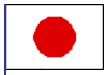
Serendipitous Recommendation for Scholarly Papers Considering Relations Among Researchers

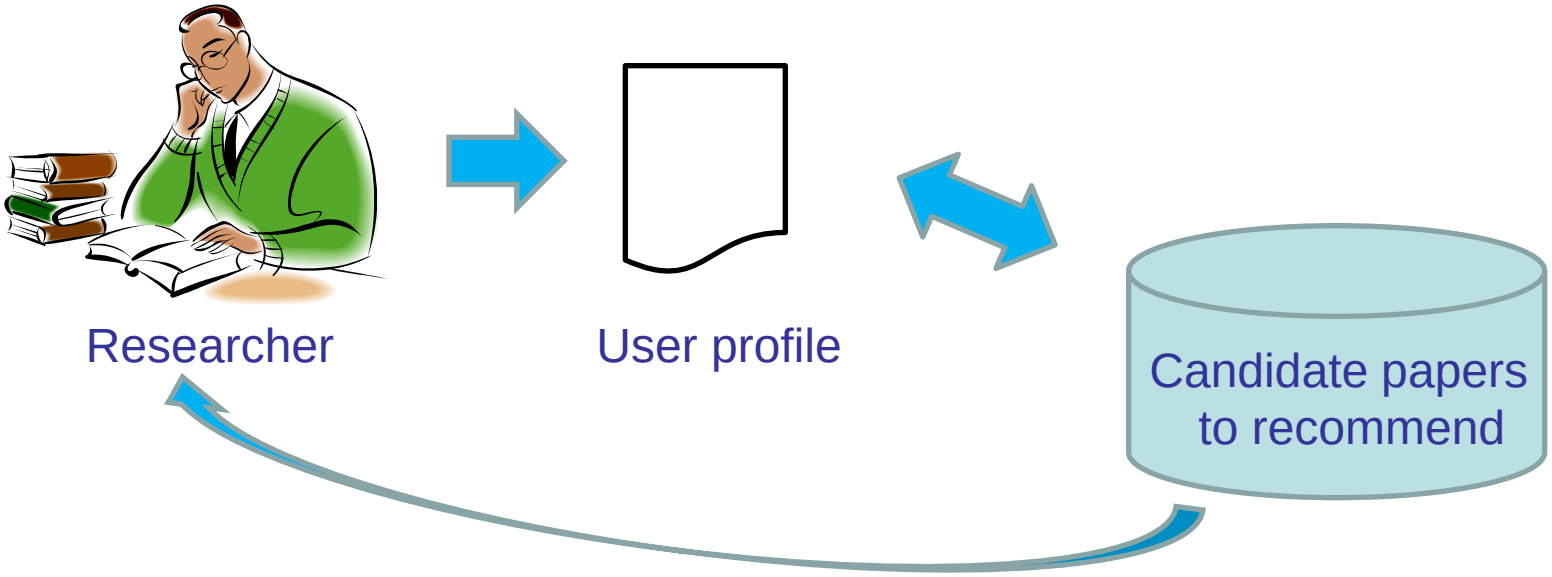



Kazunari Sugiyama, Min-Yen Kan

National University of Singapore

Introduction

Content-based Recommendation



Papers similar to each user's profile are recommended.

Introduction
[Sugiyama and Kan, JCDL'10]

Publication list

[No published papers in the past]

('11)

Broaden their range of research interests.

Junior researcher (having only one recently published paper)

Publication list

('01) ('02) -------- ('11)

Seek to apply their knowledge towards other areas.

Senior researcher (having several past published papers)

Serendipitous recommendation is important.

Introduction

When we find something serendipitous?

Advice from colleagues

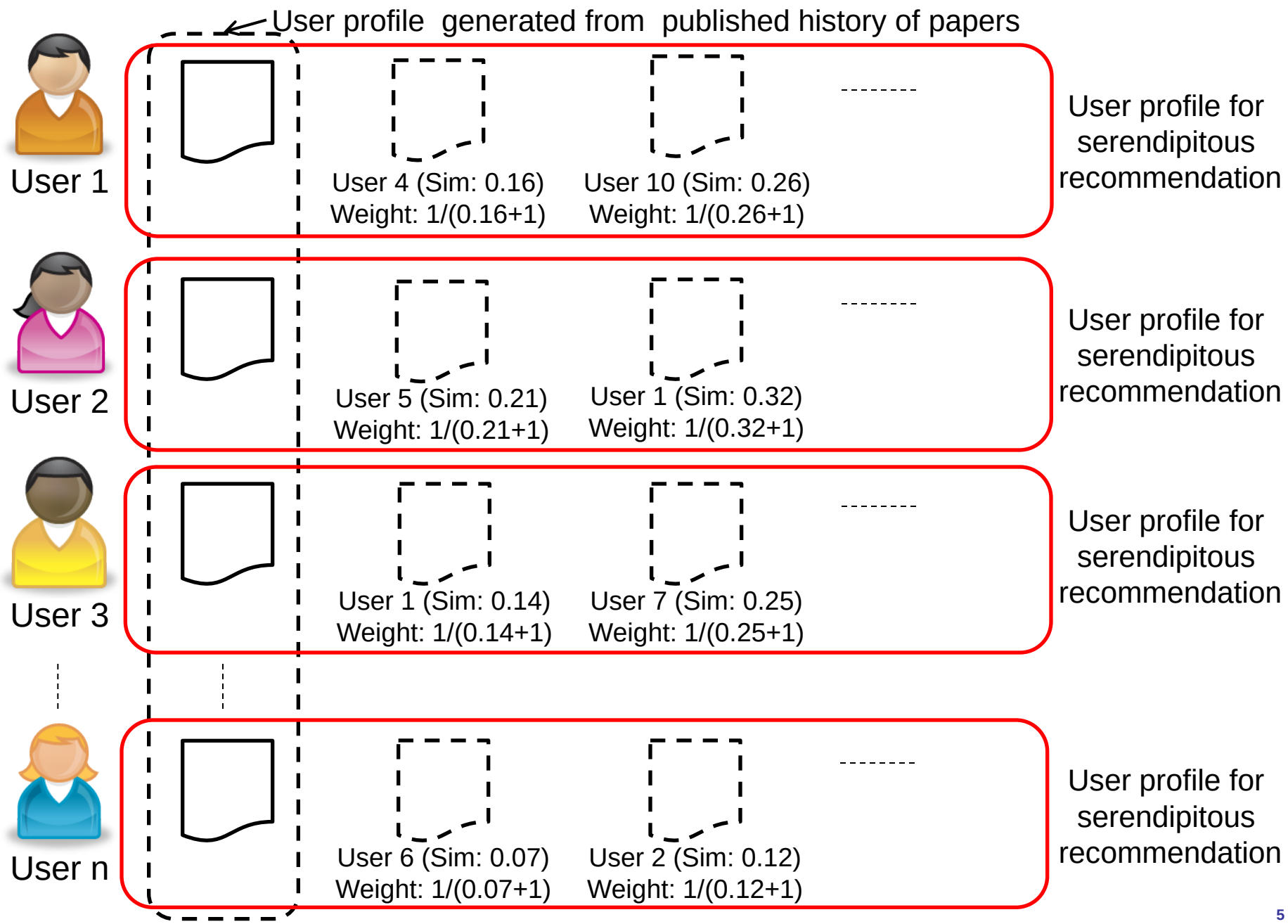
Attend seminars

Serendipitous discovery:
Interactions with others play an important role.

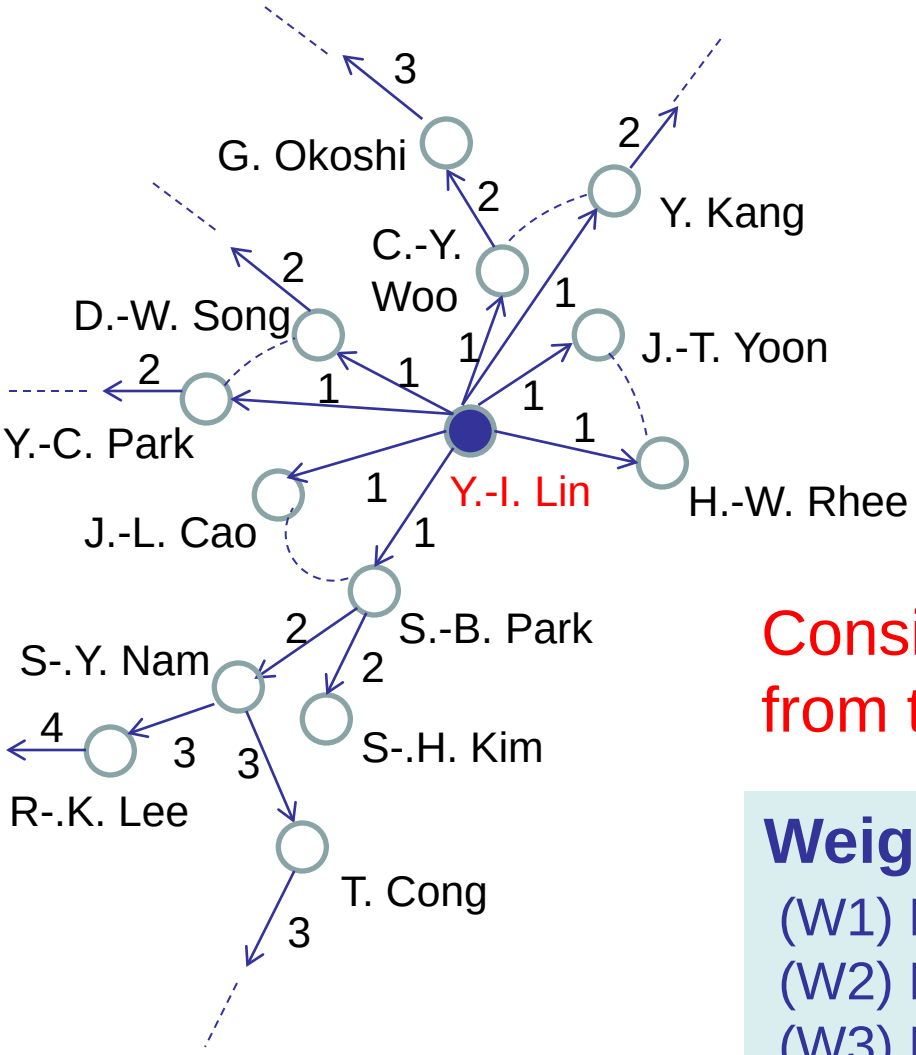
User profile construction for serendipitous recommendation with others

I. Dissimilar users
II. Co-author network

User Profile Construction via Dissimilar Users (DU)



User Profile Construction via Co-author Network (CAN)

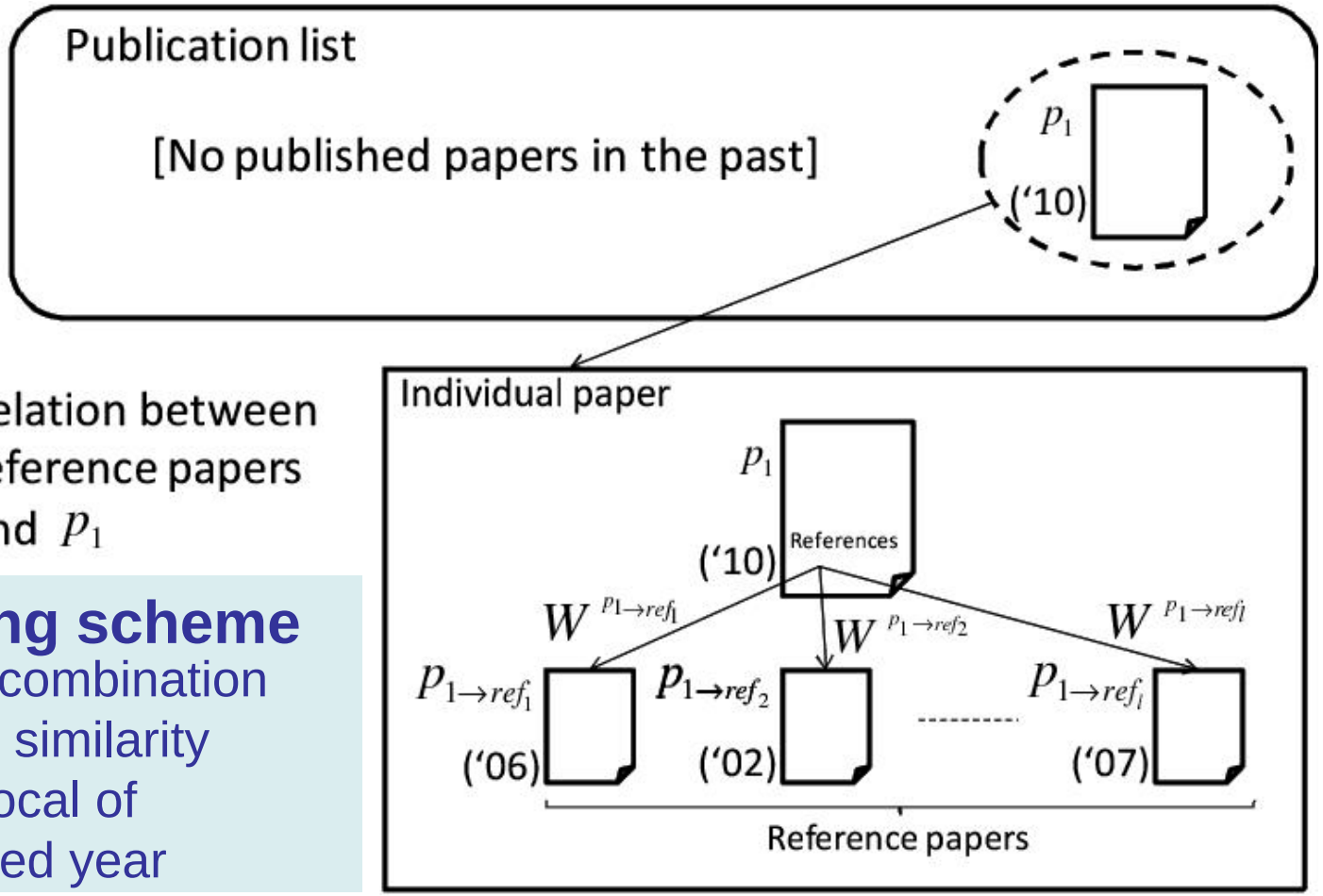


Consider only radial network from the target researcher, “Y.I. Lin”

Weighting scheme

(W1) Linear Combination (LC)
(W2) Reciprocal of Path Length (RCP-PL)
(W3) Reciprocal of Similarity (RCP-SIM)
(W4) Product of W2 and W3 (RCP-PLSIM)

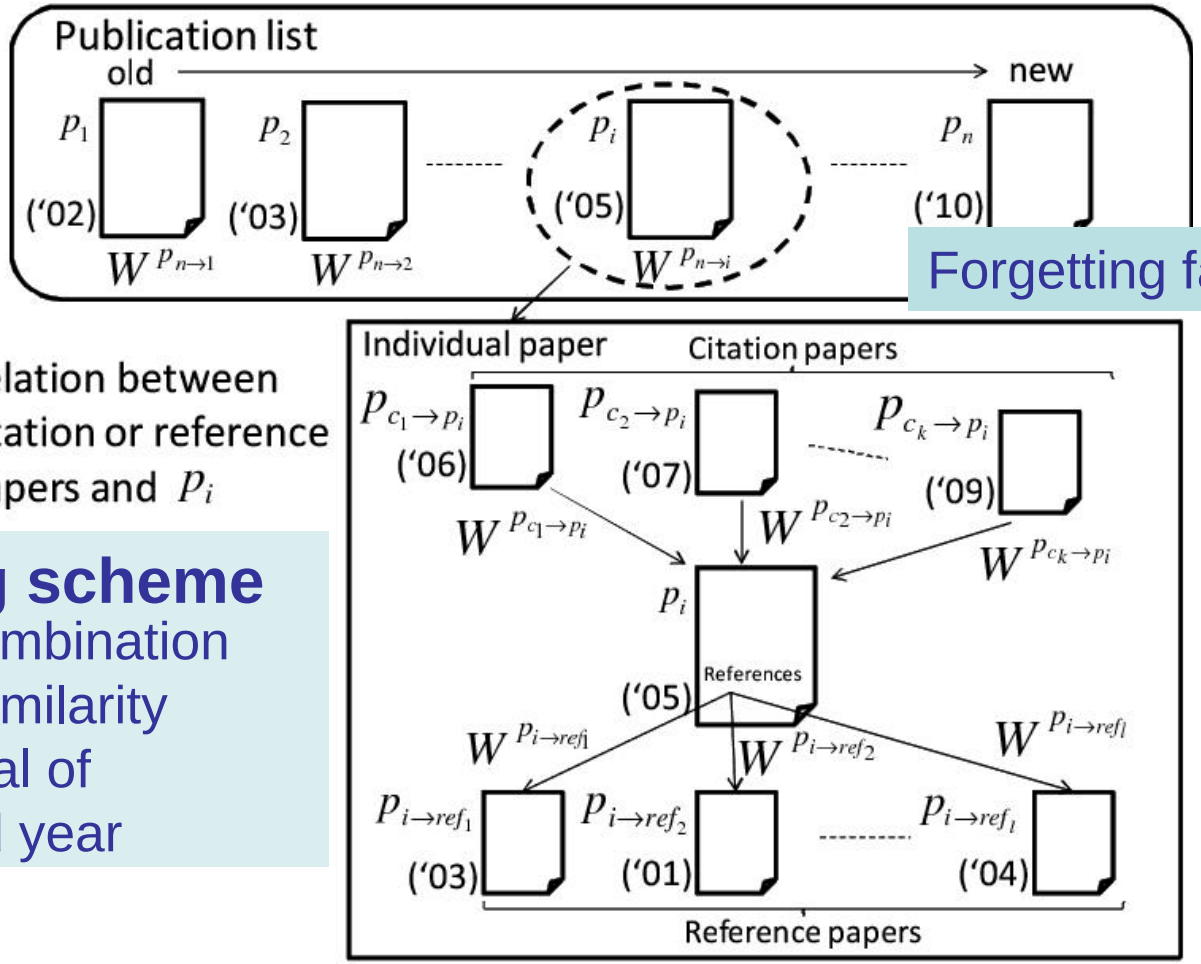
Basic User Profile Construction (Junior Researchers)



Weighting scheme

- Linear combination
- Cosine similarity
- Reciprocal of published year

Basic User Profile Construction (Senior Researchers)



Weighting scheme

- Linear combination
- Cosine similarity
- Reciprocal of published year

Experiments

Experimental Data

- **Researchers**
  - 15 junior researchers
  - 13 senior researchers

    NLP and IR researchers who have publication lists in DBLP

- **Candidate Papers to Recommend**

  ACL Anthology Reference Corpus [Bird et al., LREC'08]

  Includes information about citation and reference papers

Evaluation Measure

- Normalized Item Novelty (nITN@10)

Normalized Item Novelty

Item novelty (*ITN*) [Zhang and Hurley, RecSys'08]

$$ITN = \frac{1}{N}\sum_{j=1}^{N} d(\mathbf{P}_u^{srdp}, \mathbf{F}^{p_{rec_j}})$$

Monotone increasing

$\mathbf{P}_u^{srdp}$: User profile

$\mathbf{F}^{p_{rec_j}}$: Feature vector of the candidate paper to recommend

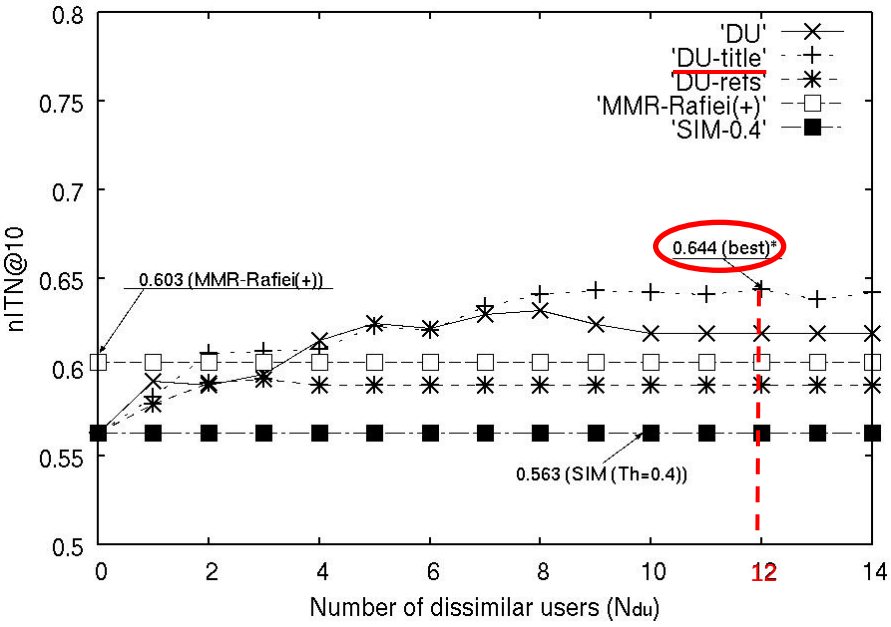
Normalized item novelty (n*ITN*)

$$nITN = \frac{1}{N}\sum_{j=1}^{N} \frac{d(\mathbf{P}_u^{srdp}, \mathbf{F}^{p_{rec_j}})}{\max d(\mathbf{P}_u^{srdp}, \mathbf{F}^{p_{rec_j}})}$$

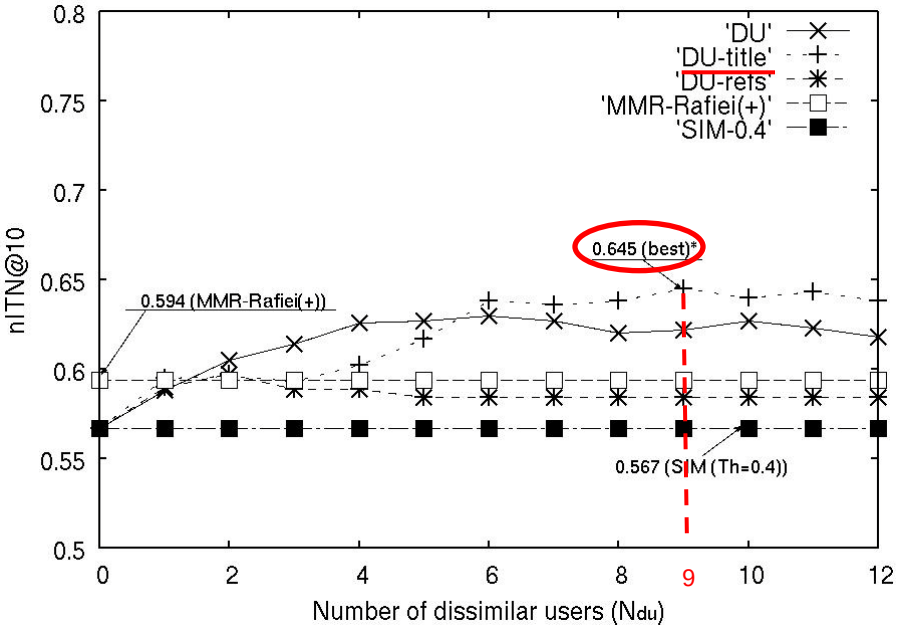
Avoid monotone increasing

Results with Dissimilar Users (DU)

[Item novelty@10]



[Junior researchers]

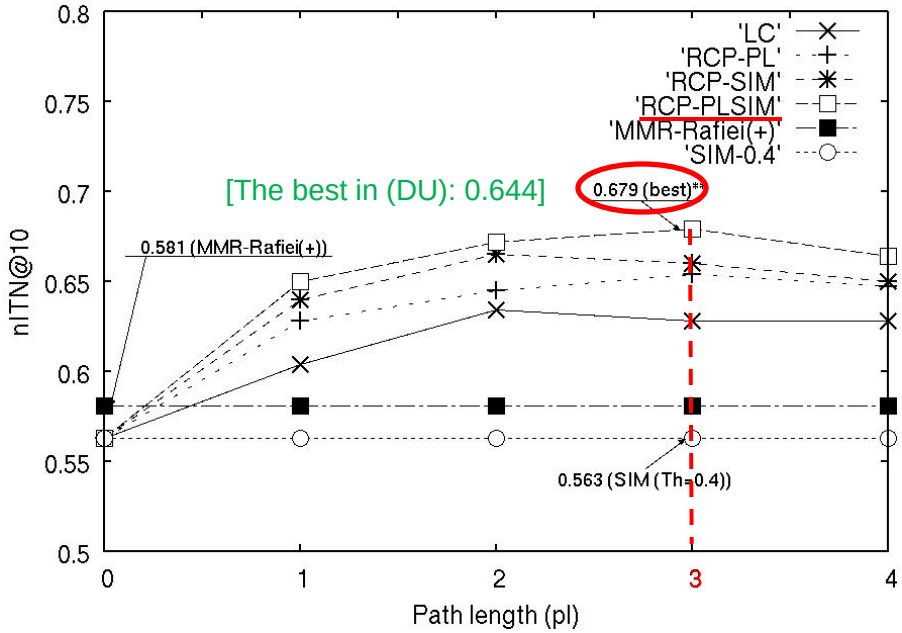


[Senior researchers]

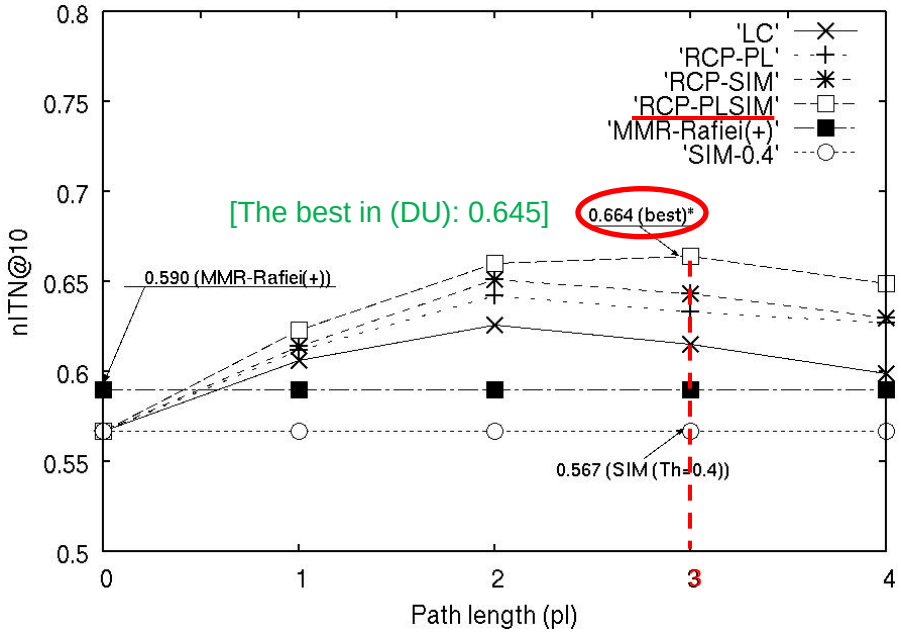
User profiles that contain a variety of topics can be constructed by using more dissimilar users.

Results with the Co-author Network (CAN)

[Item novelty@10]



[Junior researchers]



[Senior researchers]

(CAN) is more effective approach to constructing user profile for serendipitous recommendation rather than (DU).

Examples of Serendipitous Recommendation

	Recommended topics
Junior researcher (Major research topic: discourse analysis)	• Noun phrase chunking • Collocation • Term Recognition
Senior researcher (Major research topic: sentiment analysis)	• Knowledge acquisition • Relation extraction in named entities • Relation extraction in biomedical text

It is highly possible that he can discover something new and helpful.

Conclusion

- **Propose constructing user profile for serendipitous recommendation that considers relations among researchers**
  - **Dissimilar users**
  - **Co-author network**
- **Observe serendipitous nature of the recommendation**

Future Work

- **Construct user profile that provides highly accurate recommendation of serendipitous papers**
- **Expand the kinds of candidate papers to recommend**

***Thank you very much*!**