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]

Junior researcher (having only one recently published paper)

Senior researcher (having several past published papers)

Serendipitous recommendation is important.
Introduction

When we find something serendipitous?

Advice from colleagues

That’s nice idea!

How about using this technique?

Attend seminars

We can also use this approach even if the research topic is different from ours!

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 generated from published history of papers

User 1
- User 4 (Sim: 0.16)
  Weight: 1/(0.16+1)
- User 10 (Sim: 0.26)
  Weight: 1/(0.26+1)

User 2
- User 5 (Sim: 0.21)
  Weight: 1/(0.21+1)
- User 1 (Sim: 0.32)
  Weight: 1/(0.32+1)

User 3
- User 1 (Sim: 0.14)
  Weight: 1/(0.14+1)
- User 7 (Sim: 0.25)
  Weight: 1/(0.25+1)

User n
- User 6 (Sim: 0.07)
  Weight: 1/(0.07+1)
- User 2 (Sim: 0.12)
  Weight: 1/(0.12+1)
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)

Publication list

[No published papers in the past]

Relation between reference papers and $P_1$

Weighting scheme
- Linear combination
- Cosine similarity
- Reciprocal of published year
Basic User Profile Construction (Senior Researchers)

Forgetting factor

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]

\[
ITN = \frac{1}{N} \sum_{j=1}^{N} d(P_{u}^{sr dp}, F^{P_{rec_j}})
\]

\[
\begin{align*}
P_{u}^{sr dp} & : \text{User profile} \\
F^{P_{rec_j}} & : \text{Feature vector of the candidate paper to recommend}
\end{align*}
\]

Monotone increasing

Normalized item novelty (nITN)

\[
nITN = \frac{1}{N} \sum_{j=1}^{N} \frac{d(P_{u}^{sr dp}, F^{P_{rec_j}})}{\max d(P_{u}^{sr dp}, F^{P_{rec_j}})}
\]

Avoid monotone increasing
Results with Dissimilar Users (DU)

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)

(CAN) is more effective approach to constructing user profile for serendipitous recommendation rather than (DU).
### Examples of Serendipitous Recommendation

<table>
<thead>
<tr>
<th>Junior researcher (Major research topic: discourse analysis)</th>
<th>Recommended topics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Noun phrase chunking</td>
</tr>
<tr>
<td></td>
<td>• Collocation</td>
</tr>
<tr>
<td></td>
<td>• Term Recognition</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Senior researcher (Major research topic: sentiment analysis)</th>
<th>Recommended topics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Knowledge acquisition</td>
</tr>
<tr>
<td></td>
<td>• Relation extraction in named entities</td>
</tr>
<tr>
<td></td>
<td>• Relation extraction in biomedical text</td>
</tr>
</tbody>
</table>

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!