Exploiting Innocuous Activity for Correlating Users Across Sites

Oana Goga, Howard Lei, SHK. Parthasarathi, Gerald Friedland, Robin Sommer, Renata Teixeira

WWW 2013, May 15th
A lot of buzz about privacy

- Sharing settings Facebook offers
- Google’s terms of service
- Tracking cookies
What about linkage attacks?

What precisely enables such linkage attacks to succeed?
Our approach...

Exploit user **public online activity** to link accounts

- Attack model is a targeted individual
- Link on *where, when* and *what* a user posts
- Data sets: Yelp, Flickr and Twitter
Where a user posts

Location Profile

Twitter locations
Yelp locations
When a user posts

Timing Profile
What a user posts
Attack model

Targeted account

how similar?

Candidate list
Data sets

• Three social networks: Yelp, Twitter, Flickr

• Two types of data sets
  • Ground truth data set
    • Yelp-Twitter: 1,889
    • Flickr-Twitter: 13,629
  • Candidate Twitter list data set: 232,924

San Francisco, San Diego, New York, Chicago, Los Angeles
Linkage on location profile

- Represent a user’s location profile
  - Histogram of location units
- Measure similarity between location profiles
  - Cosine distance
Location profile with grids
Accuracy of linking with grids
Location profile with zip codes
Accuracy of linking with zip codes

Grid
Zip Code

False Positive Rate
True Positive Rate

0% 10% 100%
0 0.2 0.4 0.6 0.8 1

Grid
Zip Code
Location profile with clusters
Accuracy of linking with clusters
Number of linked accounts in practice

What does a 60% true positive rate at a 1% false positive rate mean?

1. 10 Flickr in 100 Twitter accounts
   - 16 linked accounts (6 true and 10 false positives)

2. 1 Flickr in 75,000 Twitter accounts (Bay Area)
   - 750 linked accounts (60% chances to find it)
What makes two accounts easy to link?

- **vulnerable**
  - Users with more than 3 common locations (80% of users)

- **medium vulnerable**
  - Add 2 unrelated locations

- **protected**
  - 95% of account pairs have no common locations
  - Add 7 unrelated locations

---

![Diagram showing False Positive Rate vs. True Positive Rate with different categories: vulnerable, medium vulnerable, and protected.](image_url)
### Accuracy when combining all features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Flickr to Twitter</th>
<th>Yelp to Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>60%</td>
<td>44%</td>
</tr>
<tr>
<td>Location + timing + language</td>
<td>63%</td>
<td>44%</td>
</tr>
<tr>
<td>Username</td>
<td>77%</td>
<td>7%</td>
</tr>
<tr>
<td>Location + timing + language + username</td>
<td>88%</td>
<td>44%</td>
</tr>
</tbody>
</table>
Conclusion

• Location can be exploited to link accounts
  • Location profiles should capture points of interest
  • Coarse grain location is enough
• Automatic reposting can be exploited to link accounts
• Language is not accurate enough to link accounts
Location profile with zip codes
Linkage on timing profile

• Method
  • The timing profile is a list of timestamps
  • Number of timestamp matches is the similarity metric
  • Need a threshold for a timestamp match
    – To big introduces noise/ to small might miss good matches

• Results
  • 6 common timestamps enough to narrow down to 50 accounts
Summary of findings

• Location and timing, when available, are powerful features to select candidate matches
  – Narrow down to 1,000 from 26,000 accounts for location
  – Narrow down to 10 from 26,000 accounts for timing

• Analysis of properties of user postings
  – Common zip codes make the matching easier
  – Adding locations to the targeted account is effective to make the matching harder
Performance metric

Rank ground truth user is 1

Targeted account
Performance metric

Targeted account

Rank ground truth user is 4
Performance of matching with location profile

- 35% of Flickr and Yelp accounts can be matched to a set of 250 Twitter accounts.
- 40% of Flickr and 60% of Yelp accounts can be matched to a set of 1000 Twitter accounts.
Properties that impact the performance of matching

- Do not post in the same zip codes on Yelp and Twitter
- Add more posts to Yelp to correct past mistakes

![Graph showing the relationship between the number of common zip codes and the probability to narrow to 500 accounts.](image)