Detection of Spam Tipping Behaviour on Foursquare

Anupama Aggarwal†, Prof P. Kumaraguru “PK”‡, Prof J. Almeida*

†Indraprastha Institute of Information Technology (IIIT-Delhi, India)
* Universidade Federal de Minas Gerais (UFMG, Brazil)
Who Am I?

• Assistant Professor, IIIT-Delhi
• Ph.D. from School of Computer Science, Carnegie Mellon University
• Research interests
  • Privacy, e-Crime, Online Social Media, Usable security
• Founded and manage PreCog, precog.iiitd.edu.in
• @ponguru, fb/ponnurangam.kumaraguru
Location Based Social Network

33 Million Users *

3.5 Billion checkins *

31% of mobile social media users use Foursquare *

* As of January 2013
Foursquare 101

Location Sharing OSN

Foursquare 101

Tip: Suggested Activity for a Venue

Tip can be Liked or Saved
Spam Tips

‣ Tips unrelated to Venue

Cisco left a tip at Baskin Robbins
Jan 3 - Pantai Medical Centre, Kuala Lumpur, Malaysia

“Buy the original XanGo mangosteen juice at best price
http://www.x1concept.com
Spam according to Foursquare ToS

- Tips with links to websites selling software, realtor contact info, a listing for your business, or other promotion

- Tips with inappropriate language or negativity directed at another person

- Unauthorized or unsolicited advertising, junk
Contributions

- Characterizing irregular user behaviour
  - We observed different categories of spam users
  - We characterize features distinguishing these spam users

- Automatic detection of spammers
  - Distinguish between spam and legitimate Foursquare users
  - Cluster spam users into different categories according to their behaviour
Data Crawling

2,400,594 tips
613,298 users
Observed Categories of Spam Users

- **Marketing**: These users post tips to promote and advertise a specific product/brand/venue/external URL
- **Malicious**: Such Foursquare users post external URLs in Tips which direct to spam/phishing/malware websites
- **Abusive / Derogatory**: These users try to deface or bad-mouth another person
- **Self Promotion**: These users try to draw attention to themselves
## Ground Truth Data

### Annotation Portal

<table>
<thead>
<tr>
<th>Business Name</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>The NoName Bar AKA Drink (Nightlife Spot)</td>
<td>Professional &amp; Other Places</td>
</tr>
<tr>
<td>Workahol (Professional &amp; Other Places)</td>
<td>Shop &amp; Service</td>
</tr>
<tr>
<td>Darcy do you work with us? Lol I'm confused.</td>
<td>Shop &amp; Service</td>
</tr>
<tr>
<td>Very nice stuff unlike most tanning salons.</td>
<td>Shop &amp; Service</td>
</tr>
<tr>
<td>Danque Tan WSP (Shop &amp; Service)</td>
<td>Shop &amp; Service</td>
</tr>
<tr>
<td>Honestly the slowest bk known to man... Nice staff though.</td>
<td>Shop &amp; Service</td>
</tr>
<tr>
<td>Burger King (Food)</td>
<td>Shop &amp; Service</td>
</tr>
<tr>
<td>Animal Party!</td>
<td>Residence</td>
</tr>
<tr>
<td>Dustins Crib! (Residence)</td>
<td></td>
</tr>
<tr>
<td>Hands down the best &quot;fast food&quot; Asian restaurant you will find. Everything is made fresh for you and tastes like heaven. They even have beer or wine if you're feeling frisky!</td>
<td>Shop &amp; Service</td>
</tr>
<tr>
<td>Pal Wei Asian Diner (Food)</td>
<td></td>
</tr>
<tr>
<td>Frozen Margarita Packets!</td>
<td>Shop &amp; Service</td>
</tr>
<tr>
<td>R/C Liquors (Shop &amp; Service)</td>
<td></td>
</tr>
<tr>
<td>Parmesan Garlic is what's up. Teehee!</td>
<td>Shop &amp; Service</td>
</tr>
<tr>
<td>Buffalo Wild Wings (Food)</td>
<td></td>
</tr>
<tr>
<td>So much fun! People.....is the point for them to be a &quot;dick&quot; obviously...</td>
<td>Shop &amp; Service</td>
</tr>
<tr>
<td>Dick's Last Resort (Food)</td>
<td></td>
</tr>
<tr>
<td>Staff is awesome. Ownership is friendly and doesn't treat you like a number. Prices are extremely reasonable plus free water and coffee!</td>
<td>Shop &amp; Service</td>
</tr>
<tr>
<td>Relax And Tan (Shop &amp; Service)</td>
<td></td>
</tr>
</tbody>
</table>

Mark this Foursquare in either one or two of the following categories.
(Choose the categories which are most appropriate. Also, do not forget to tickmark the unrelated Tips above):
- Legitimate User
- Marketing / Advertising
- Malicious
- Self Promotion
- Abusive / Derogatory
- I can't decide

2,000 Legitimate users
1,900 Spammers
Features used to detect Spammers

- **User Attributes**
  - Properties of the Foursquare user profile and his checkins

- **Social Attributes**
  - Friends network of the Foursquare user under inspection

- **Content Attributes**
  - Details about Tips posted by the Foursquare user
## Features used

<table>
<thead>
<tr>
<th>Category</th>
<th>$\chi^2$ rank</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Attributes</td>
<td>1</td>
<td>Number of Tips</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Ratio of Check-ins and Tips</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Number of Check-ins</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Number of Badges</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>Number of Mayorships</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>Ratio of Check-ins and Badges</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>Number of Photos posted</td>
</tr>
<tr>
<td>Social Attributes</td>
<td>6</td>
<td>Number of Friends</td>
</tr>
<tr>
<td>Content Attributes</td>
<td>2</td>
<td>Similarity score of Tips</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Number of URLs posted</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Average number of words in Tips</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>Average number of characters in Tips</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Ratio of number of likes and number of Tips</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Average number of spam words in Tips</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>Average number of phone-numbers posted in Tips</td>
</tr>
</tbody>
</table>
Few Observations

- Spammers post same/similar Tips on multiple venues
- A large fraction of spam Tips contain URLs
- Spam Tips may also have phone numbers
- Legitimate users have more Friends
- Spammers have very few Friends but large number of Tips
Relation b/w Tips and Checkins

Irregular User Behaviour
Tips Distribution

Legitimate users

Spammers
## Classification Results

<table>
<thead>
<tr>
<th>Classification Algorithm</th>
<th>Precision (Spam)</th>
<th>Precision (Safe)</th>
<th>Recall (Spam)</th>
<th>Recall (Safe)</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN</td>
<td>83.2%</td>
<td>86.6%</td>
<td>86.3%</td>
<td>83.5%</td>
<td>84.89%</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>88.1%</td>
<td>89.2%</td>
<td>88.3%</td>
<td>85.8%</td>
<td>89.53%</td>
</tr>
<tr>
<td>Random Forest</td>
<td><strong>89.3%</strong></td>
<td><strong>90.2%</strong></td>
<td><strong>88.3%</strong></td>
<td><strong>90.3%</strong></td>
<td><strong>89.76%</strong></td>
</tr>
</tbody>
</table>
Detection of Spam Classes

- Expectation-Maximization (EM) clustering
- Spammers Categories -
  - Advertising / Marketing
  - Self Promotion
  - Abusive
  - Malicious
Detection of Spam Classes

- Clustering Accuracy for spammer categories -

<table>
<thead>
<tr>
<th>Category</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advertising</td>
<td>88.23%</td>
</tr>
<tr>
<td>Self-Promotion</td>
<td>87.23%</td>
</tr>
<tr>
<td>Abusive</td>
<td>78.88%</td>
</tr>
<tr>
<td>Malicious</td>
<td>0%</td>
</tr>
</tbody>
</table>
Conclusion

- Analyzed spammers behaviour on Foursquare

- We obtained an accuracy of 89.76% with Random Forest classifier to distinguish spammers from legitimate users

- We classified the spammers into four broad categories

- We were able to detect users belonging to Advertising, Self-promotion and Abusive categories with an accuracy of 88.23%, 87.23% and 78.88%
Future Work

- Refine our methodology by use of other classification algorithms
- Use multiclass classification to detect users in any of the spam categories
- Correlation of content and the URLs posted by different users can help us in identifying several spam campaigns on Foursquare
Do you know these venues?

- DBSocial (SIGMOD)
- HotSocial (KDD)
- PSOSM (WWW)
- SNS (Eurosys)
- WOSN (SIGCOMM/Usenix)
- WOSS (VLDB)
October 7 – 8, 2013
Boston, USA
http://cosn.acm.org/
Paper submission due: June 28, 2013
Thank You!

Questions?

pk@iiitd.ac.in
Thank You!

Questions?

pk@iiitd.ac.in
Backup Slides
Data Crawling Detail

1. Twitter Stream

2. Tweets with 4sq.com

3. Venue ID

4. Tips posted at that Venue

5. Users who post Tips

6. Twitter users + Tweets

Foursquare users + Tips

Tweets:
- Joseph Handschiegel @shioxypes
  I'm at The Double: An Urban Tavern (Chicago, IL) 4sq.com/LPlyEl
  from Chicago, IL
- Josh Smith @joshsmithboston
  I'm at Ri Ra Irish Pub and Restaurant (Portsmouth, NH) w/ 3 others
  4sq.com/L_bswiH

Tips:
- Go to Broken Yolk Cafe (Downtown) and have a pomegranate mimosa and stare at their big ass fan.
- Try the Al Pastor taco. Sooooo good.
Gender Ratio

Gender Ratio in Foursquare Data

- Male: 63.4%
- Female: 36.6%
Linked Accounts

- Facebook: 70.29%
- Twitter: 55.24%
- Intersection: 17.97%

43.45%