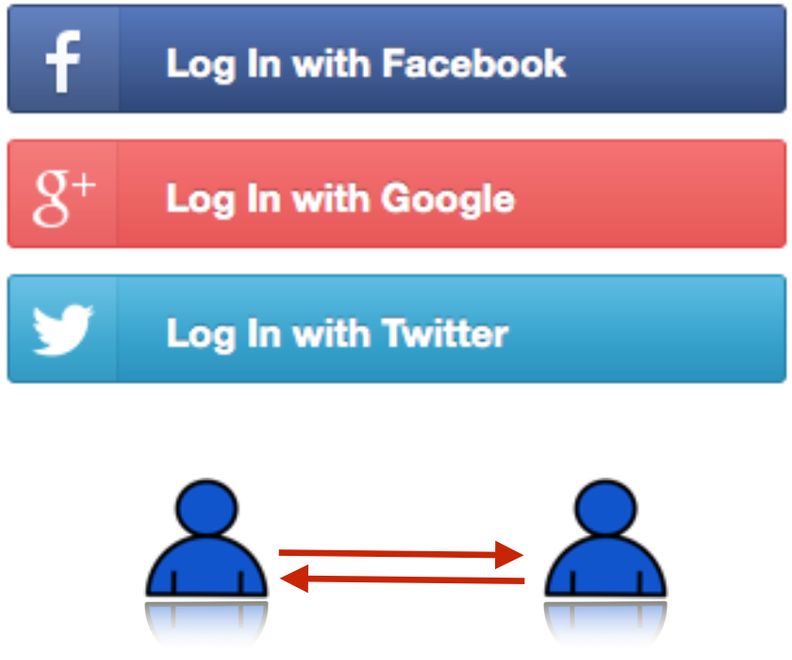
Social Bootstrapping

How **Pinterest** and **Last.fm** Social Communities Benefit by Borrowing Links from **Facebook**

C. Zhong[1], M. Salehi[2], S. Shah[3], M. Cobzarenco[4], N. Sastry[1], M. Cha[5]

[1]**King's College London,** [2]University of Tehran, [3]UC Berkeley, [4]Last.fm, [5]KAIST

10 April 2014



Social Bootstrapping

How **Pinterest** and **Last.fm** Social Communities Benefit by Borrowing Links from **Facebook**

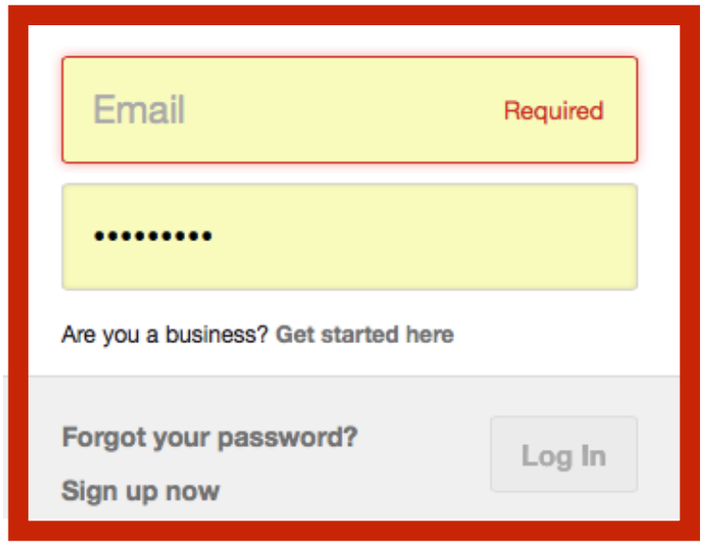
C. Zhong[1], M. Salehi[2], S. Shah[3], M. Cobzarenco[4], N. Sastry[1], M. Cha[5]

[1]**King's College London,** [2]University of Tehran, [3]UC Berkeley, [4]Last.fm, [5]KAIST

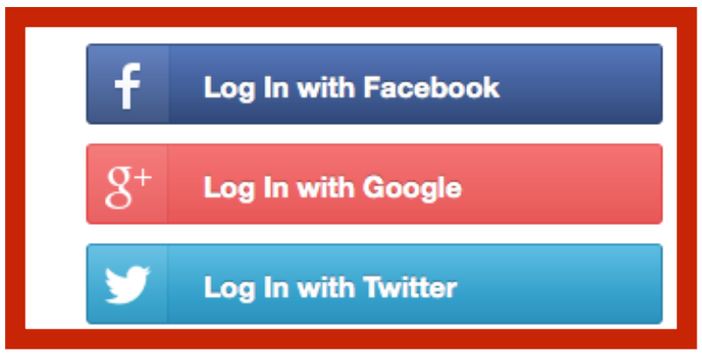
10 April 2014

The Dilemma for New Websites: How to construct social network?

- Option 1:
  - Create entirely new social network
- Option 2:
  - Social Bootstrapping



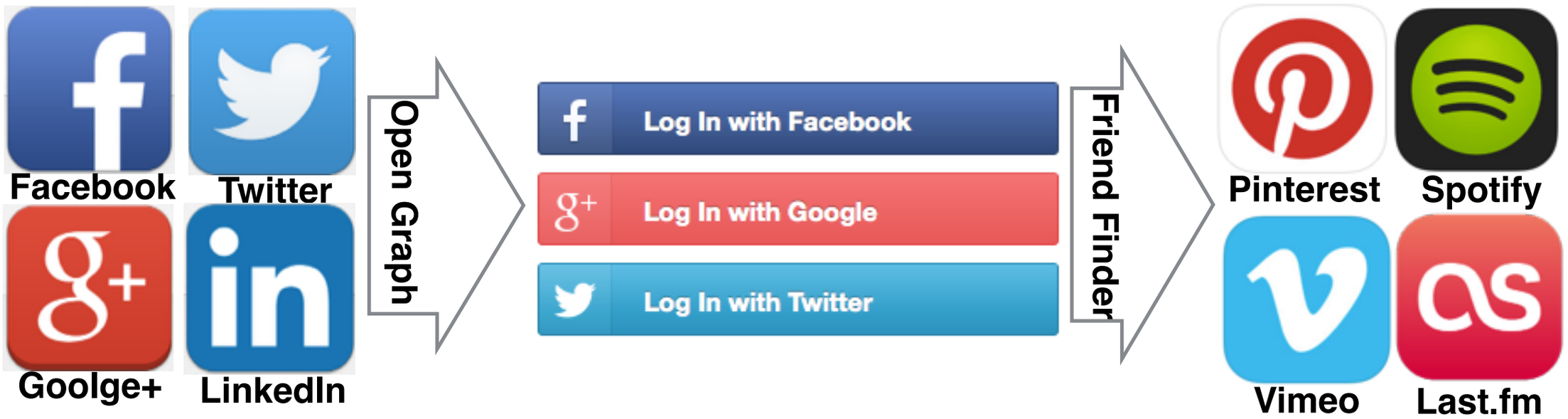
Option 1



Option 2

Social Bootstrapping

- The process of **copying** links from established social networks (*source network*) onto a third-party website (*target network*).



Source Network

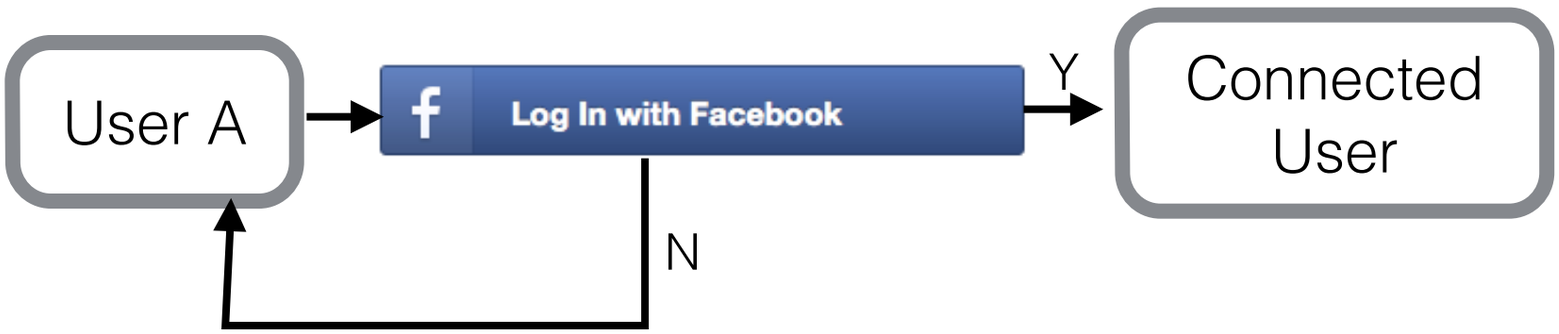
Target Network

Social Bootstrapping in action: Friend Finder in Pinterest

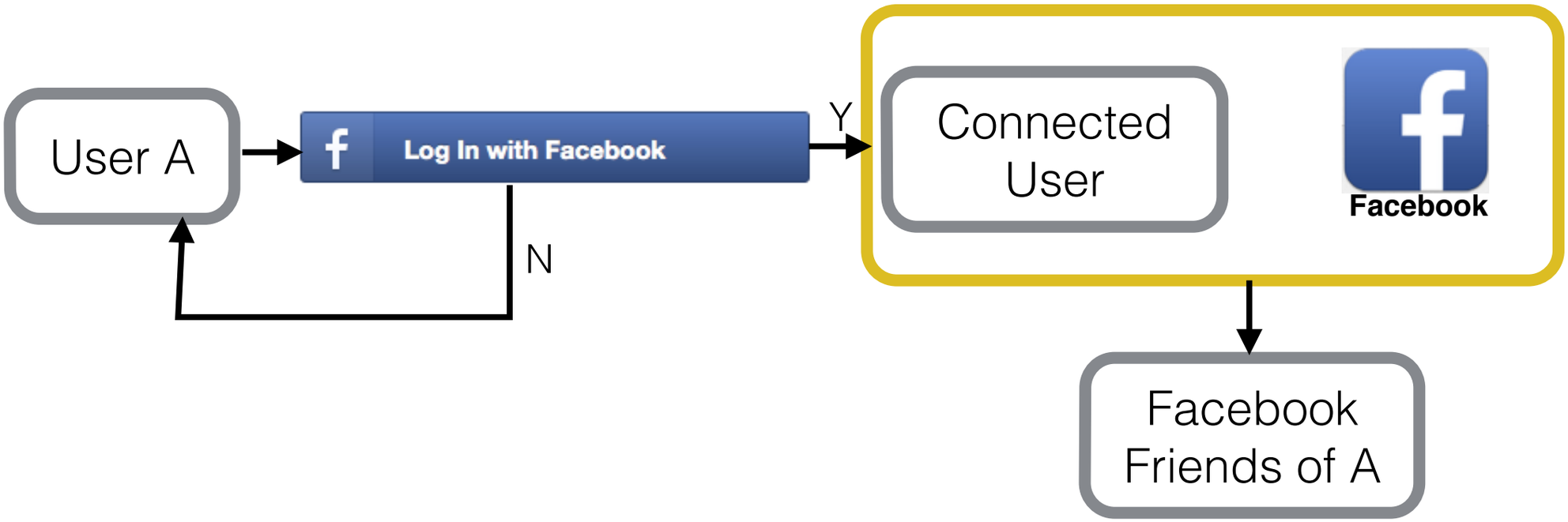
Social Bootstrapping in action: Friend Finder in Pinterest

User A

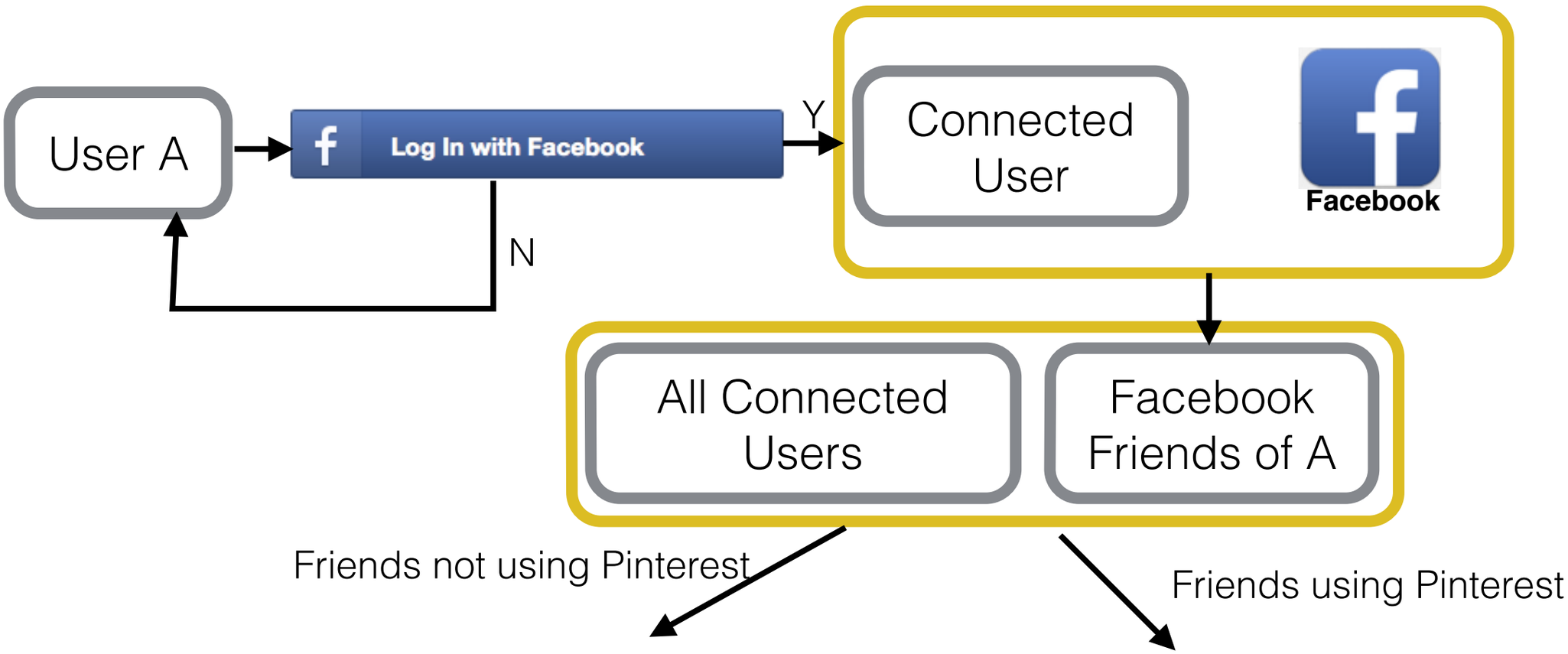
Social Bootstrapping in action: Friend Finder in Pinterest



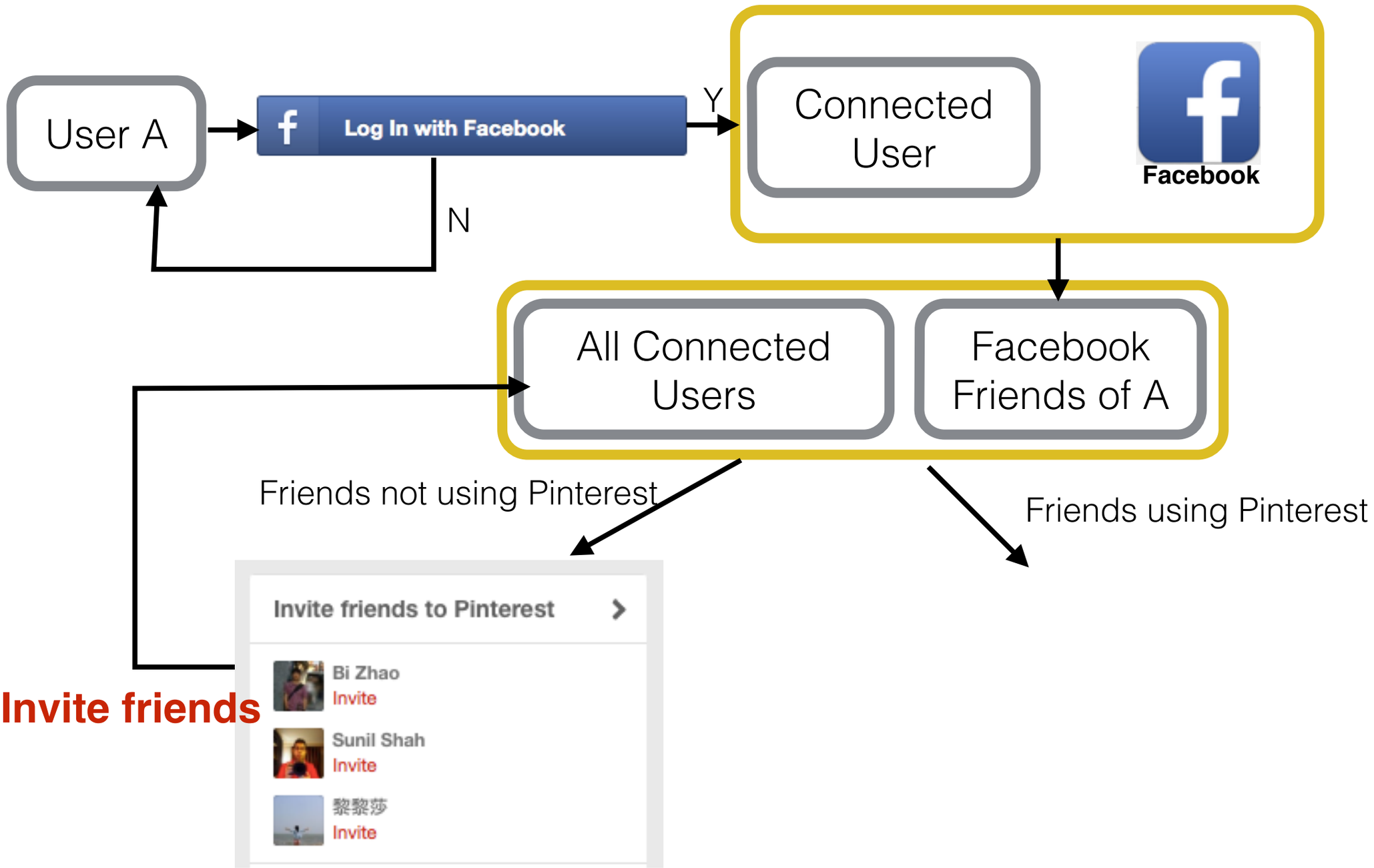
Social Bootstrapping in action: Friend Finder in Pinterest



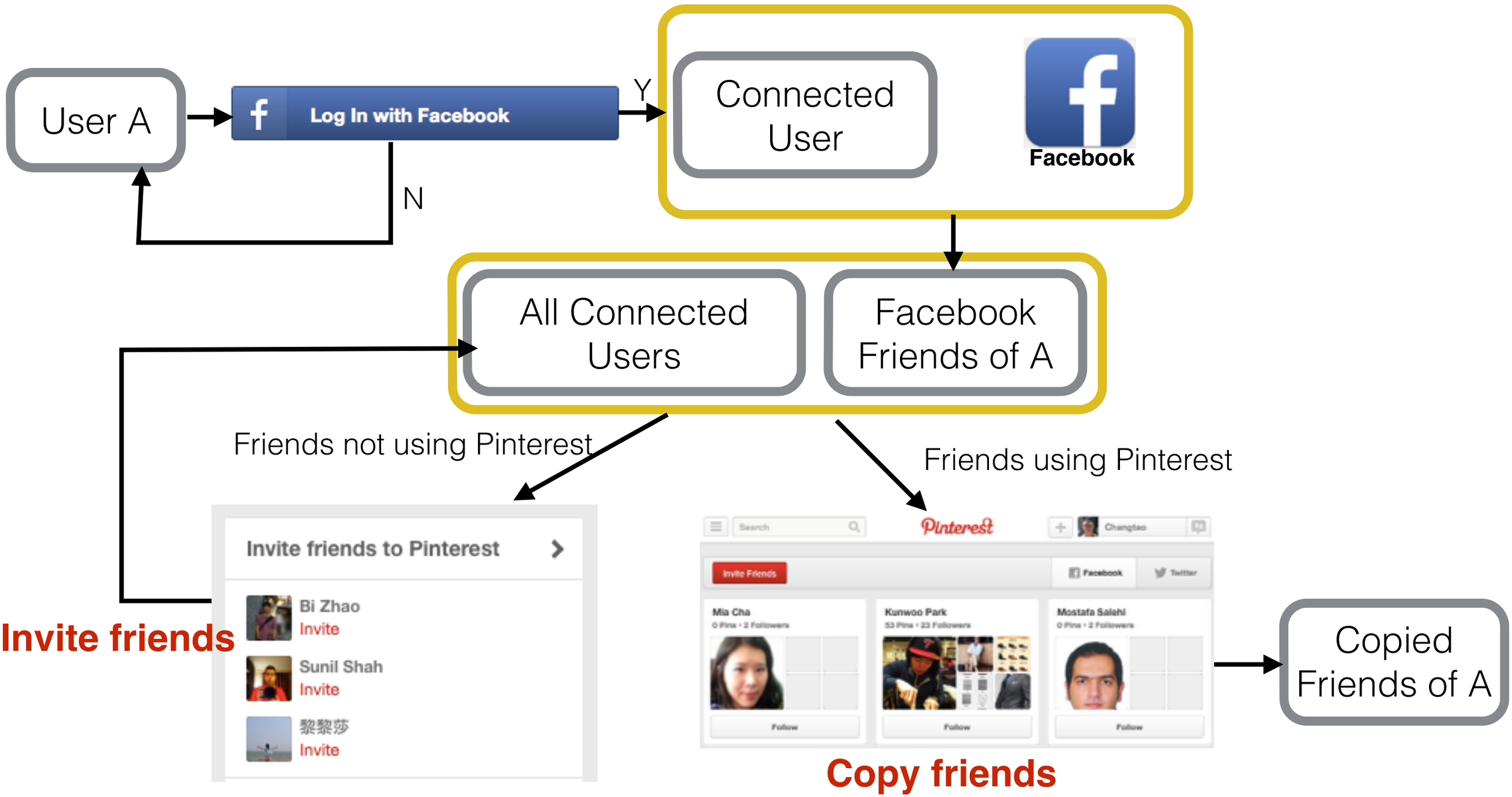
Social Bootstrapping in action: Friend Finder in Pinterest



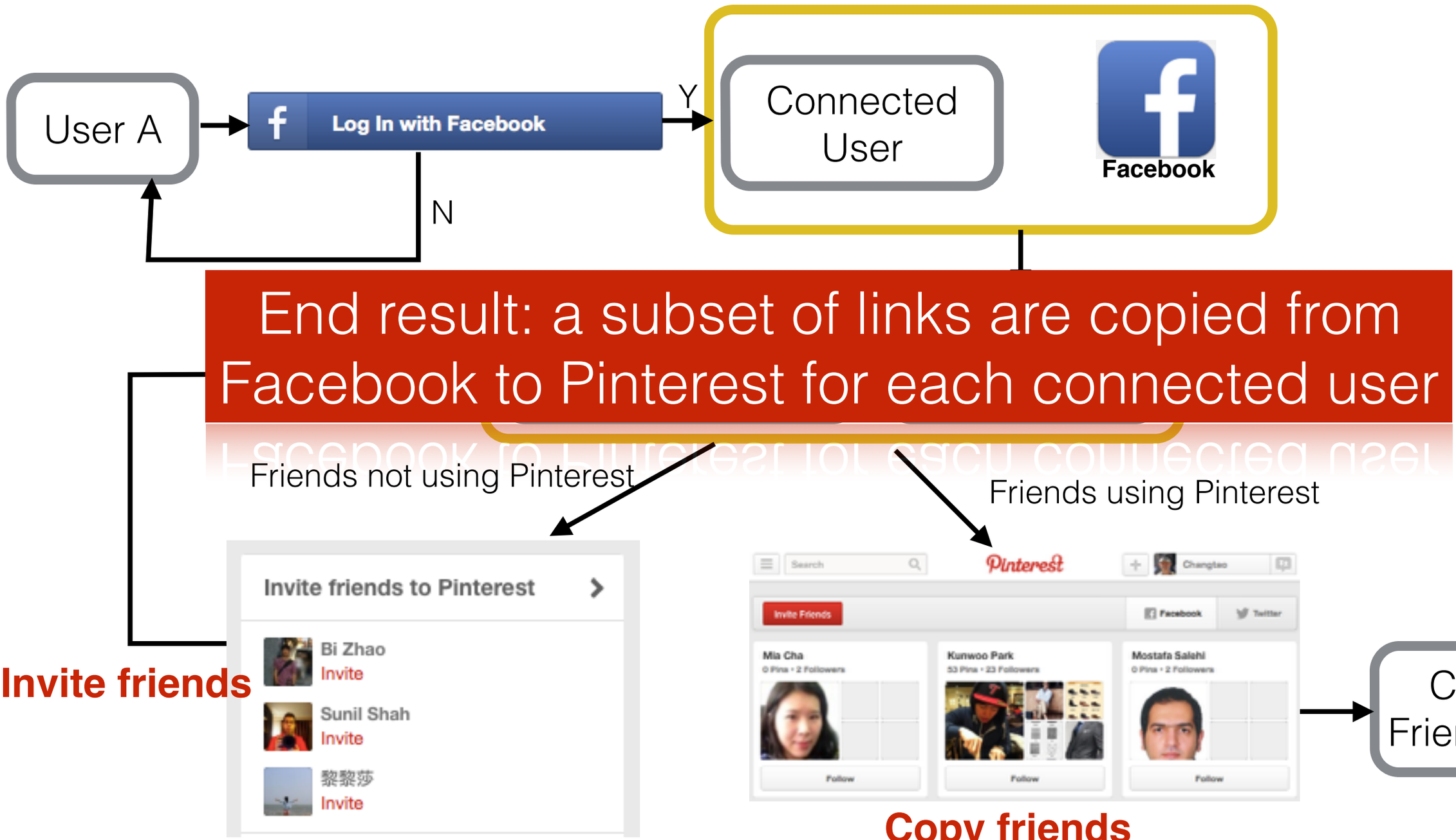
Social Bootstrapping in action: Friend Finder in Pinterest



Social Bootstrapping in action: Friend Finder in Pinterest

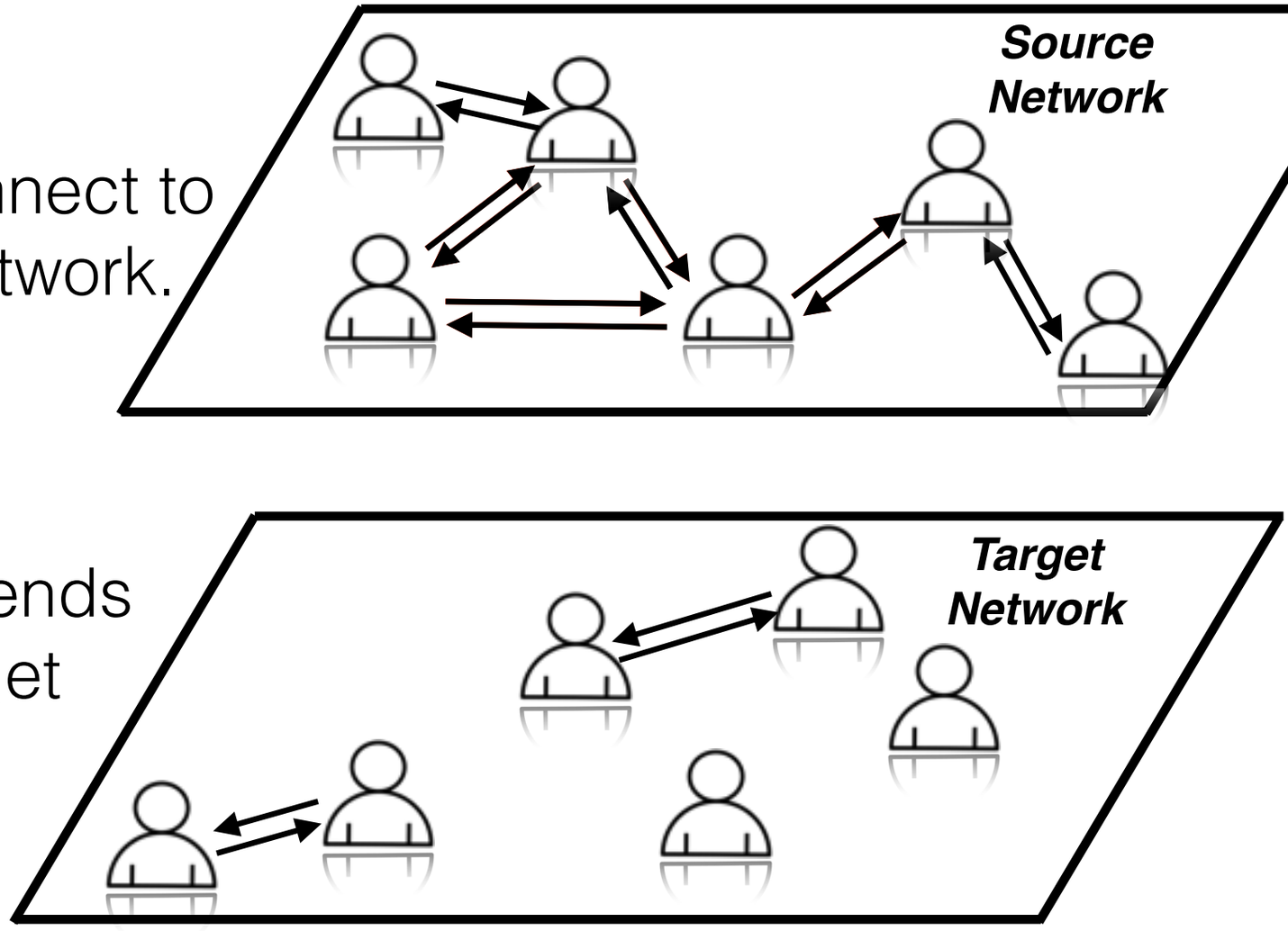


Social Bootstrapping in action: Friend Finder in Pinterest



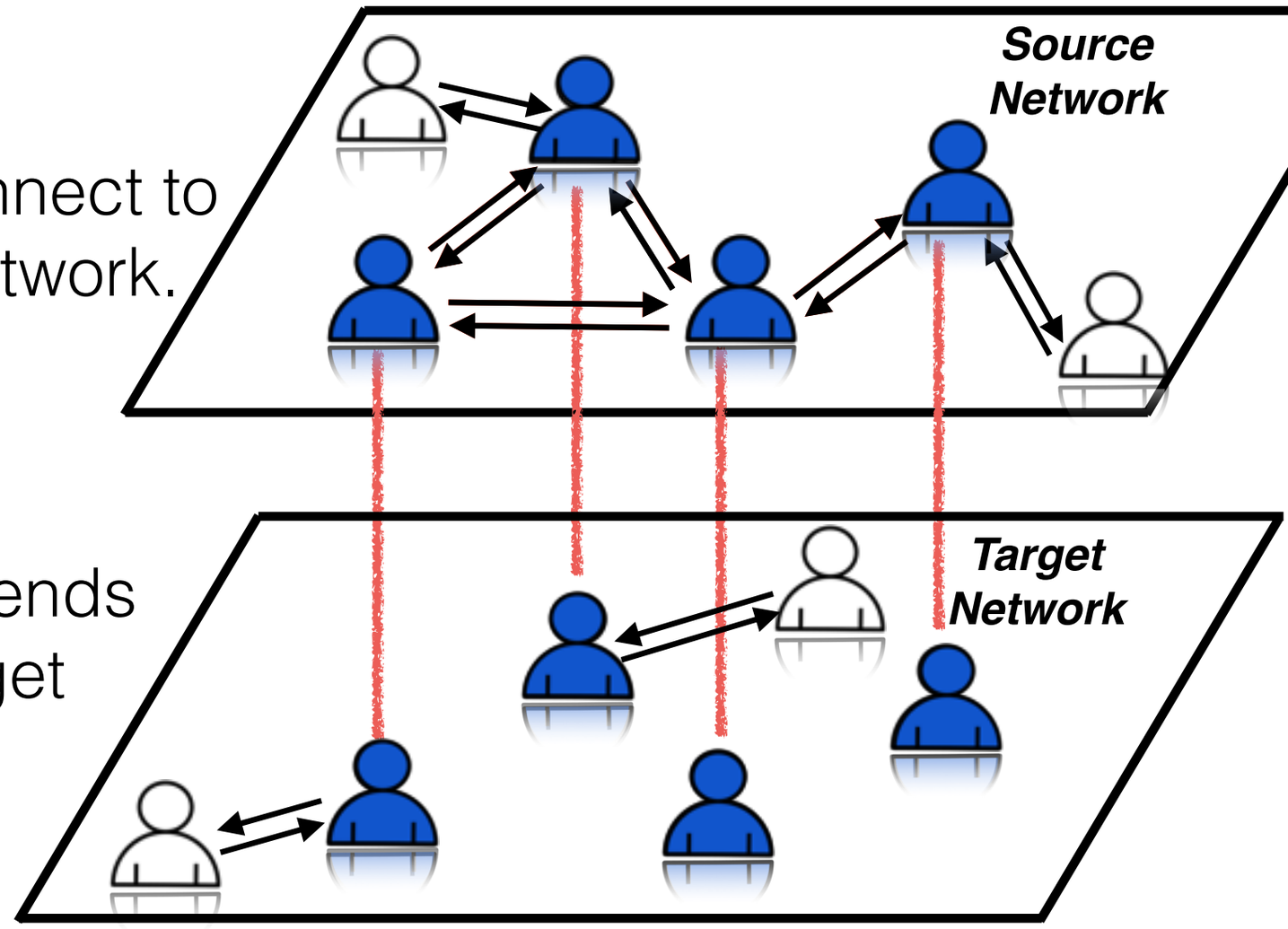
Analytical Model: Link Bootstrapping Sampling

- **Node sampling**:
  - Users in target network connect to their accounts in source network.
- **Link sampling**:
  - Connected users import friends from source network to target network.



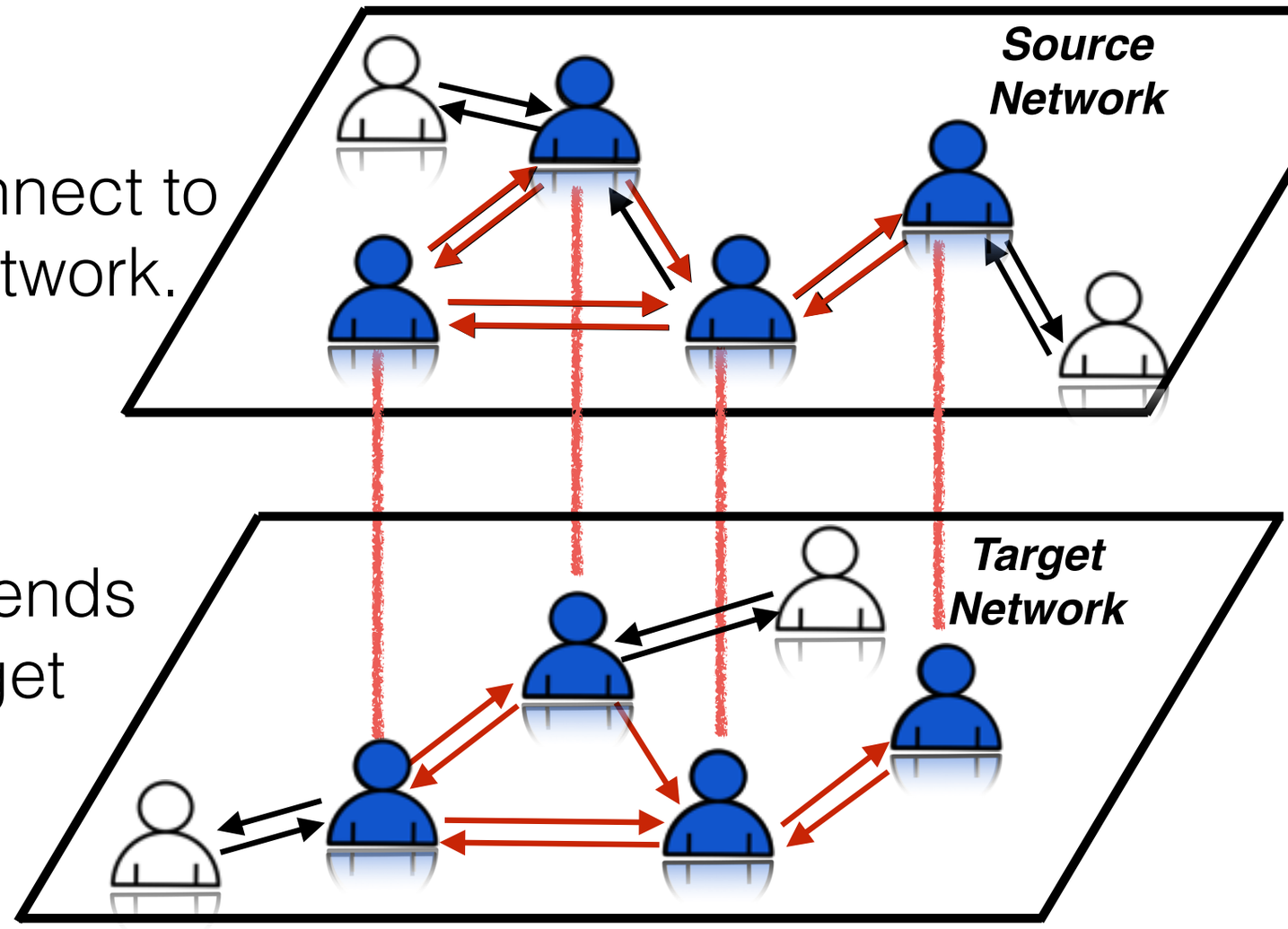
Analytical Model: Link Bootstrapping Sampling

- **Node sampling**:
  - Users in target network connect to their accounts in source network.
- **Link sampling**:
  - Connected users import friends from source network to target network.

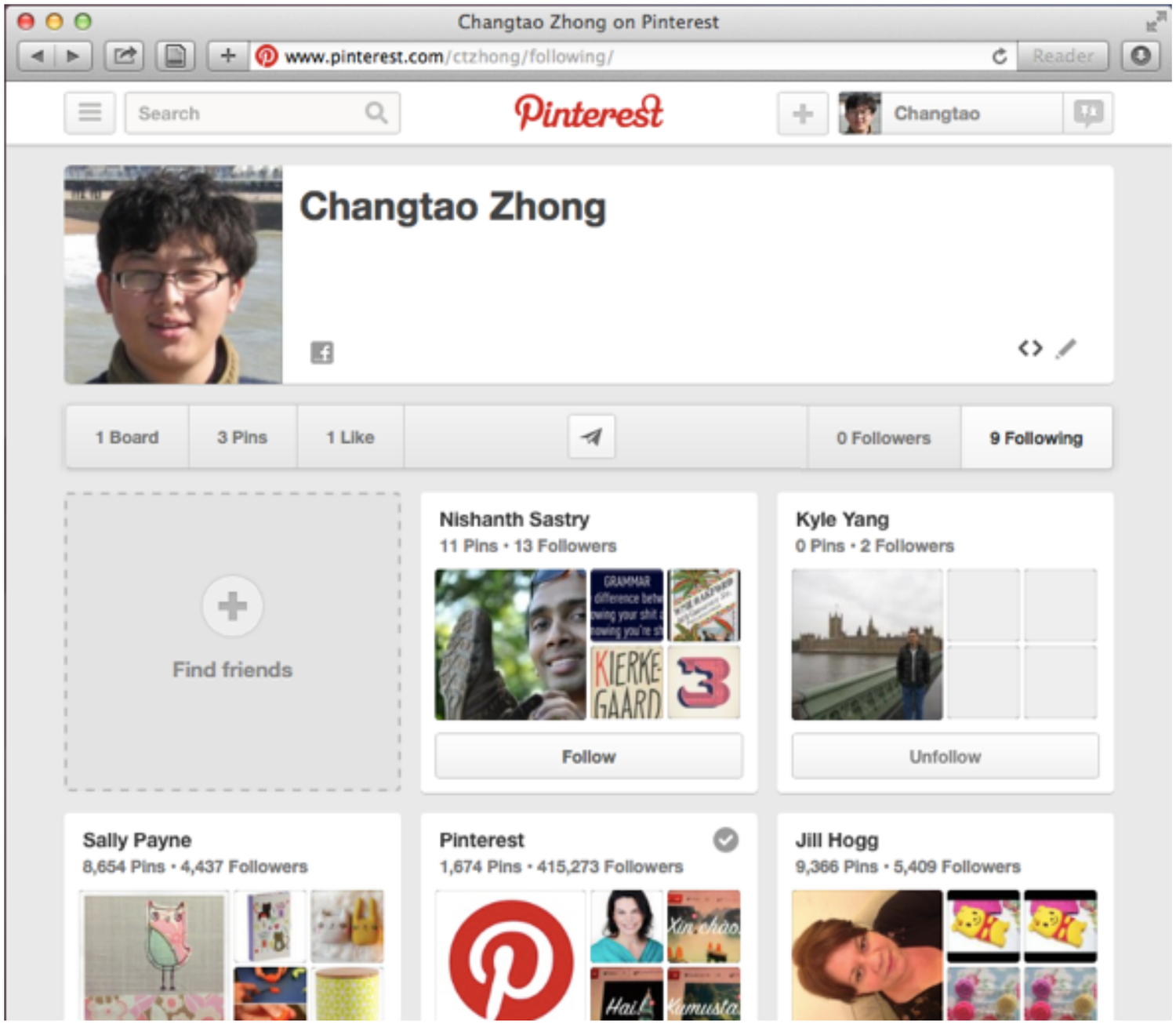


Analytical Model: Link Bootstrapping Sampling

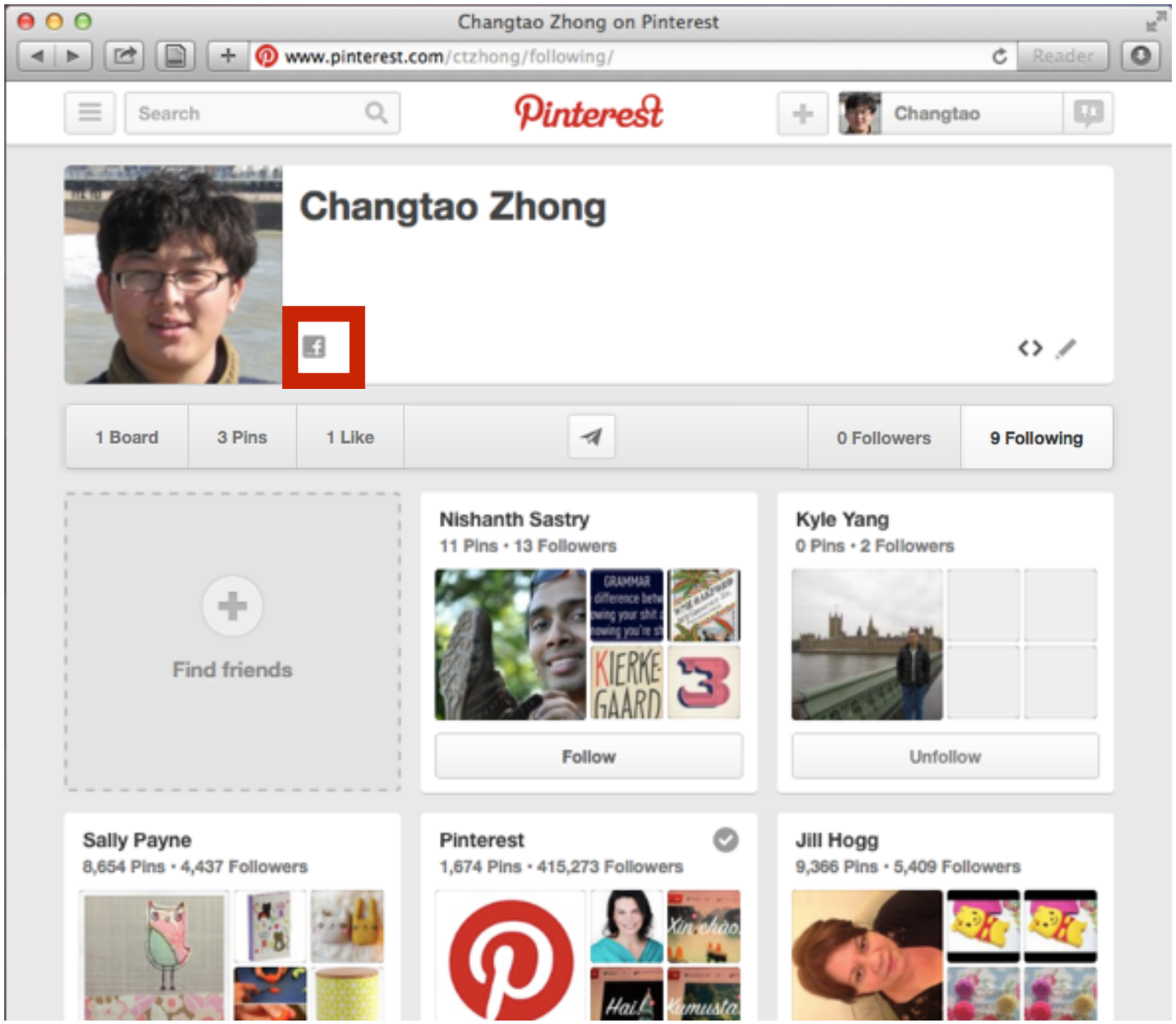
- **Node sampling**:
  - Users in target network connect to their accounts in source network.
- **Link sampling**:
  - Connected users import friends from source network to target network.



Collecting Pinterest Data

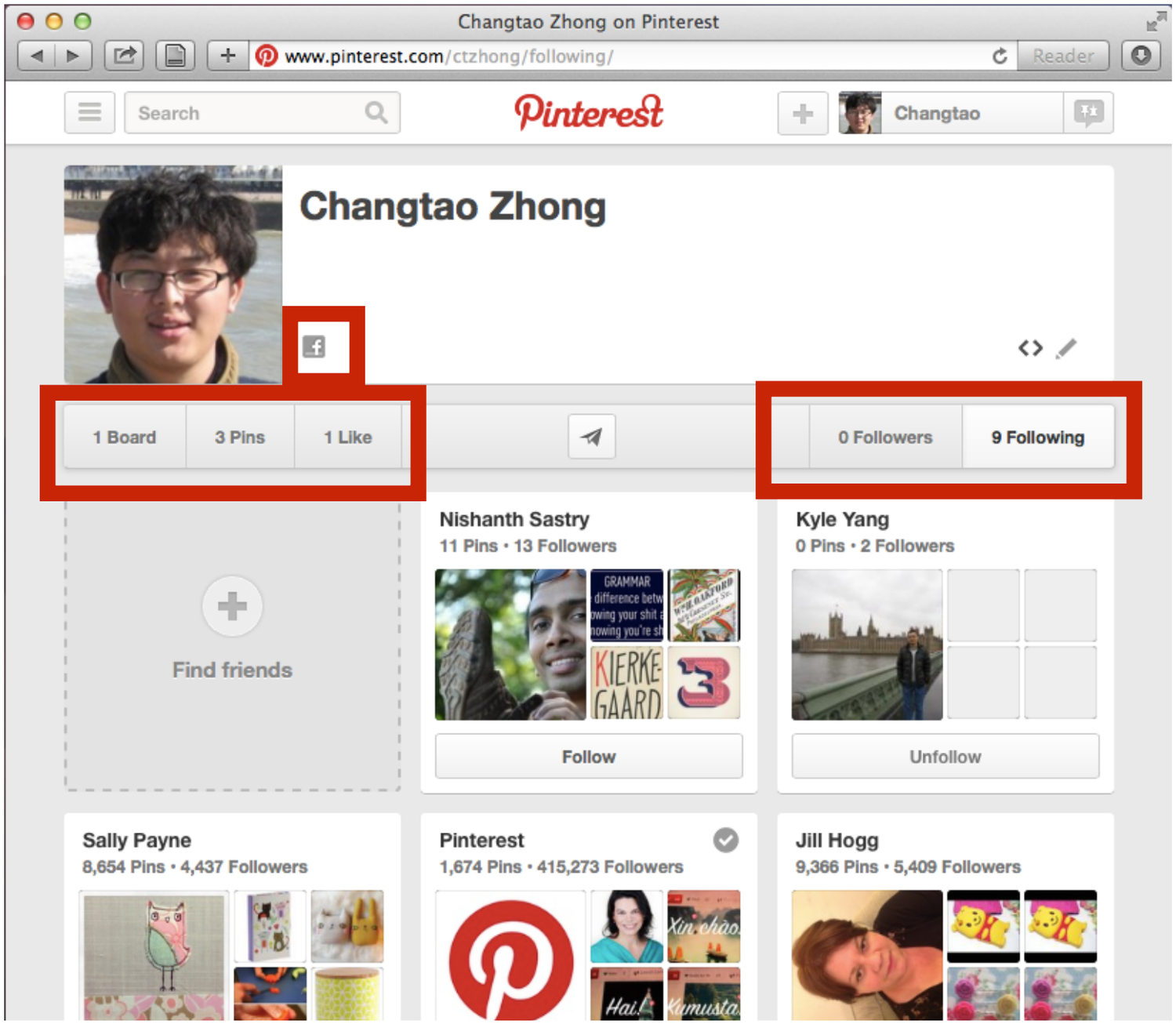


Collecting Pinterest Data

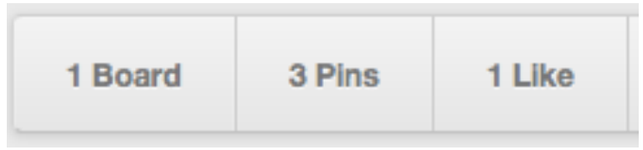


Facebook handle

Collecting Pinterest Data



Facebook handle



Activity info

Social Network

Pinterest Dataset

Social Curation Dataset

www.inf.kcl.ac.uk/staff/nrs/projects/cd-gain/dataset.html

Reader

{Social}Curation

HOME RESEARCH DATASETS PAPERS

Dataset

Terms

Datasets

An anonymized version of the Pinterest dataset used in our WWW14 and ICWSM13 papers is being made available to the research community. **If you are interested in using this data, please send us an email at netsys@kcl-dot-AC-dot-uk to get the link where you can download the data.** Note that sending the email indicates that you accept our terms and conditions in the following section. Please indicate which of following parts of the dataset you need in the email.

1. **Pinterest network**: A snowball sampled social graph of Pinterest, crawled in Apr 2013.
2. **Facebook network**: The Facebook social graphs of users who appear in the Pinterest activites dataset below.
3. **Fb-copied network**: The subset of Pinterest network that only contains links common to both Pinterest and Facebook.
4. **User information**: The basic statistics (such as the number of pins, likes, followees and followers) of Pinterest users.
5. **Pinterest activities**: Repin and like activity in Pinterest during 03-21 Jan, 2013.

Pinterest Dataset

Social Curation Dataset

www.inf.kcl.ac.uk/staff/nrs/projects/cd-gain/dataset.html

Reader

{Social}Curation

HOME RESEARCH DATASETS PAPERS

Dataset

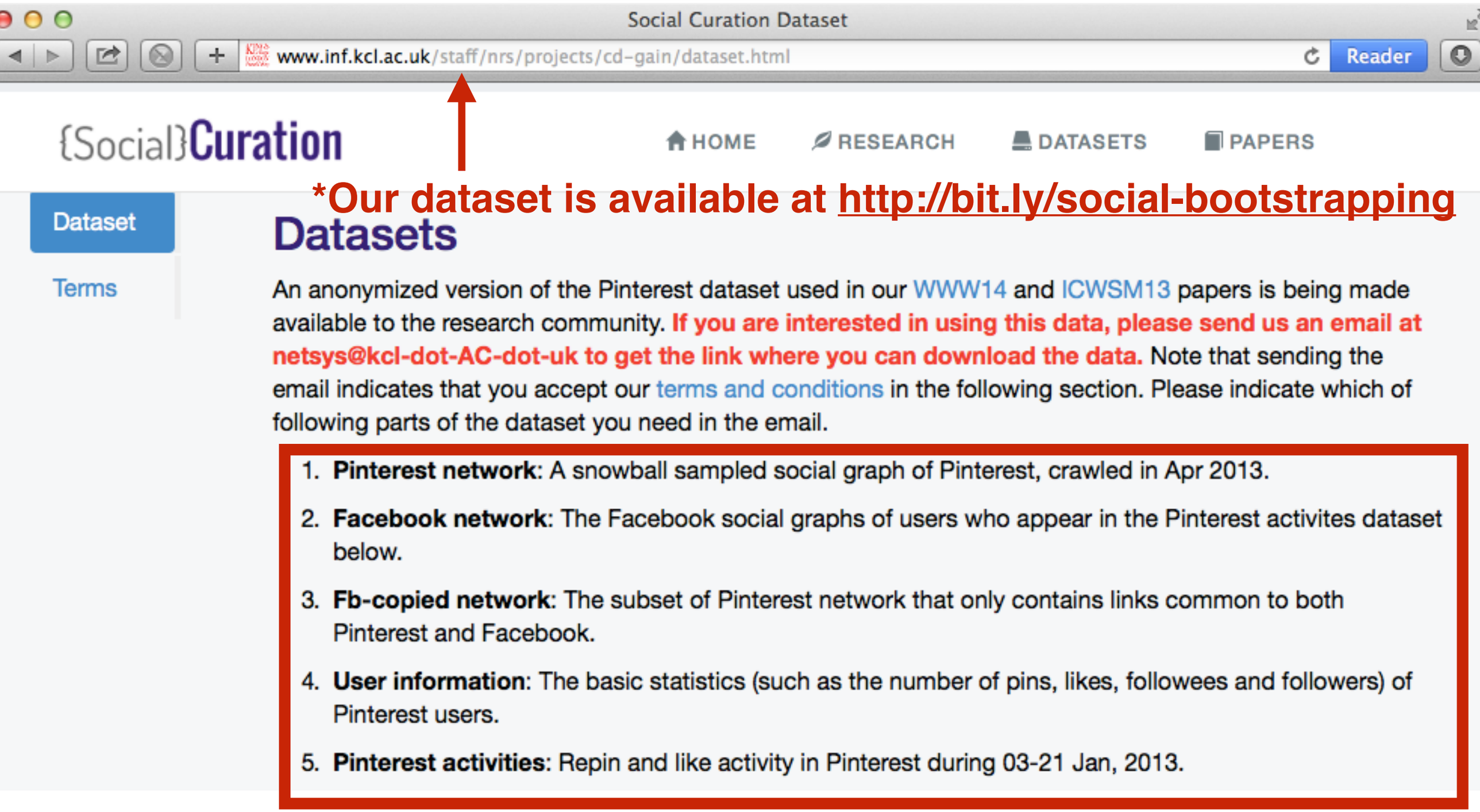
Terms

Datasets

An anonymized version of the Pinterest dataset used in our WWW14 and ICWSM13 papers is being made available to the research community. **If you are interested in using this data, please send us an email at netsys@kcl-dot-AC-dot-uk to get the link where you can download the data.** Note that sending the email indicates that you accept our terms and conditions in the following section. Please indicate which of following parts of the dataset you need in the email.

1. **Pinterest network**: A snowball sampled social graph of Pinterest, crawled in Apr 2013.
2. **Facebook network**: The Facebook social graphs of users who appear in the Pinterest activites dataset below.
3. **Fb-copied network**: The subset of Pinterest network that only contains links common to both Pinterest and Facebook.
4. **User information**: The basic statistics (such as the number of pins, likes, followees and followers) of Pinterest users.
5. **Pinterest activities**: Repin and like activity in Pinterest during 03-21 Jan, 2013.

Pinterest Dataset



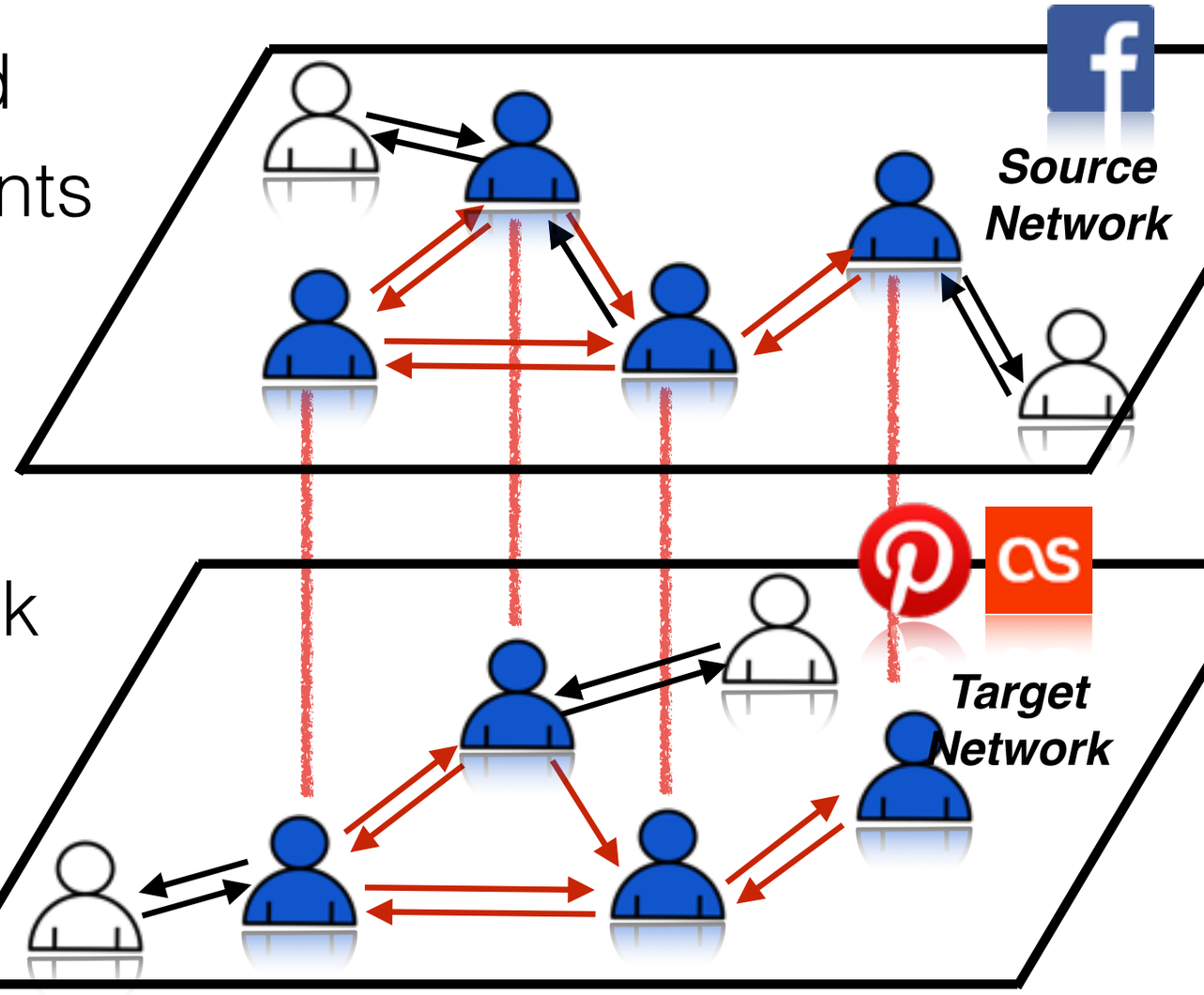
Datasets*: Pinterest and Last.fm

- **Connected users**
  - Users that have connected with their Facebook accounts

    40m / 68m 600k

- **Copied links**
  - Links copied from Facebook

    1b / 3.8b 2.8m



***Our dataset is available at http://bit.ly/social-bootstrapping**

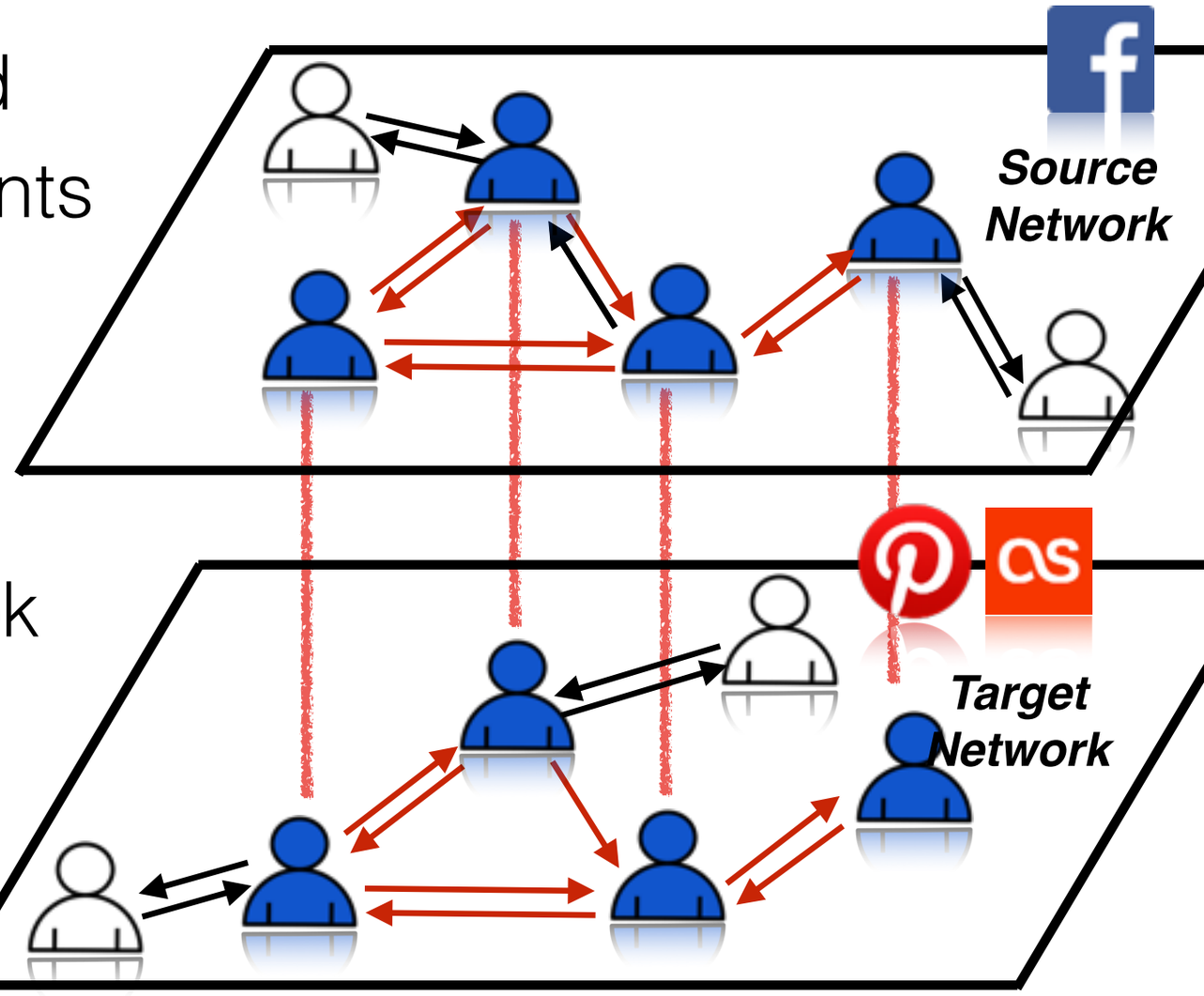
Datasets*: Pinterest and Last.fm

- **Connected users**
  - Users that have connected with their Facebook accounts

40m / 68m 600k

- **Copied links**
  - Links copied from Facebook

1b / 3.8b 2.8m



***Our dataset is available at http://bit.ly/social-bootstrapping**

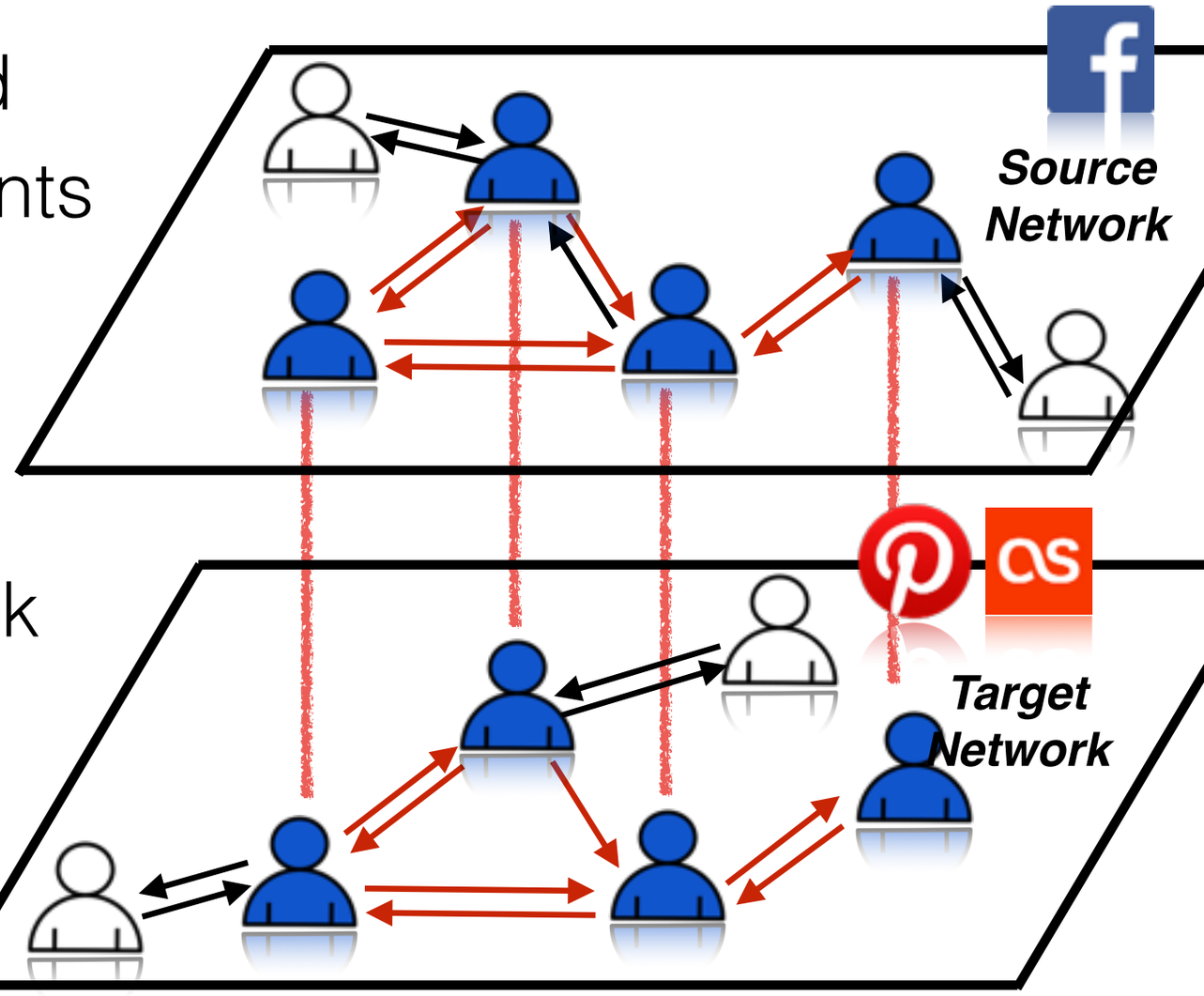
Datasets*: Pinterest and Last.fm

- **Connected users**
  - Users that have connected with their Facebook accounts

40m / 68m 600k

- **Copied links**
  - Links copied from Facebook

1b / 3.8b 2.8m



***Our dataset is available at http://bit.ly/social-bootstrapping**

Social Bootstrapping has advantages on paper

✓ Can instantly bootstrap from a mature network.

- Facebook has 10 years of history; Twitter 8 years.[1]

✓ Not "yet another" network fighting for user attention

- 71% of online adults are now Facebook users[2]

[1]Wikipedia. http://en.wikipedia.org/wiki/Facebook, http://en.wikipedia.org/wiki/Twitter
[2]Pew Research. http://www.pewinternet.org/2013/12/30/social-media-update-2013/

Social Bootstrapping has advantages on paper

✓ Can instantly bootstrap from a mature network.

- Facebook has 10 years of history; Twitter 8 years.[1]

✓ Not "yet another" network fighting for user attention

- 71% of online adults are now Facebook users[2]

BUT: Different networks have different purposes. Does copying create a good social community on the target website?

[1]Wikipedia. http://en.wikipedia.org/wiki/Facebook, http://en.wikipedia.org/wiki/Twitter
[2]Pew Research. http://www.pewinternet.org/2013/12/30/social-media-update-2013/

Outline

- *Q: Does copying create a good social community on the target website?*
  - **Structural Benefits:** Copying helps users get started with "good" structure with more social interactions
  - **"Weaning"**: Beyond Bootstrapping, active and influential users wean from Facebook to create new links natively

Structural Benefits

Our analyses show that:

- **Reciprocity:** *Copied > Native*
- **Clustering:** *Copied > Native*
- **Connectivity:** *Giant Connected Component appears easily*

Structural Benefits

Our analyses show that:

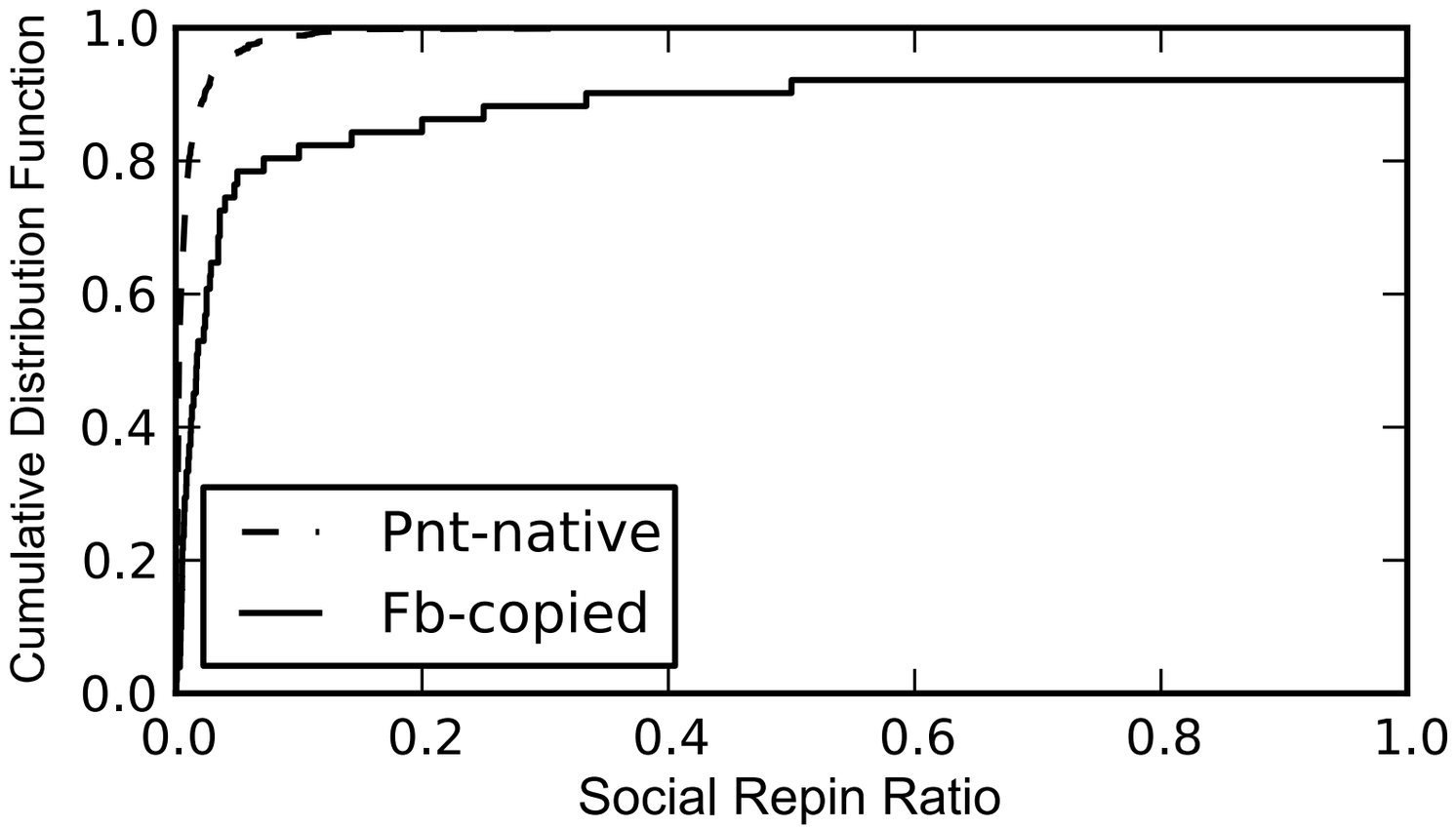
- **Reciprocity:** *Copied > Native*
- **Clustering:** *Copied > Native*
- **Connectivity:** *Giant Connected Component appears easily*

Copying links results in a **stronger** and **denser** social structure.

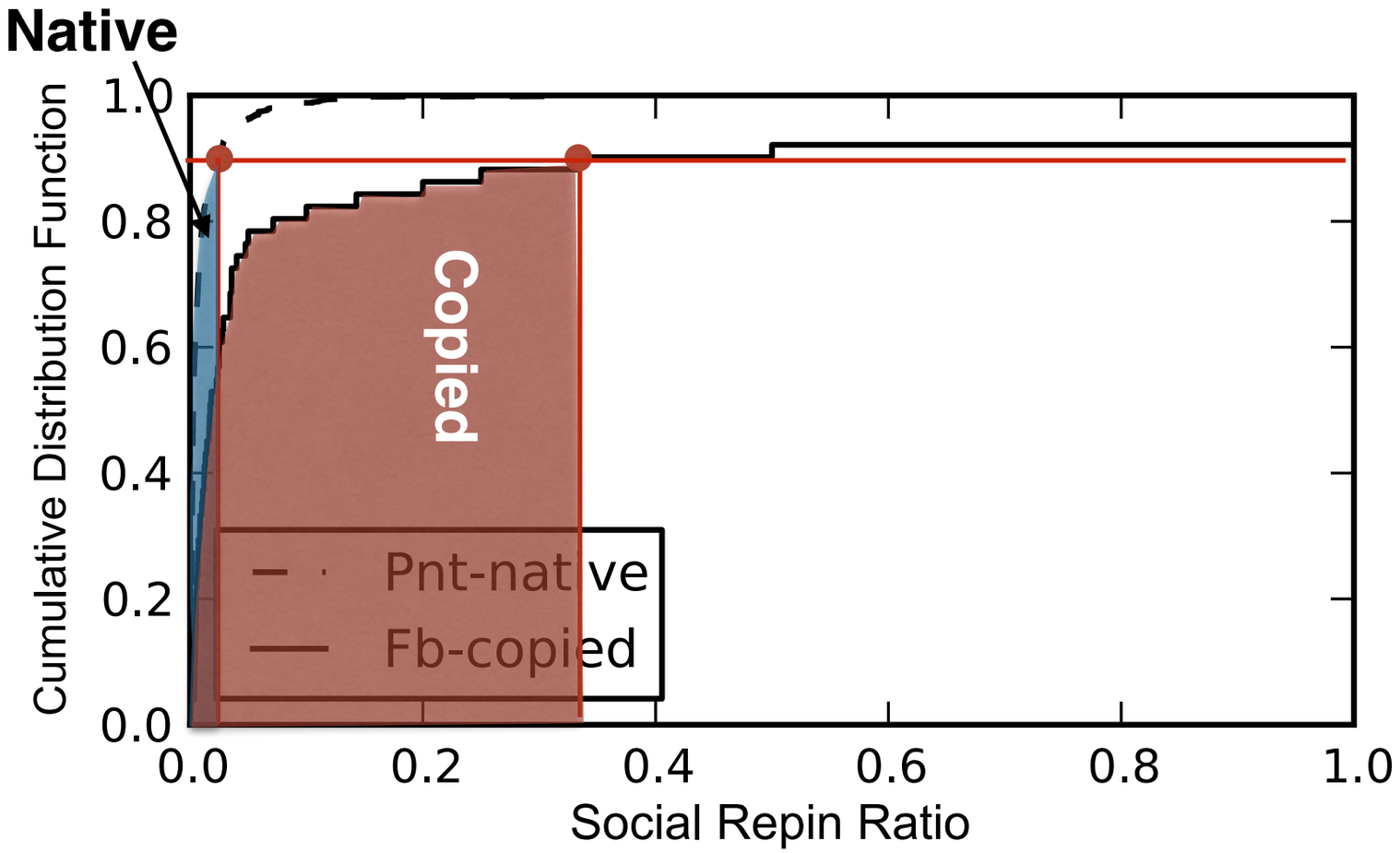
Structural benefits → Social interaction?

- **Repin** (the most popular activity on Pinterest):
  - Put images published by others into one's own collections.
- **Social Repins**: Repins made by friends.

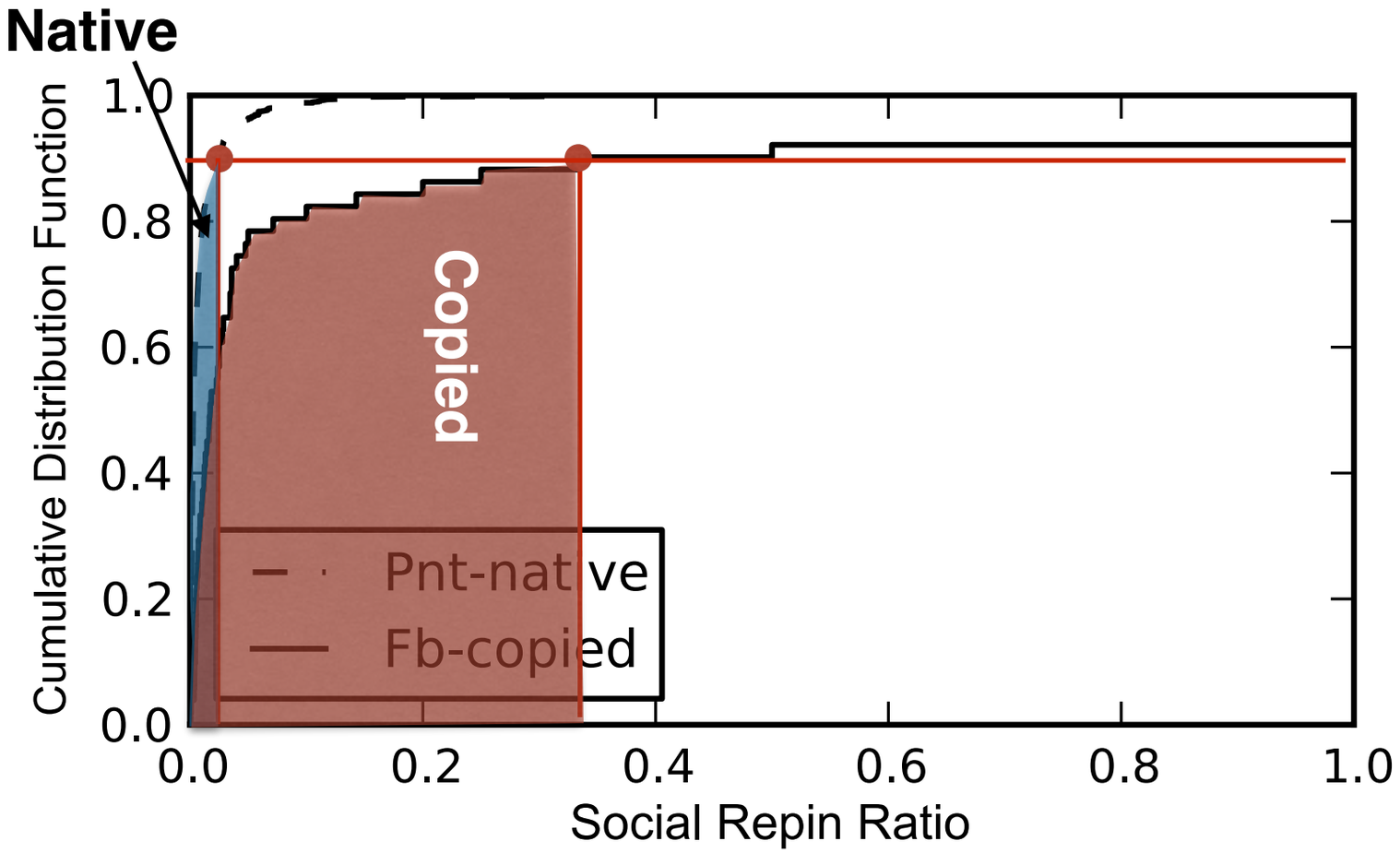
Copied links richer in social repins



Copied links richer in social repins



Copied links richer in social repins

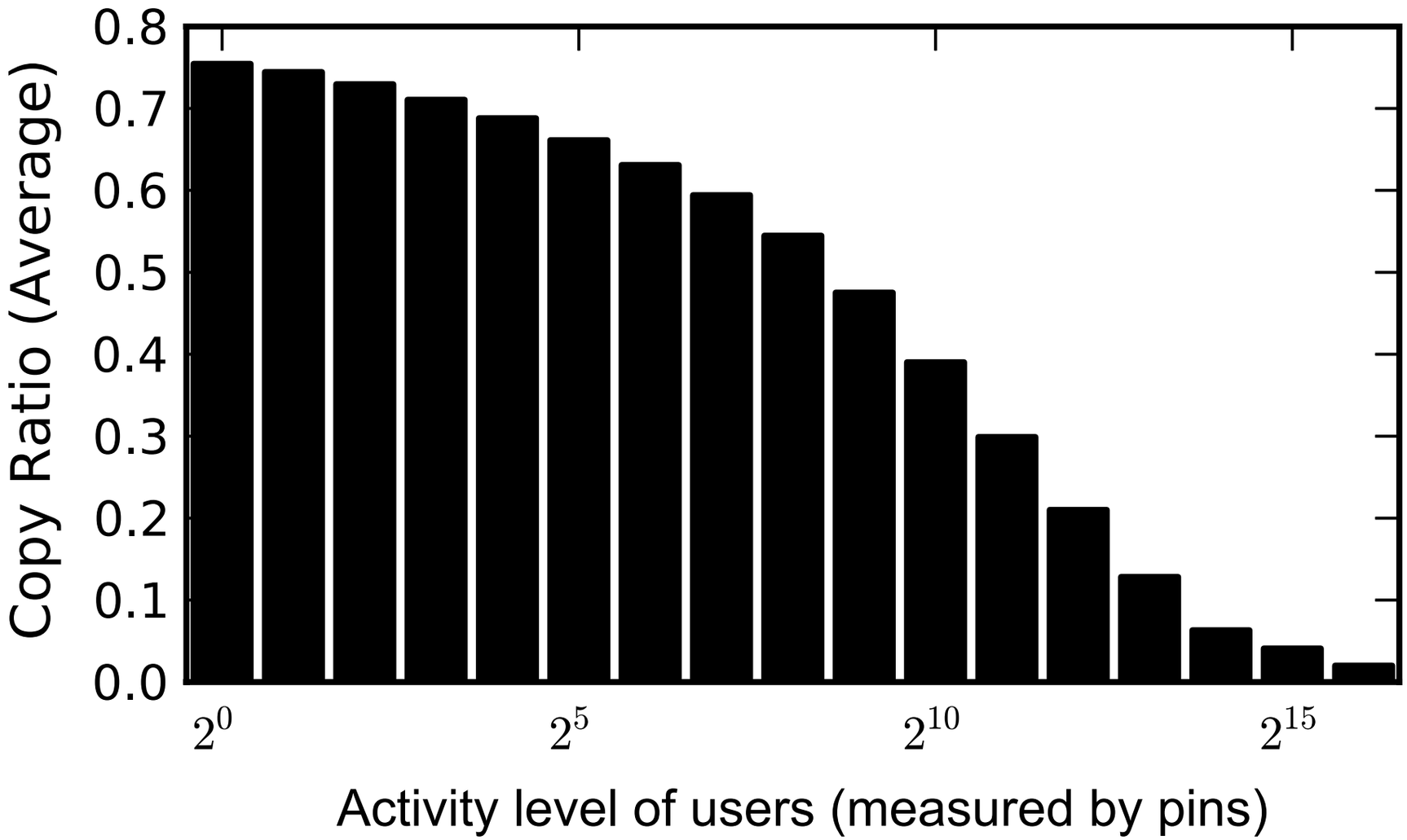


Copying creates networks which is **good** for social interaction

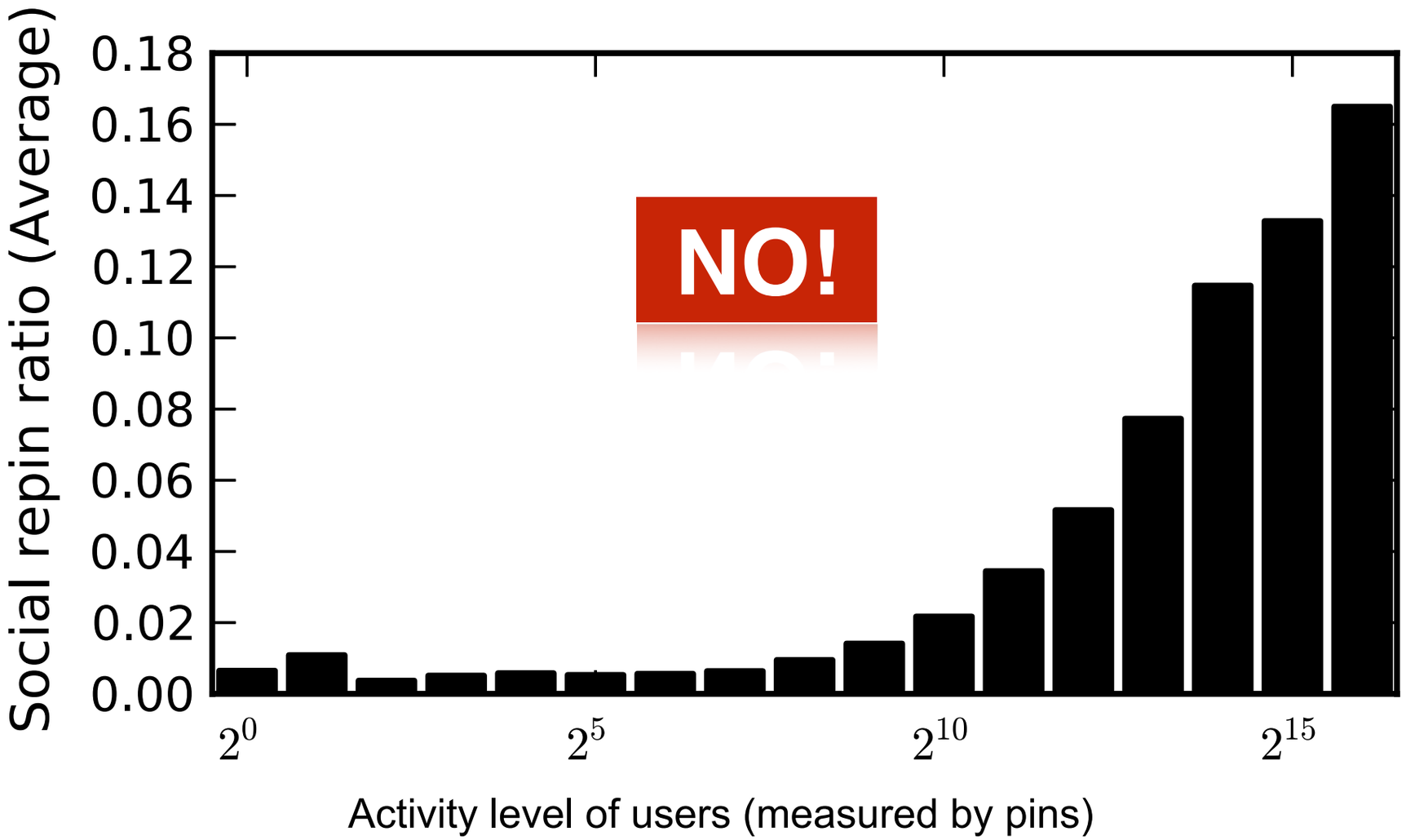
Outline

- *Q: Does copying create a good social community on the target website?*
  - **Structural Benefits:** Copying helps users get started with “good” structure with more social interactions
  - **“Weaning”**: Beyond Bootstrapping, active and influential users wean from Facebook to create new links natively

Active/influential users copy less

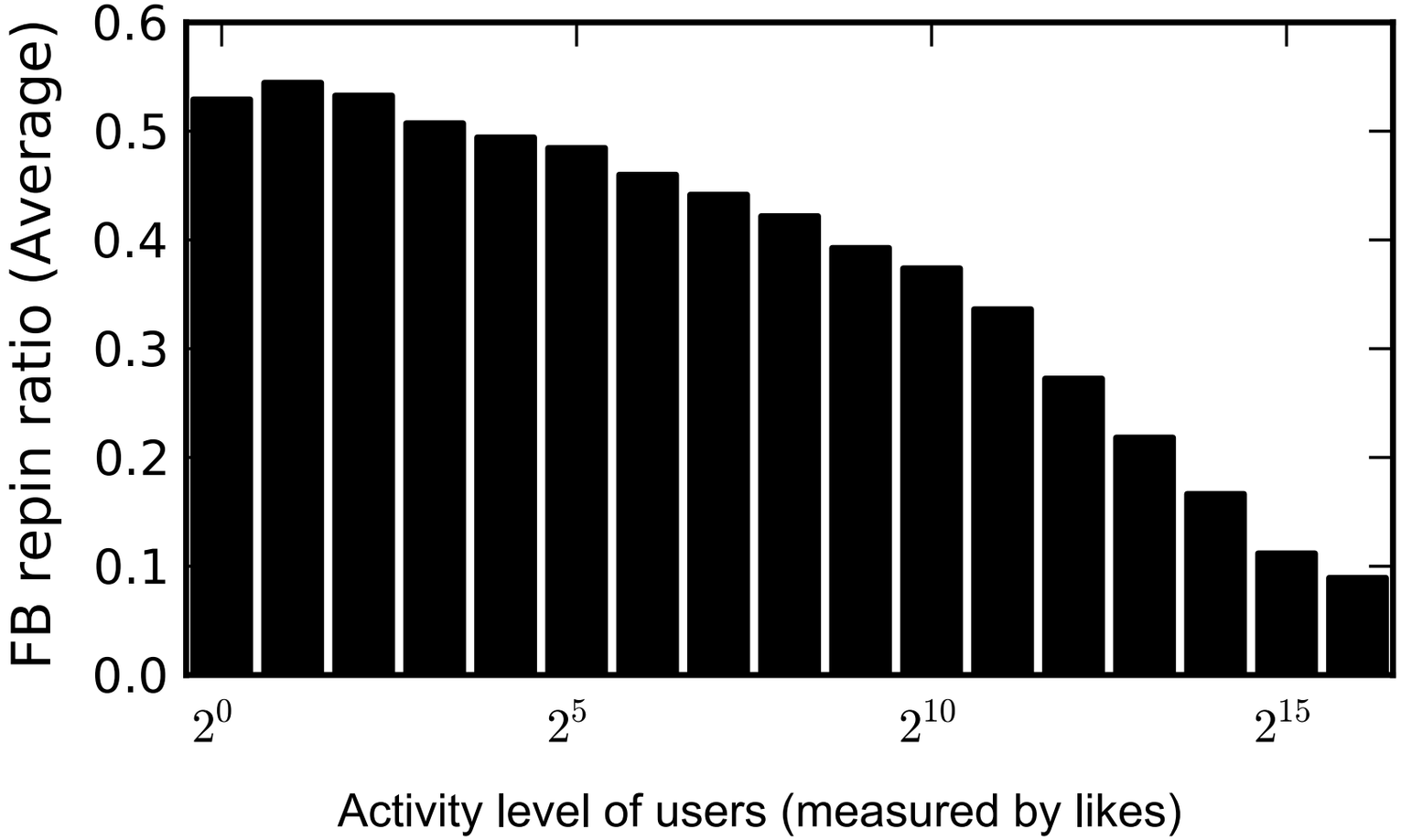


Are active users less social?



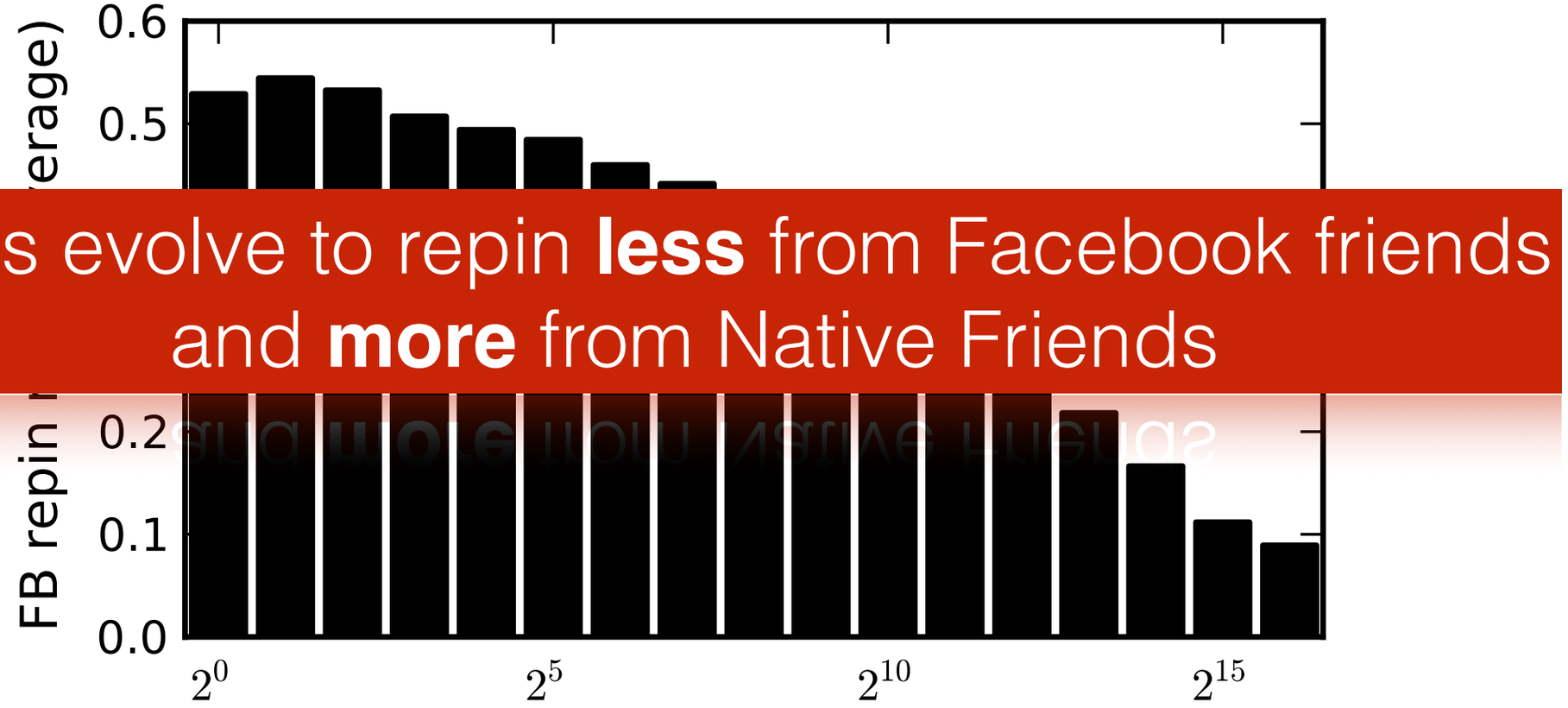
Weaning from Facebook

$$\textbf{FB Repin Ratio} = \frac{\text{Repins of Facebook friends}}{\text{Repins of Facebook } \underline{\textit{and}} \text{ Native friends}}$$



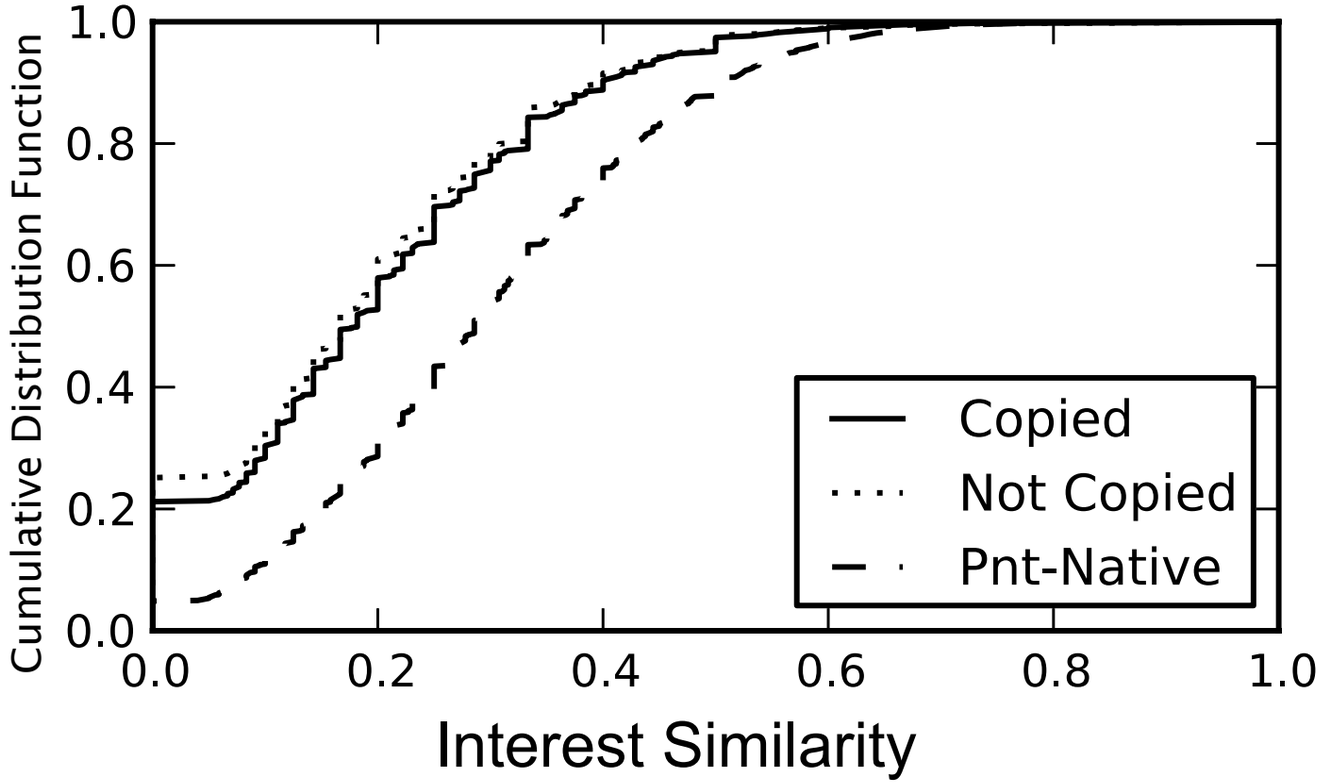
Weaning from Facebook

$$\textbf{FB Repin Ratio} = \frac{\text{Repins of Facebook friends}}{\text{Repins of Facebook } \underline{\textit{and}} \text{ Native friends}}$$

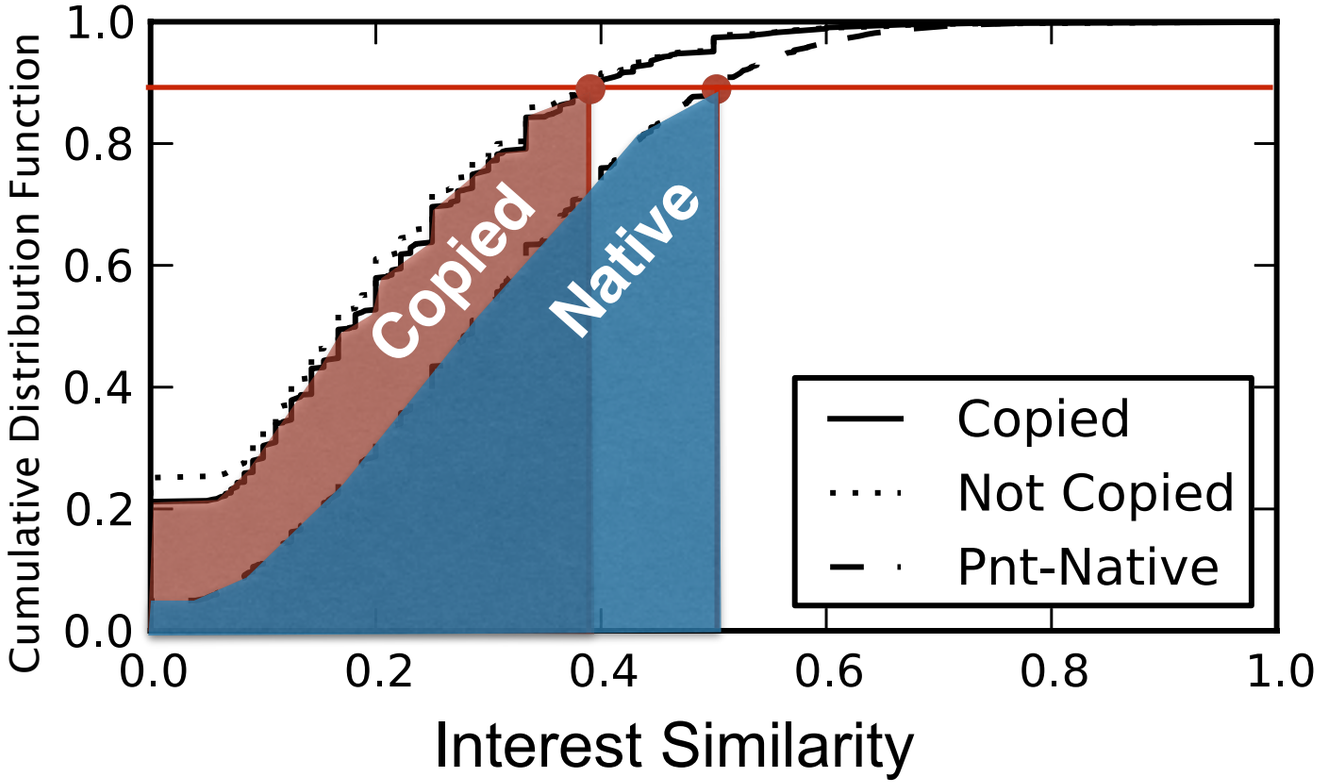


Users evolve to repin **less** from Facebook friends and **more** from Native Friends

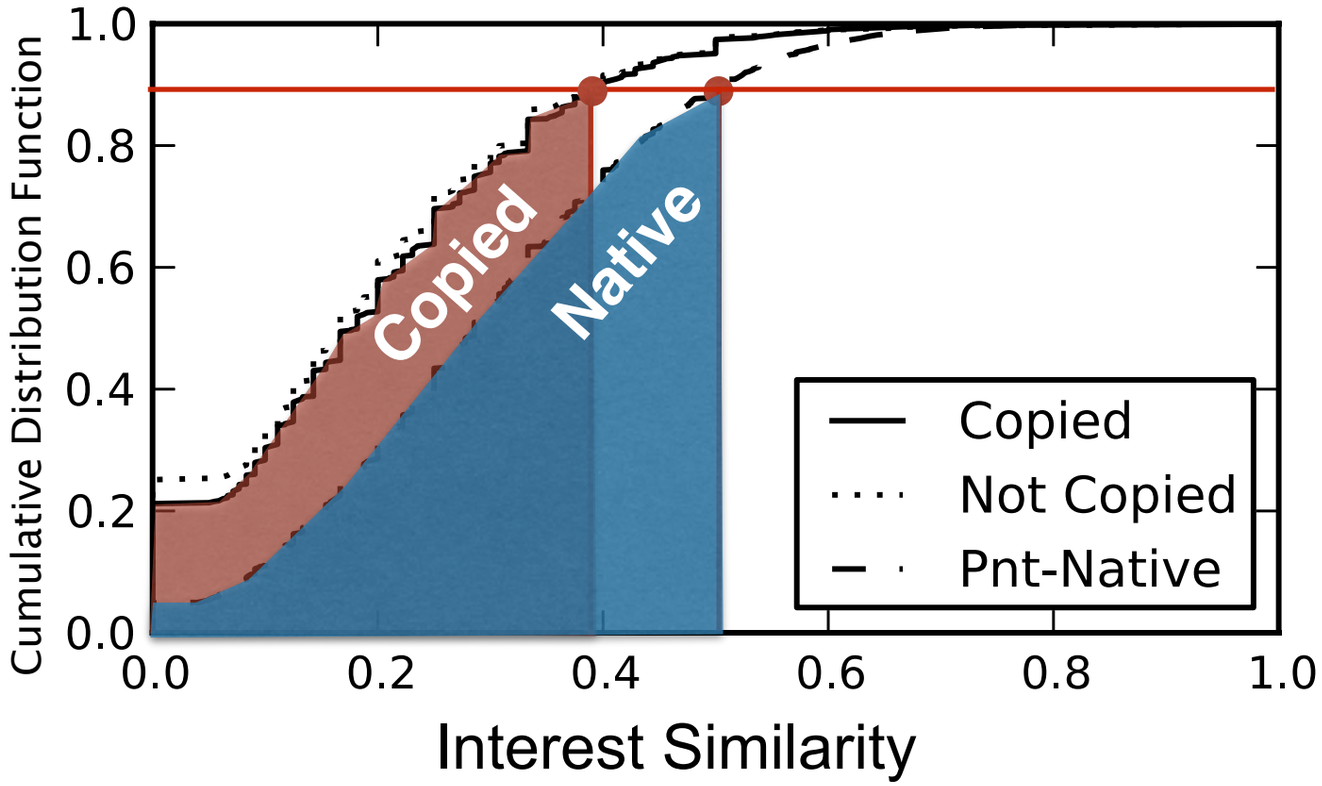
Why do active/influential users wean from copied to native friends?



Why do active/influential users wean from copied to native friends?



Why do active/influential users wean from copied to native friends?



Because **native friends match their interests more** than friends copied from Facebook

Answers

Q: Does social bootstrapping by copying links from Facebook create a good social community on the target website?

- Copying is **useful to initiate social interaction**
- Taking a long-term view, **active/influential users tend to move away** from copied social links and build social relationships natively.

Answers

Q: Does social bootstrapping by copying links from Facebook create a good social community on the target website?

- Copying is **useful to initiate social interaction**
- Taking a long-term view, **active/influential users tend to move away** from copied social links and build social relationships natively.

Copying and building links natively are both equally important to the success of target website.

Thank you!

Our dataset is available at

http://bit.ly/social-bootstrapping

Social Bootstrapping: How Pinterest and Last.fm Social Communities Benefit by Borrowing Links from Facebook

C. Zhong[1], M. Salehi[2], S. Shah[3], M. Cobzarenco[4], N. Sastry[1], M. Cha[5]

[1]**King's College London,** [2]University of Tehran, [3]UC Berkeley, [4]Last.fm, [5]KAIST