Predicting Anchor Links between Heterogeneous Social Networks

Sina Sajadmanesh
Hamid R. Rabiee
Ali Khodadadi

Digital Media Lab
Sharif University of Technology

August 2016

2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)
Emergence of New Social Networks
New networks try to invite users of older ones

Find ideas from friends

⭐ Best ideas for you  ⬇ Add more friends

Try adding a few more friends to help us pick better boards for you

Find and invite anyone

Enter name or email
New networks try to invite users of older ones

Invitations can be more intelligent!
Networks connect to each other

He used Pinterest to dive deep
Welcome to Pinterest, the world’s catalog of ideas

Continue with Facebook

or

Email
Create a password

Sign up

Creating an account means you’re okay with
Pinterest’s Terms of Service, Privacy Policy

50+ billion ideas to explore | 15 seconds to sign up (free!)
Networks connect to each other
Some information is shared between networks.
Networks can use information from each other.

INFO YOU PROVIDE TO THIS APP:

Public profile (required)
Sina Sajadmanesh, profile picture, 21+ years old, male and other public info

Friend list
Navid Ḵazaei, Mohammad Ḵazaei, and 186 others

Email address
sina.sajadmanesh@gmail.com
Problem: Anchor Link Prediction

Who is likely to join target network?
Problem: Anchor Link Prediction

By joining new users, new anchor links emerge.
Why do users join a new network?

**Personal Factors**
- Attractiveness of the target network
- Dissatisfaction from the source network

**Social Factors**
- Number of friends in the target network
- Intimacy between friends in the target network
Why do users join a new network?

- Attractiveness of the target network
- Dissatisfaction from the source network
- Number of friends in the target network
- Intimacy between friends in the target network

Factors:
- Personal Factors
- Social Factors
Meta-Path-Based Approach

**Network Schema:**
- Meta structure of a heterogeneous network

**Meta-Path:**
- A path defined over network schema

Sample Meta-Path:

Sample Network Schema:

Heterogeneous information
Number of friends in the target network

#instances of this meta-path = #friends in the target network

Will he join?
Connector Meta-Paths

Similarity Extension

Similar users perform similar actions

Will he join?
How to model similarity?

Similarity meta-paths

Social

Spatial

Temporal

Textual
“similarity” as a measure of “intimacy”

Intimacy between friends in the target network

Will he join?
Recursive Meta-Paths

Similarity Extension again

This kind of meta-path begins with a user and returns to himself

Will he join?
Classification

9 connector meta-paths
Path-Count
9 features
Path-Count
9x9x9 recursive meta-paths
729 features
Dataset

Twitter as source network:
- Containing about 5k users with a total of 8m tweets

Foursquare as target network:
- Containing about 3.5k users with a total of 48k tips

Common users:
- About 3k shared users

Ground truth:
- Positive Samples: Common users who joined twitter before foursquare
- Negative Samples: Non-anchor users
Experiment Settings

- **Comparison methods:**
  - **CICF:** consistent incidence co-factorization
  - **CMP:** connector meta-paths only
  - **RMP:** recursive meta-paths only
  - **CRMP:** both connector and recursive meta-paths

- **Experiment setup:**
  - 1936 positive samples
  - 1941 negative samples
  - 5-fold cross-validation with linear SVM
Experiment Results

Effect of heterogeneous information

Accuracy

<table>
<thead>
<tr>
<th></th>
<th>CMP</th>
<th>RMP</th>
<th>CRMP</th>
<th>CICF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogeneous</td>
<td>0.70</td>
<td>0.65</td>
<td>0.70</td>
<td>0.50</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>0.75</td>
<td>0.72</td>
<td>0.80</td>
<td>0.60</td>
</tr>
</tbody>
</table>

AUC

<table>
<thead>
<tr>
<th></th>
<th>CMP</th>
<th>RMP</th>
<th>CRMP</th>
<th>CICF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogeneous</td>
<td>0.70</td>
<td>0.65</td>
<td>0.70</td>
<td>0.50</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>0.75</td>
<td>0.72</td>
<td>0.80</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Effect of remaining anchor links

Accuracy

AUC
Effect of newness of target network

Accuracy

AUC
Experiment Results

Effect of similarity extension

Accuracy

AUC
Conclusion

Problem:

- Anchor link prediction
- Different from conventional link prediction

Method:

- A meta-path-based approach
- Connector and Recursive meta-paths model different aspects of social factors

Future Works:

- To model personal factors as well
- To predict the time of link creation
Thank You!

Any Questions?

sajadmanesh@ce.sharif.edu
rabiee@sharif.edu
khodadadi@ce.sharif.edu