A Weibo Bot-users Indentification Model Based on Random Forest

LIU Kan¹, †, YUAN Yunying¹, LIU Ping²

1. School of Information and Safety Engineering, Zhongnan University of Economics and Law;
2. School of Information Management, Wuhan University;
† Corresponding author, E-mail: liukan@znufe.edu.cn

NLPCC 2014
What's the Problem?

- Lower credibility on online social network

  traditional water army
  - manually manipulated
  - easy to be distinguished
  - weak targeted ability
  - low efficiency

  advanced bot-users
  - high-level automation
  - strong disguise power
  - targeted ability to release
  - high efficiency
Key idea

- Got after Experiencing & Analysis

Framework for bot-users identification model
Data

- get users from investigation & API
- manually classify the data based on predefined rules
- Tool: java, R

表 1 用户的基本信息表
Table 1 Typical attributions for users

<table>
<thead>
<tr>
<th>属性</th>
<th>Id</th>
<th>Followers</th>
<th>Friends</th>
<th>Mutual_F</th>
<th>Comment</th>
<th>Commented</th>
<th>SendM</th>
<th>ReceiveM</th>
</tr>
</thead>
<tbody>
<tr>
<td>说明</td>
<td>用户 UID</td>
<td>粉丝数</td>
<td>关注数</td>
<td>互粉数</td>
<td>评论数</td>
<td>被评论数</td>
<td>发私信数</td>
<td>收私信数</td>
</tr>
</tbody>
</table>

表 2 微博信息表
Table 2 Typical attributions for Weibo

<table>
<thead>
<tr>
<th>属性</th>
<th>MID</th>
<th>Time</th>
<th>Content</th>
<th>Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>说明</td>
<td>微博 MID</td>
<td>创建时间</td>
<td>微博内容</td>
<td>发布平台</td>
</tr>
</tbody>
</table>

NLPCC 2014
Behavior-based

- **CCE (Corrected Conditional Entropy)**
  - measuring the regularity of user's behavior
  - treat time intervals of tweet from every user as a sequence $X = \{X_i\}$

Entropy: $H(X_1, \ldots, X_m) = E[I(x)] = - \sum_{x_1, \ldots, x_m} P(x_1, \ldots, x_m) \log P(x_1, \ldots, x_m)$

CE: $CE(X_m / X_{m-1}) = H(X_m / X_1, \ldots, X_{m-1}) = H(X_1, \ldots, X_m) - H(X_1, \ldots, X_{m-1})$

CCE: $CCE_m = CCE(X_m / X_{m-1}) = CE(X_m / X_{m-1}) + perc(X_m) \times EN(X_1)$

Final CCE: $CCE_u = \text{MIN}\{CCE_2, CCE_3, \ldots, CCE_m\}$

m: the length of the sequence

Entropy when $m = 1$

Percentage of unique sequences

NLPCC 2014
Behavior-based

- CCE

Fig. 2 Cumulative Distribution Function of CCE
Content-based

- **Similarity**
  - measuring ratio of weibo with repeated content

\[
\text{Similarity}_i = \frac{\text{IdenticalWeibo}_i}{\text{TotalWeibo}_i}
\]

TotalWeibo: # of all compared weibos
IdenticalWeibo: # of weibos with exactly the same content

\[
\text{Similarity}_u = \frac{1}{n} \sum_{i=1}^{n} \text{Similarity}_i
\]

n: # of weibos the targeted users released
Content-based

- Content-based

Fig. 3 Cumulative Distribution Function of Similarity

NLPCC 2014
Relation-Based

- Follow: reputation & MutualRatio
  - measuring one-way following relationship
    \[ \text{Reputation}_u = \frac{\text{Followers}_u}{\text{Friends}_n + \text{Followers}_u} \]
  - measuring bidirectional following relationship
    \[ \text{MutualRatio}_u = \frac{\text{Mutual}_F_u}{\text{Friends}_n + \text{Followers}_u - \text{Mutual}_F_u} \]
Relation-Based

- Relation-based

Fig. 4 Cumulative Distribution Function of Reputation

Fig. 5 Cumulative Distribution Function of MutualRatio
Relation-Based

- MentionRatio
  - measuring ratio of mentions in all weibos

\[ \text{MentionRatio}_u = \frac{\text{Mention}_u}{\text{TotalWeibo}_u} \]

![Cumulative Distribution Function of MentionRatio](image)

Fig. 6  Cumulative Distribution Function of MentionRatio
Relation-Based

- **CommentRatio**
  - *measuring difference between # of comments made and comments received*

![Graph showing cumulative distribution function of CommentRatio]

Fig. 7  Cumulative Distribution Function of CommentRatio
Relation-Based

- **MessageRatio**
  - measuring ratio of sending messages

$$Message_u = \frac{SendM_u}{SendM_u + ReceiveM_u}$$

![Message cumulative distribution function graph]

**Fig. 8** Cumulative Distribution Function of Messages
Platform-based

- **NumOfPlatform**
  - measuring the diversity in platform

---

Fig. 9  Bar graph of platform number
Classifier

- Re-expound Problem

\[ u\{CCE_u, \text{Similarity}_u, \text{Reputation}_u, \text{Mutual}_u, \text{MentionRatio}_u, \text{CommentRatio}_u, \text{Message}_u, \text{NumOfPlatform}_u \} \xrightarrow{\text{classify}} \{\text{Bot, Ordinary}\} \]

- Reasons for choosing Random Forest
  - not sensitive to correlation
  - not sensitive to outlier
  - easily get the importance of every feature
Classifier

- Random Forest
  - random features
  - random samples

NLPCC 2014
Classification

- test the efficiency

<table>
<thead>
<tr>
<th>Table</th>
<th>The prediction results for test data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results</td>
<td>Ordinary-users(30)</td>
</tr>
<tr>
<td>Precision</td>
<td>0.967</td>
</tr>
<tr>
<td>Recall</td>
<td>0.967</td>
</tr>
</tbody>
</table>

NLPCC 2014
Classification

- importance

Fig. 9 The mean decrease accuracy of each feature
Summary

Contributions
- specifically analyse the features of bot-users in Weibo
- a novel method to distinguish bot-users from ordinary ones
- an empirical study of the method's effectiveness

Future work
- Considering the semantic features
- Adding Graphic techniques
- Extending in others areas, like forum, E-commence websites
- ...
Thank you!
Scenario Experience

offer registered accounts

ask to spread fake information

bot-users
想一想我居然没吃过炸酱面，不知道正宗的是什么样，不过我也不喜欢吃那种可吃可不吃的面食，今天吃了一碗东北人做的，味道不好极了，决定以后还是吃拌面的好，要不吃完还得吃个馊补充......人有悲欢离合，食有阴晴圆缺。
想一想我居然没吃过炸酱面，不知道正宗的是

bcgfxdvxc
想一想我居然没吃过炸酱面，不知道正宗的是什么样，不过我也不喜欢那种可吃可不吃的面食，今天吃了一碗东北人做的，味道不好极了，决定以后还是吃拌面的好，要不完了还得吃个馍补充......人有悲欢离合，月有阴晴圆缺。

11月22日22:39 来自 微博 weibo.com

收藏 转发 评论

娓娓江夏
想一想我居然没吃过炸酱面，不知道正宗的是什么样，不过我也不喜欢那种可吃可不吃的面食，今天吃了一碗东北人做的，味道不好极了，决定以后还是吃拌面的好，要不完了还得吃个馍补充......人有悲欢离合，月有阴晴圆缺。

11月22日12:17 来自 微博 weibo.com

收藏 转发 评论

wo野丫头
想一想我居然没吃过炸酱面，不知道正宗的是什么样，不过我也不喜欢那种可吃可不吃的面食，今天吃了一碗东北人做的，味道不好极了，决定以后还是吃拌面的好，要不完了还得吃个馍补充......人有悲欢离合，月有阴晴圆缺。

11月22日00:02 来自 微博 weibo.com

收藏 转发 评论
激光切割，水冷设计，配备单芯单PCB卡皇GeForce GTX TITAN SLI，顶级性能同时保证主机稳定性。真土豪，不解释！http://t.cn/zRp6X2D
Related Work

- Water army
  - Has been well studied
  - Detect water army usually through URL, name, etc.

- Bot-users
  - hasn't been widely studied
  - mainly focus on the tweet

- Difference between Weibo and Tweet
  - forms: more diversive, including pics, videos, etc.
  - topics: more entertainment, while mainly news in tweet
  - behavior: higher frequency

Still requires further improvements!
Scenario Experience

offer registered accounts

ask to spread fake information

bot-users
Scenario Experience

offer registered accounts

bot-users

ask to spread fake information
Classification

- test the efficiency

Table 2  The prediction results for test data1

<table>
<thead>
<tr>
<th>识别结果</th>
<th>机器用户（30名）</th>
<th>普通用户（30名）</th>
<th>普通用户（120名）</th>
<th>普通用户（240名）</th>
</tr>
</thead>
<tbody>
<tr>
<td>机器用户</td>
<td>29</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>普通用户</td>
<td>1</td>
<td>29</td>
<td>118</td>
<td>238</td>
</tr>
</tbody>
</table>