IBM Research - China

Temporal Recommendation on Graphs via Long- and Short-term Preference Fusion

Liang Xiang, Quan Yuan, Shiwan Zhao, Li Chen,
Xiatian Zhang, Qing Yang, Jimeng Sun

Presenter: Quan Yuan

Problem & Challenges

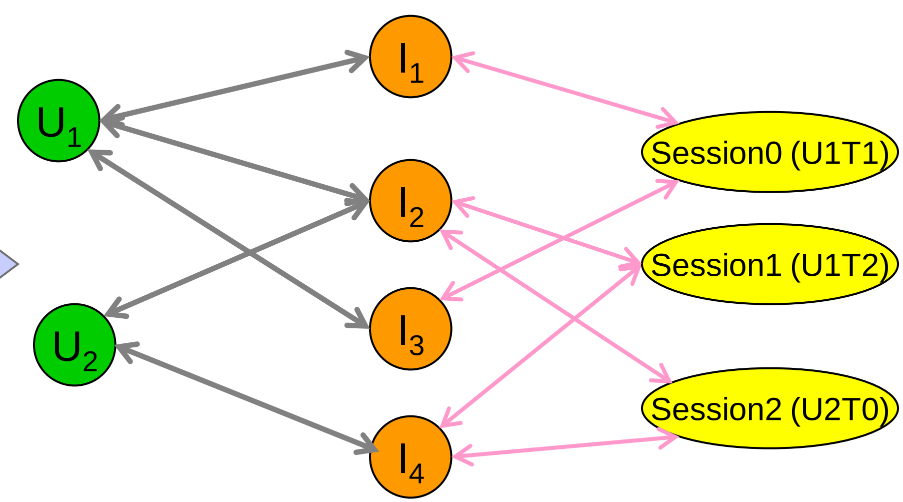
- Temporal dynamics is crucial in recommender system.
  - [Koren KDD09], [Liu IUI10], etc
- Temporal recommendation focuses more on **local** recommendation models for each user
- When modeling individual, one's behavior is usually determined by long-term interests and short-term interests
- Challenges
  - How to represent and balance users' long-term and short-term preferences?

Motivation for Session-based Temporal Graph (STG)

- input data <*user, item, time*>
- User-Item Matrix usually can be represented as a bipartite graph
- When incorporating time factors, we introduced a new type of node – "*session node*"
  - Session: dividing the time slices into bins and binding the bins with corresponding users
- Time dimension is a ***local*** effect of user, treat time as a universal dimension shared by all users is not very effective, e.g. tri-partite graph or tensor

	I1	I2	I3	I4
U1	1	1	1	0
U2	0	1	0	1

User-Item Matrix



Bipartite graph

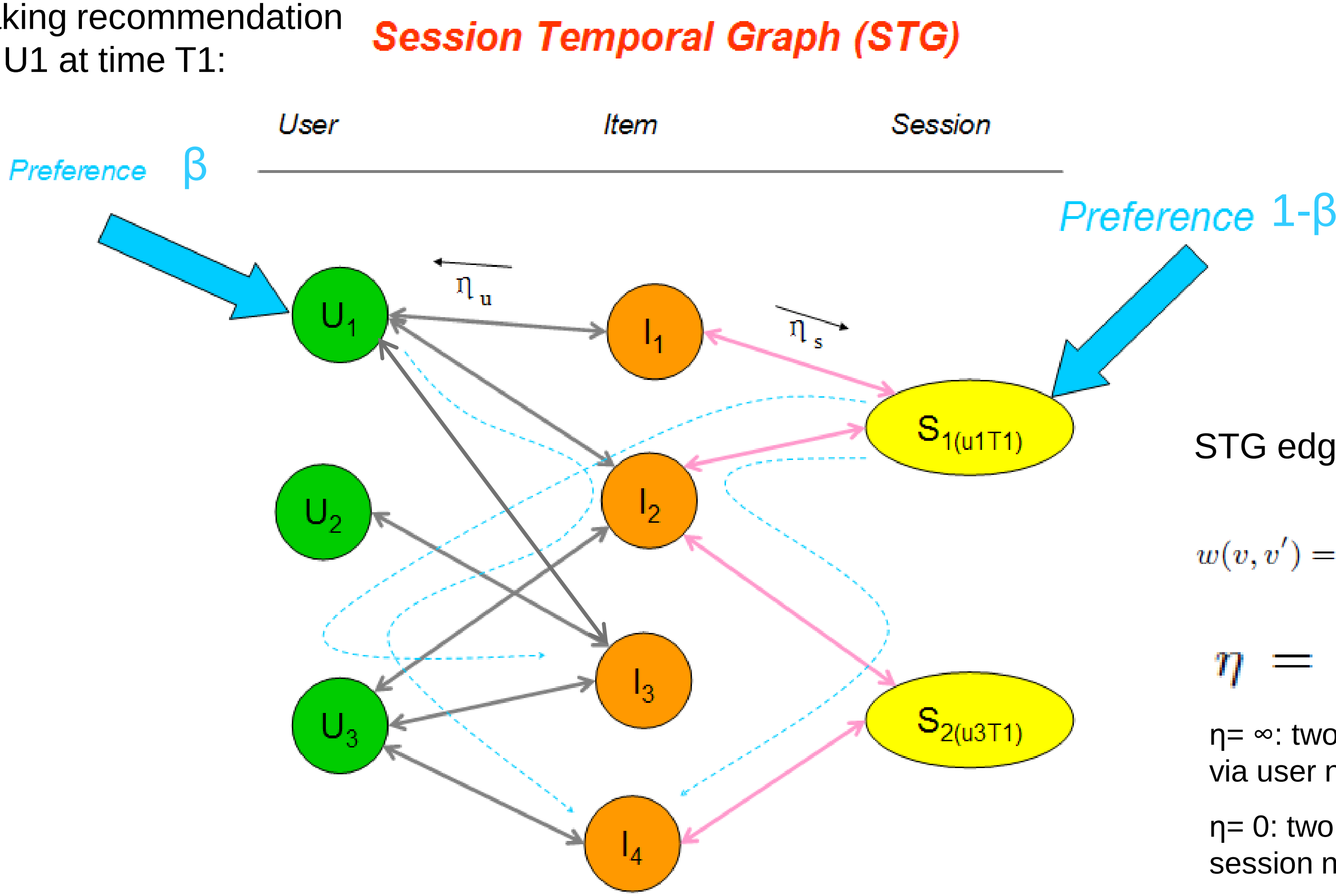
Add time factors
as session node

Injected Preference Fusion (IPF on STG)

- An algorithm based on STG, which balancing the impact of long-term and short-term preferences when making recommendation
- Basic Idea:
  - Injected Preferences into both user node (β) and session node (1 -β)
  - Then in propagation process, the preferences were propagated to an unknown item node
  - Finally, the nodes which get most preferences will be recommended to current user

Injected Preference Fusion on STG

Making recommendation for U1 at time T1:



STG edge weight definition:

$$w(v,v') = \begin{cases} 1 & v \in U \cup S, v' \in I \\ \eta_u & v \in I, v' \in U \\ \eta_s & v \in I, v' \in S \end{cases}$$

$$\eta = \eta_u / \eta_s$$

η= ∞: two items only connected via user nodes;

η= 0: two items only connected via session nodes;

Paths from user node:	user -> item -> user -> item; user -> item -> session -> item
Paths from session node:	session ->item -> user ->item; session -> item ->session -> item

Experiments

- Data sets
  - CiteULike

User bookmark a paper at some time

User	4,607
Item	16,054
User-item pair	109,346
sparsity	99.85%

  - Delicious

User bookmark a web page at some time

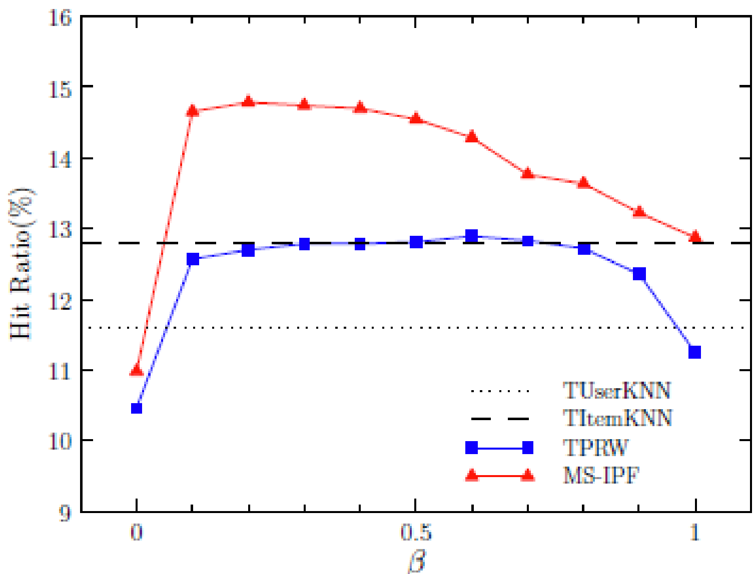
User	8,861
Item	3,257
User-item pair	59,694
sparsity	99.79%

  - Evaluation Metric

$$\text{Hit Ratio} = \frac{\sum_u I(T_u \in R(u,t))}{|U|}$$

    - Hit Ratio: Put the latest item of each user into test set, then generate a list of N (N=10) items for everyone at time t. If the test item appears in the recommendation list, we call it a hit
  - Compared Algorithms
    - Temporal User KNN
    - Temporal Item KNN
    - Temporal Personalized Random Walk

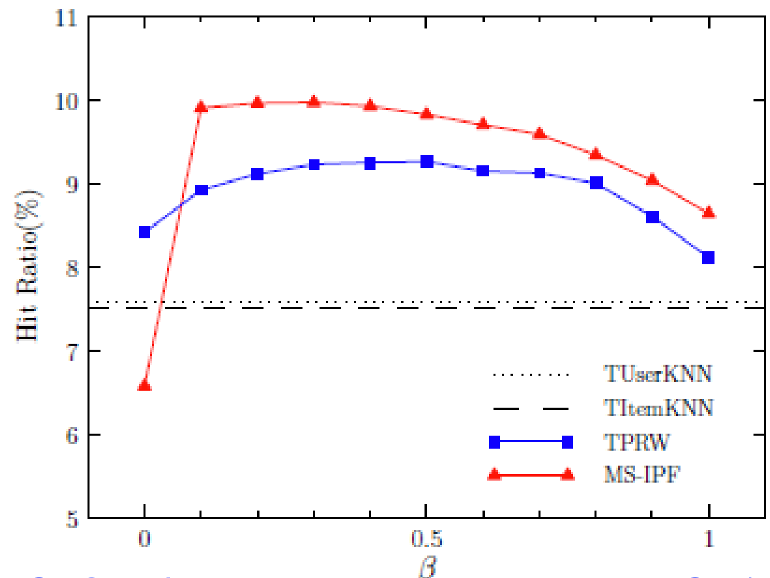
β's impact – Balance the Injected Preferences on User and Session node



β= 0, only short-term

CiteULike

β= 1, only long-term



β= 0, only short-term

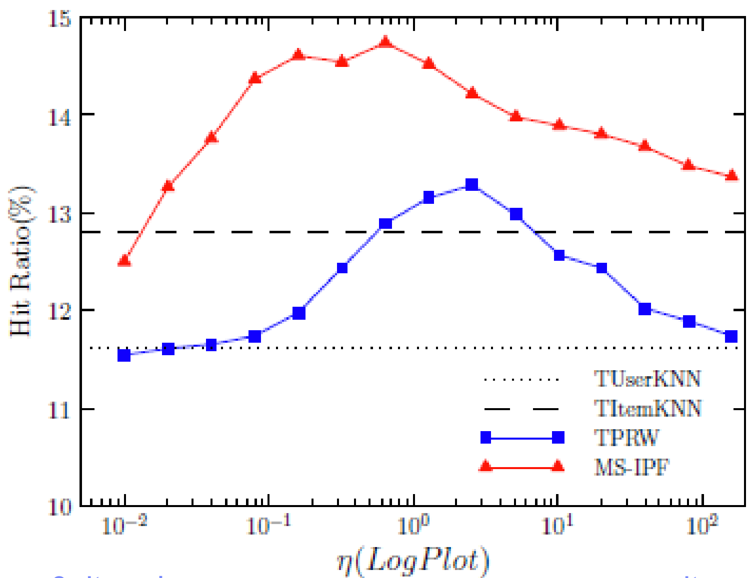
Delicious

β= 1, only long-term

- Optimal results were get when βbelongs to [0.1,0.6];
- Proves the impacts of both long-term and short-term preferences in making good recommendation

η's impact -- Control the ratio of preferences (from an item node) flow to user node against to session node

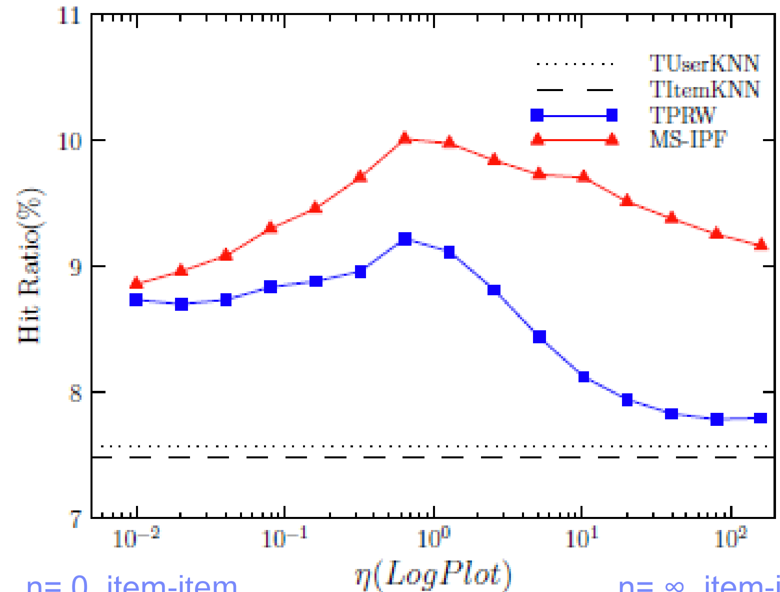
$$\eta = \eta_u / \eta_s$$



η= 0, item-item connected only by session

CiteULike

η= ∞, item-item connected only by user



η= 0, item-item connected only by session

Delicious

η= ∞, item-item connected only by user

- Proves the effectiveness of balancing long-term and short-term preferences in propagation process
- Since X-axis is the logarithm value, it means the optimal hit ratio can be get for a wide range of η

Session size's impact on Hit Ratio of IPF

time window (days)	CiteULike	Delicious
1	13.85%	9.83%
2	13.70%	9.72%
3	13.70%	9.72%
4	13.76%	9.72%
5	13.81%	9.74%
6	13.87%	9.68%
7	13.85%	9.69%
15	13.76%	9.59%
30	13.81%	9.48%
45	13.35%	9.24%
60	13.24%	9.20%
90	13.19%	8.83%

- The result is not very sensitive to the size of time window
  - On CiteULike, the optimal time window is about one week
  - On Delicious, the optimal time window is about one day
- Users' interests on research topics (CiteULike) drift more slowly than interests on browsing web pages (Delicious), proves with our real life experience.

Overall Accuracy Comparison

Method	Hit Ratio	Improvement
TItemKNN	12.85%	–
TUserKNN	11.63%	-9.49%
TPRW	13.46%	4.75%
MS-IPF	14.78%	15.02%

CiteULike

Method	Hit Ratio	Improvement
TItemKNN	7.49%	–
TUserKNN	7.58%	1.2%
TPRW	9.39%	25.37%
MS-IPF	10.07%	34.45%

Delicious

- User Temporal Item KNN as baseline,
  - On CiteULike, MS-IPF improves TItemKNN up to 15.02%;
  - On Delicious, MS-IPF improves TItemKNN up to 34.45%

Conclusion

- Propose a Session-based Temporal Graph (STG) to incorporate temporal information on the graph
- Based on STG, we propose Injected Preference Fusion (IPF) to balance the impact of users' long-term and short-term preferences.
- Compare with other approaches on two real datasets, which confirm that STG's effectiveness for incorporating temporal data, and IPF is effective to balance users' long-term and short-term preferences.