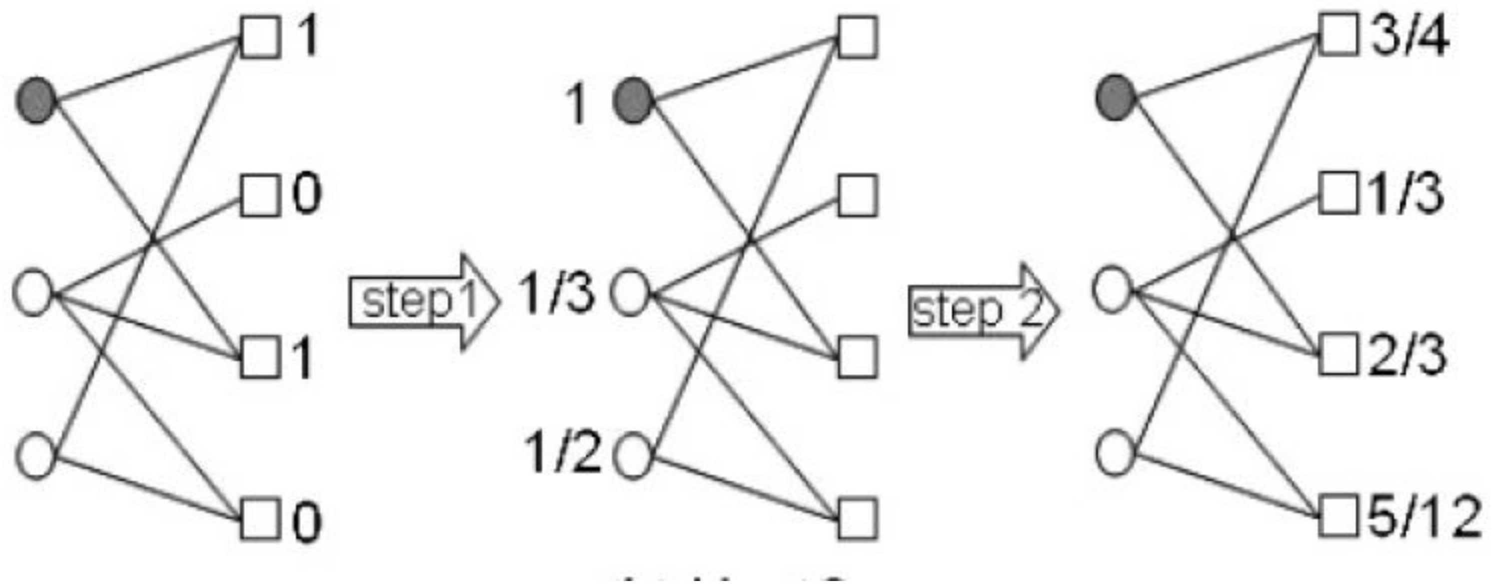


Heat Conduction (HC)

Zhang, Yi-Cheng, Marcel Blattner, and Yi-Kuo Yu. "Heat conduction process on community networks as a recommendation model." *Physical review letters*99.15 (2007): 154301.

- Assign one unit of resource to each object
- Redistribute the resource through the user-item network

- $W_{\alpha\beta} = \frac{1}{k_\alpha}\sum_{j=1}^{N}\frac{a_{j\alpha}a_{j\beta}}{k_j}$



MD vs. HC

- Mass diffusion has high recommendation accuracy yet low diversity
- Heat conduction has high diversity yet low recommendation accuracy

Hybridization with MD and HC

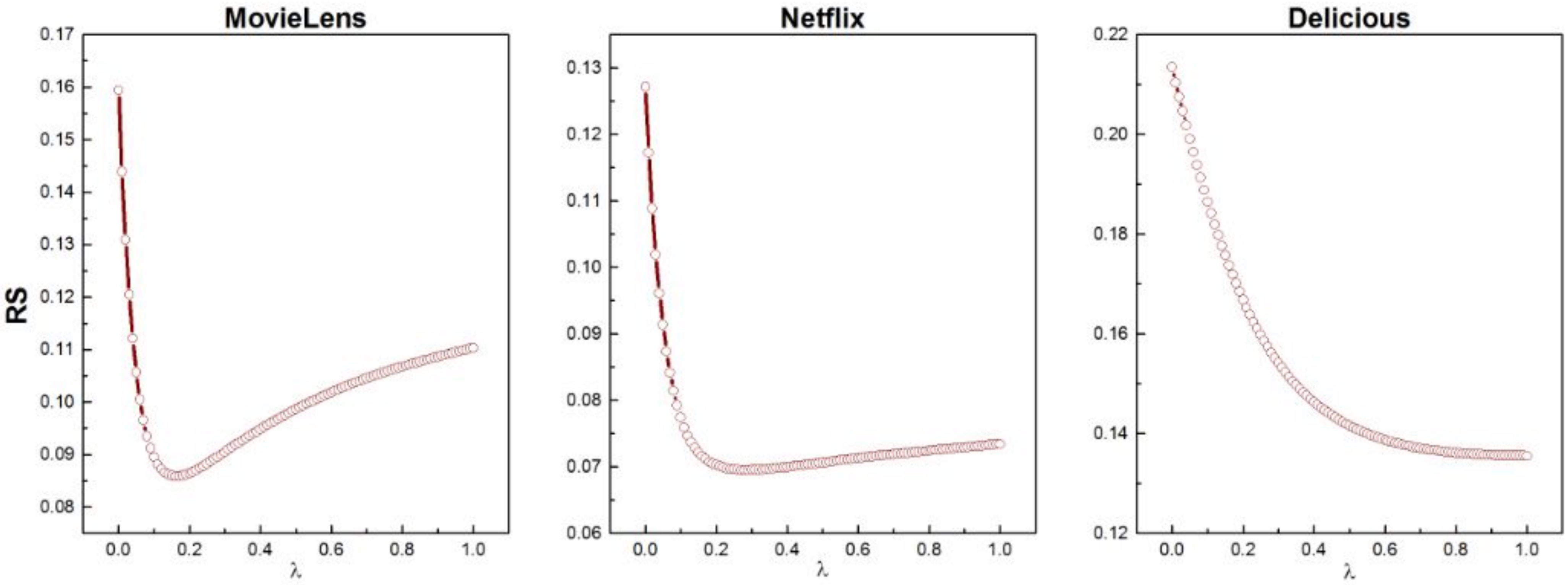
- to solve the accuracy-diversity dilemma
- Combine Mass Diffusion and Heat Conduction

- $W_{\alpha\beta}^{M+H} = \frac{1}{k_\alpha^{1-\gamma} k_\beta^\gamma} \sum_{j=1}^{N} \frac{a_{j\alpha} a_{j\beta}}{k_j}$ $\gamma \epsilon [0,1]$

- $\gamma = 0 \rightarrow \text{Heat Conduction}$
- $\gamma = 1 \rightarrow Mass\ Diffusion$

Zhou, Tao, et al. "Solving the apparent diversity-accuracy dilemma of recommender systems." *Proceedings of the National Academy of Sciences*107.10 (2010): 4511-4515.

Hybridization with MD and HC

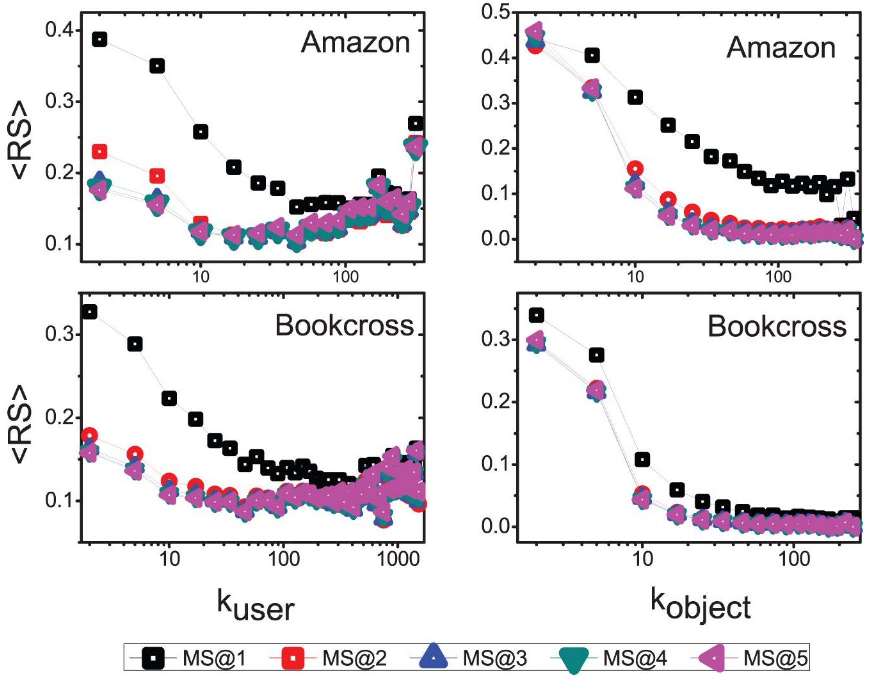


Semi-Local Diffusion(SLD)

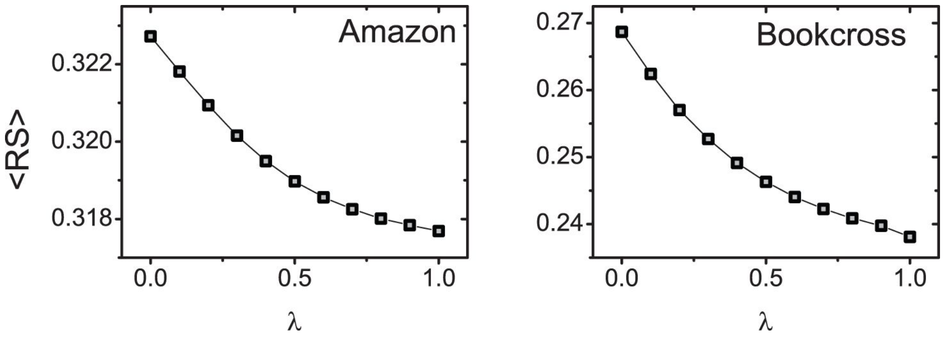
Dataset	#user	#objects	#links	sparsity
Amazon	50000	54,152	283,382	1035×10^{-4}
Bookcross	21122	203,373	504,643	1.17×10^{-4}

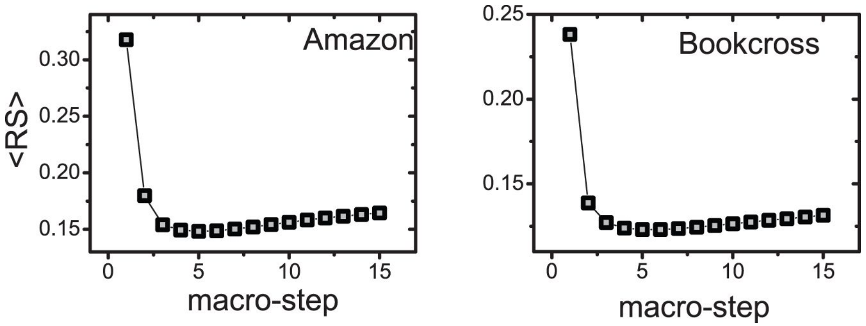
The sparsity is obtained by $\frac{\#links}{N \times M}$, where N and M are the number of users and items, respcetively.

doi:10.1371/journal.pone.0079354.t001

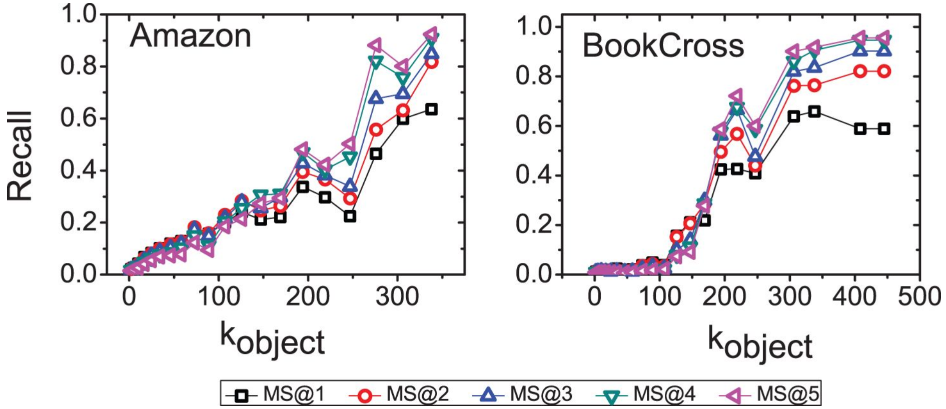


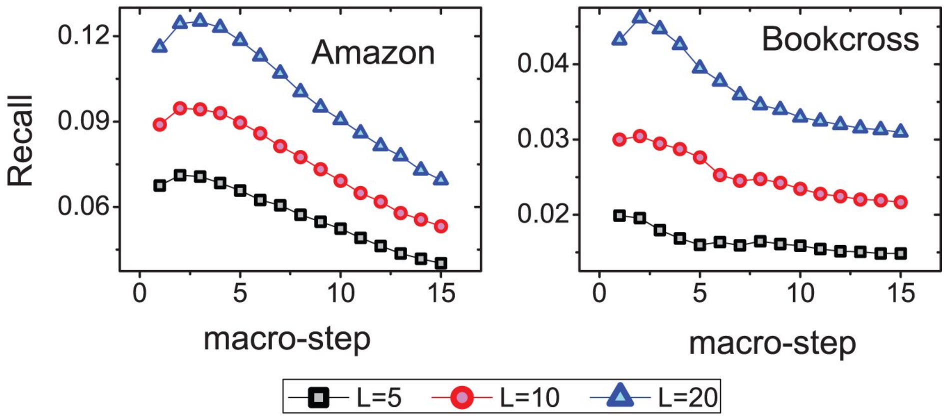
em

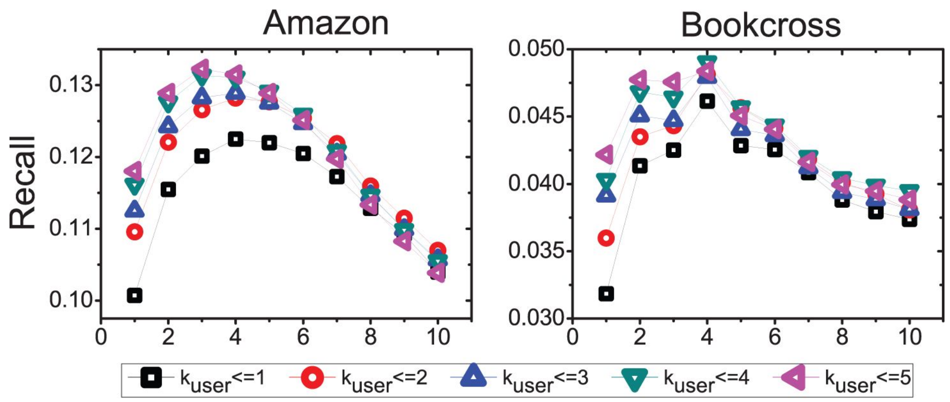


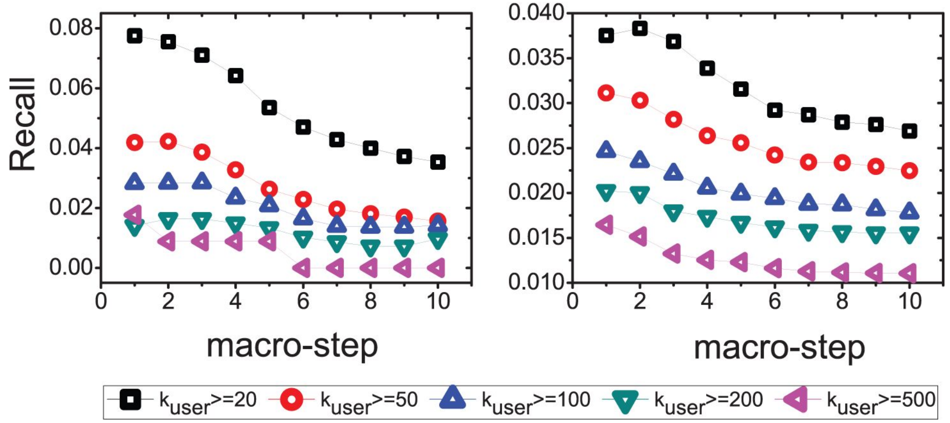


Semi-Local Diffusion(SLD)



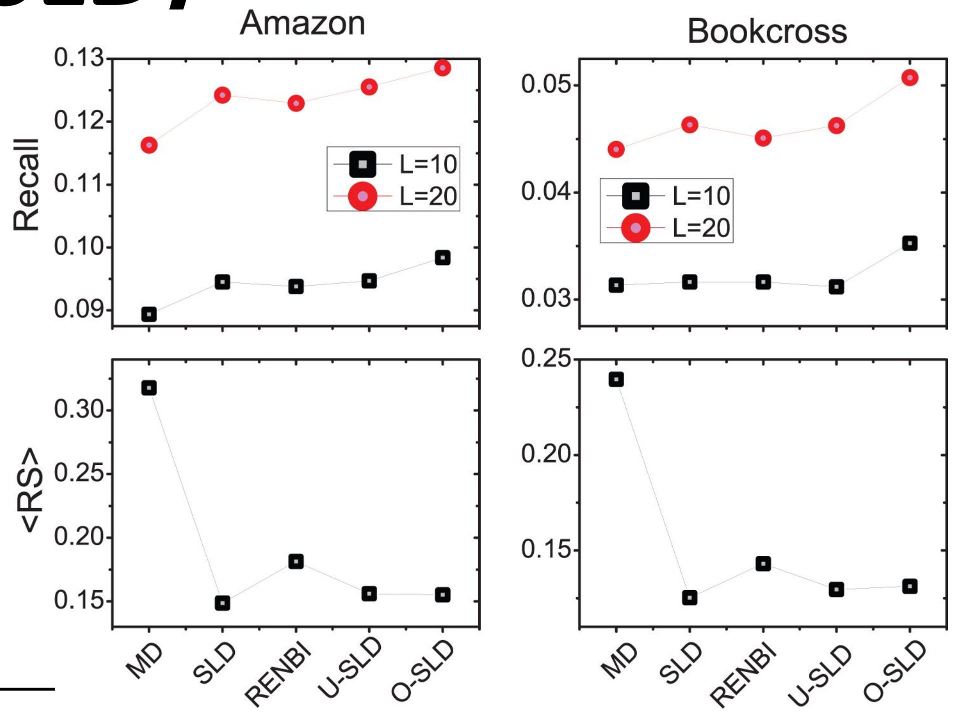






Semi-Local Diffusion(SLD)

- User-based semi-local diffusion
  - $F_\alpha^u = f_\alpha^{(1)} + \sum_{i=2}^{n} \frac{1}{(K-k_u)^\theta} f_\alpha^{(i)}$
  - $K = \max(k_u) + 1$
- Item-based semi-local diffusion
  - $F_\alpha^u = f_\alpha^{(1)} + \sum_{i=2}^{n} \frac{1}{k_\alpha^\theta} f_\alpha^{(i)}$



Amazon					
		SLD-T	**RENBI**	**U-SLD**	**O-SLD**
Recall	T	2	–	–	–
	θ	–	2	−0.9	−0.3
Ranking score	T	5	–	–	–
	θ	–	2	−1	−0.5
Bookcross					
		SLD-T	**RENBI**	**U-SLD**	**O-SLD-**
Recall	T	2	–	–	–
	θ	–	1.0	−0.6	−0.2
Ranking score	T	5	–	–	–
	θ	–	2	−1	−0.7

doi:10.1371/journal.pone.0079354.t002

Zeng, Wei, et al. "Information filtering in sparse online systems: recommendation via semi-local diffusion." *PloS one* 8.11 (2013): e79354.

Other extension

- Preferential diffusion
- Similarity-preferential diffusion
- RENBI
- Network manipulation

The long-term influence of algorithms

- most algorithms only evaluate their performance in single recommendation step
- the recommendation performance in the evolving network is overlooked
- The network structure vs. The performance of algorithm

Method

- In one macro-step of our simulation, we randomly choose 10% of users to be active. After each macro-step, we evaluate the performance of some recommendation algorithms

- Note that we do not consider the growth of the system since introducing new users or items may involve the cold start problem for them

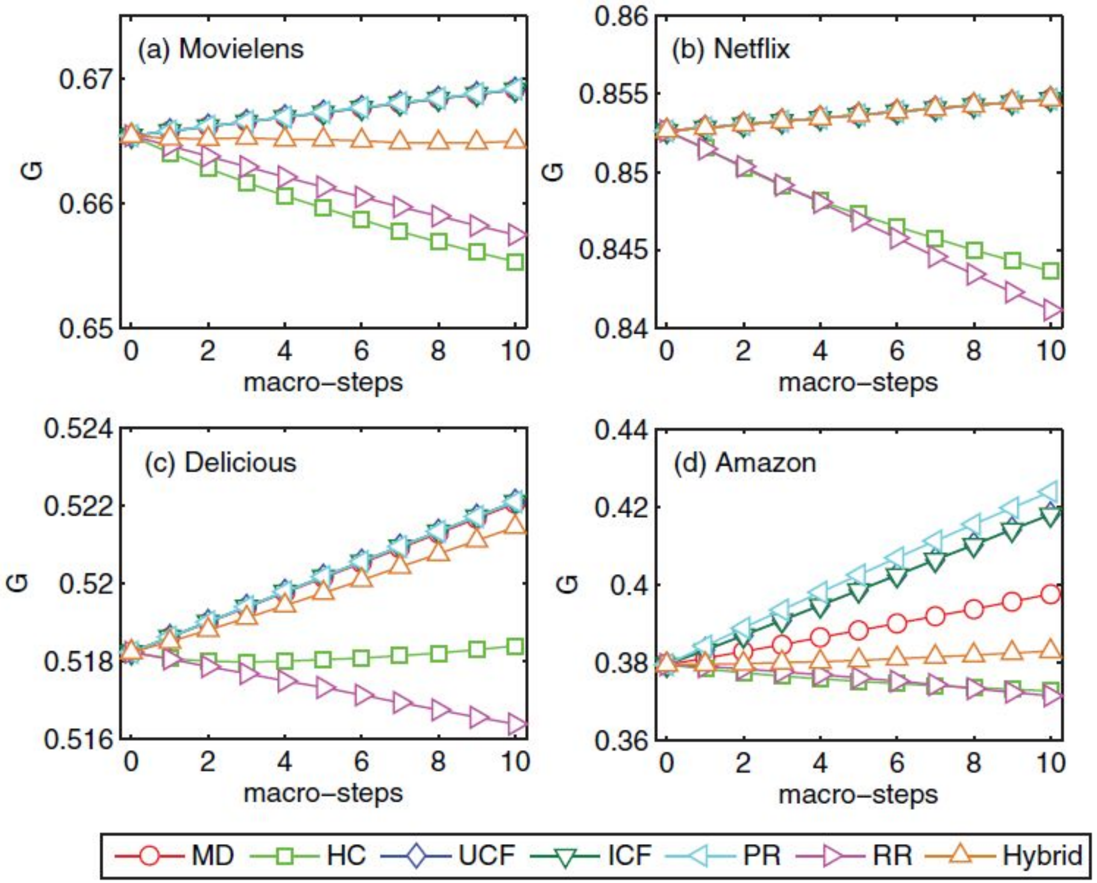
Method	Acronym
Mass Diffusion	MD
Heat Conduction	HC
User-based Collaborative Filtering	UCF
Item-based Collaborative Filtering	ICF
Popularity-based Recommendation	PR
Random Recommendation	RR

The influence on global diversification

At each step, a random user is selected as the active user, and the recommendation scores of all items are then evaluated for him/her.

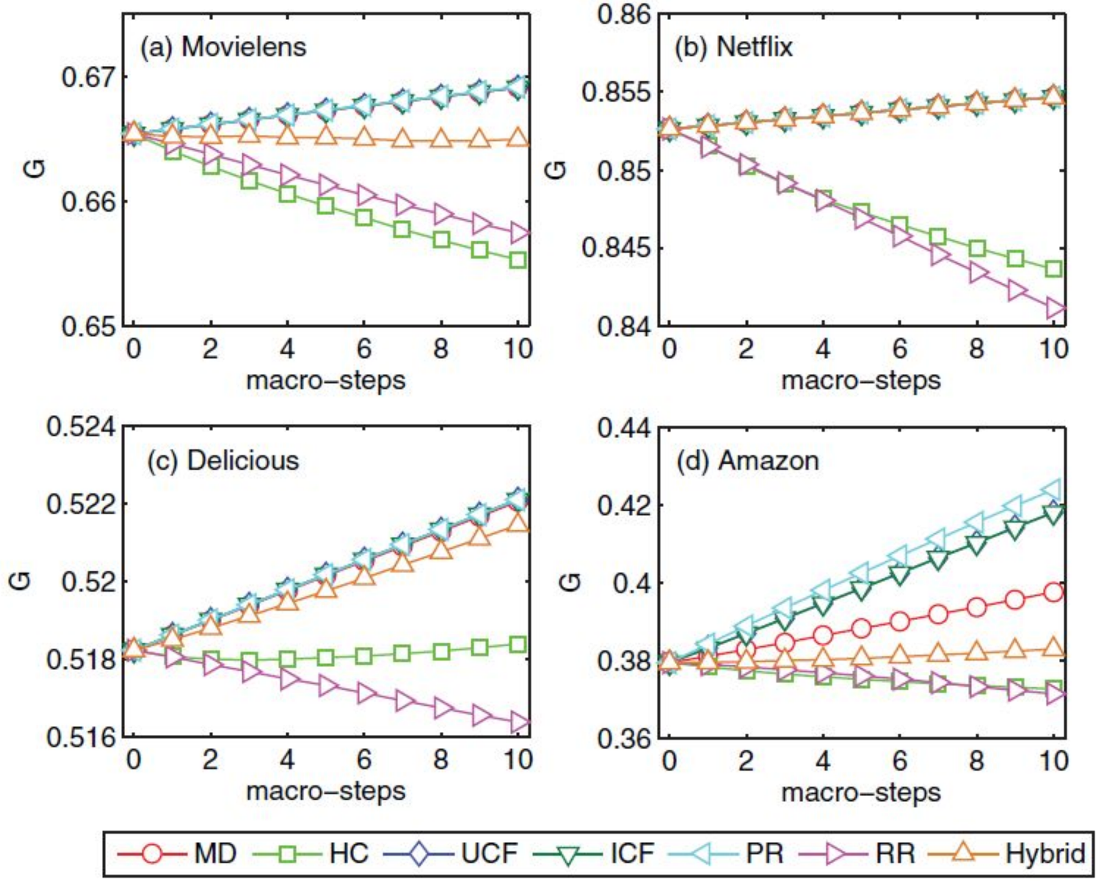
For simplicity, we assume that the active user would accept the recommendations by selecting the uncollected item with the highest score

Network	Users	Items	Links	Sparsity
Movielens	943	1682	82520	$5.20 \cdot 10^{-2}$
Netflix	10000	6000	701947	$1.17 \cdot 10^{-2}$
Delicious	10000	232657	1233997	$5.30 \cdot 10^{-4}$
Amazon	20000	66525	258911	$1.95 \cdot 10^{-4}$



Explanation

- MD, UCF and ICF only recommend the relevant items according to former choices of users.
- HC does not reinforce the popularity of hot items as MD, UCF and ICF
- On the other hand, HC is different from the uniform addition of links in RR, as it inclines to add links to items with small degree

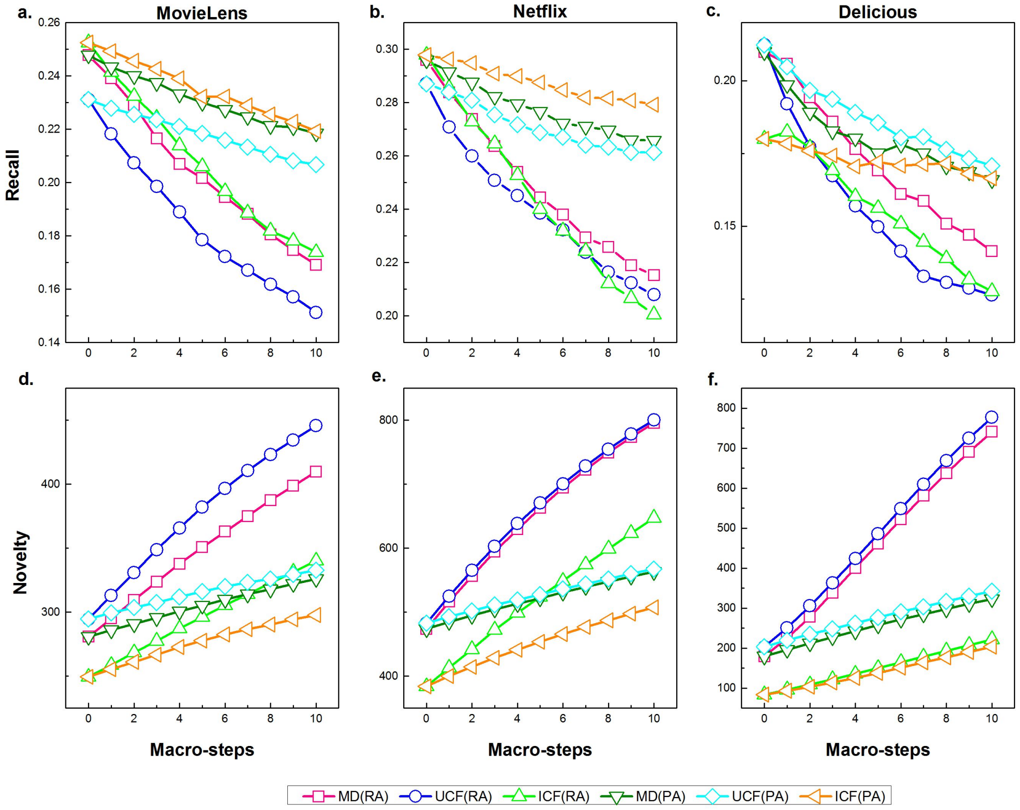


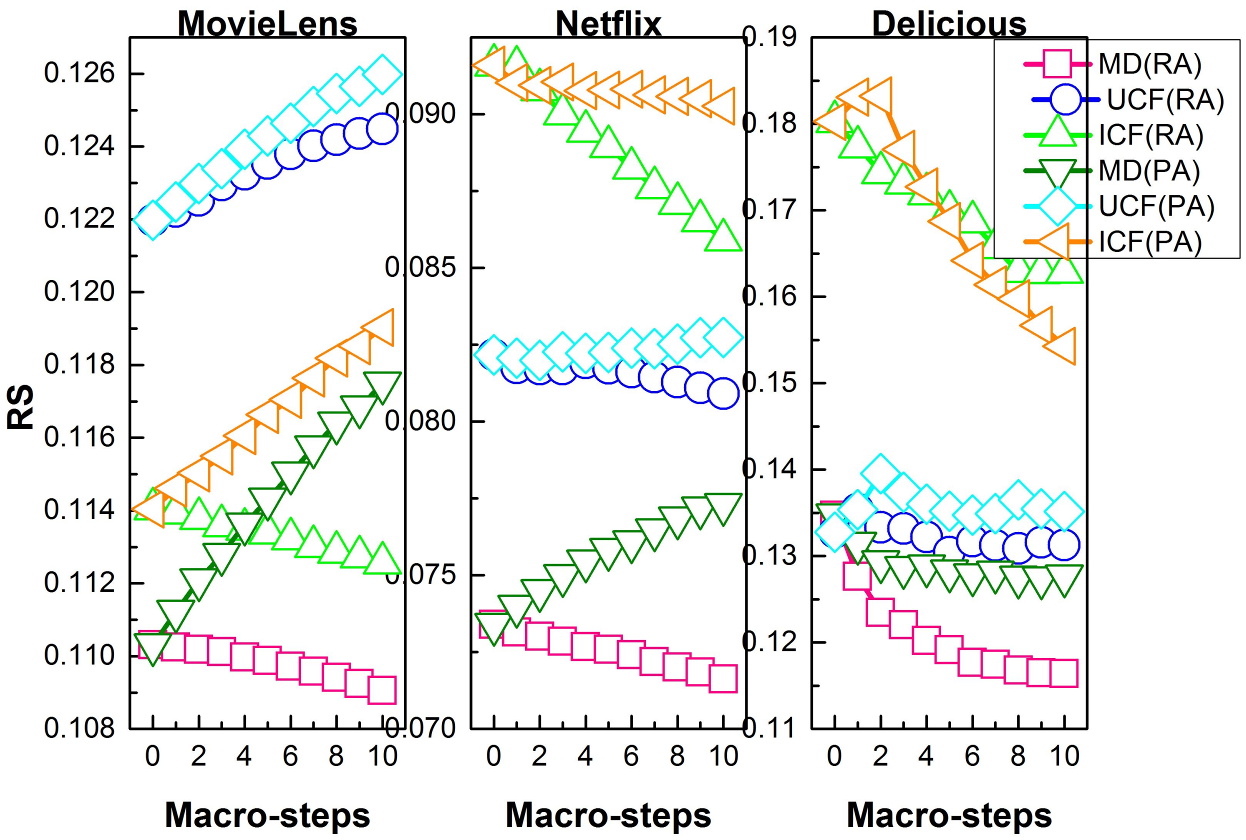
The influence on recommendation accuracy

- **Preferential attachment(PA)**
- **Recommendation algorithm(RA)**

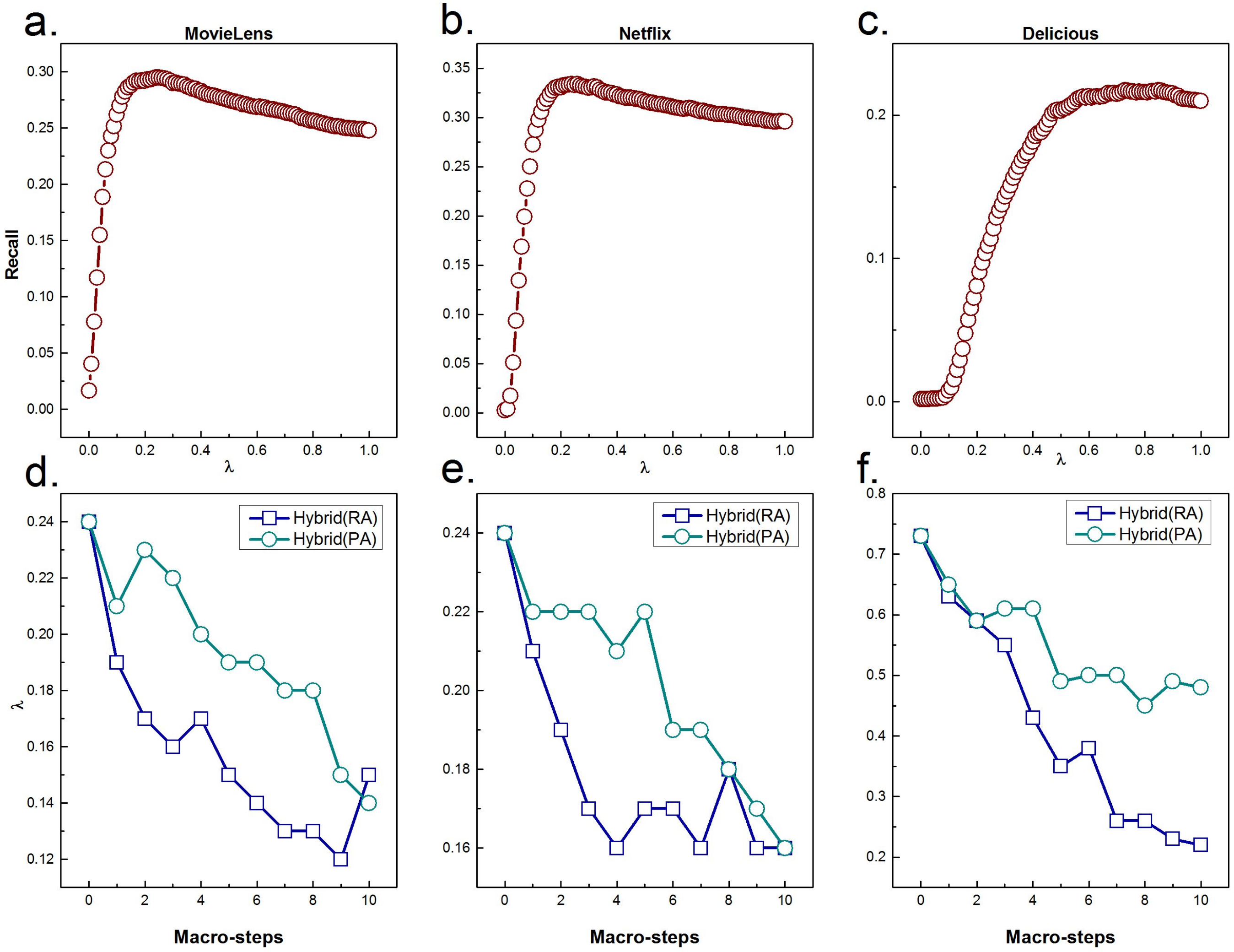
Data set	Users	Items	Links
MovieLens	943	1574	82520
Netflix	2294	1861	71074
Delicious	4733	3058	41924

The influence on recommendation accuracy





The influence on Hybrid parameter



Ongoing Work

- In a market, how to simulate the user behavior pattern(item selection).
- In a market, how to design a robust recommendation algorithm with a high quality in a long-term.
- In a market, how to give some advices on new products selection for a manufacturer.

Thank You all!