Measuring User Influence in Twitter: The Million Follower Fallacy

Authors: M Cha, H Haddadiy, F Benevenuto, K Gummadi

Presented By: Saurabh Agrawal
Agenda

- Importance of Social Media Marketing
- Why Twitter?
- Viral Marketing and Influence Theories
- Measure and Analyze User Influence
- Conclusions
- References
Importance of Social Media Marketing

Source: Social Media Examiner 2012 Social Media Marketing Industry Report
Why Twitter?

- Twitter users are more likely to follow brands
- “Do you follow brands on social networks?”

Source: Social Media Examiner 2012 Social Media Marketing Industry Report
How Twitter Works?

Typical Actions of Twitter Users:

• **Read** messages of other users

• **Tweet** their own message

• **Follow** users who post interesting stuff

• **Pass** interesting information to their followers

• **Respond** to other user’s tweets
Viral Marketing

- Uses pre-existing social media for branding
- Self-replicating viral processes
- Analogous to computer virus
- Information spreads through word-of-mouth
Influence Theories

- **Traditional Theory**
  - Emphasizes on ‘*influentials*’ who excels in persuading others
  - Target *influentials* in the network

- **Modern Theory**
  - Emphasizes on interpersonal relationship between users
  - Readiness of society to adopt to innovations
Questions about Influence

• How influence diffuses through a network?

• Does influence vary across time and topics?

• Can influence in one area be transferred to others?

• How can an ordinary user gain influence in short time?

Real World Scenario:
TV advertisement features a famous actor to sell computers. Is it an effective marketing strategy?
Methodology to Study Influence

Steps involved in studying influence patterns:

1. Define Influence Metrics
   - Eg: Indegree, Mention, Retweet Influence
2. Collect Data
   - Eg: Twitter dataset
3. Measure Influence
   - Eg: User X has the retweet influence rank of 1
4. Analyze Results
   - Eg: Celebrities are the most mentioned users
Influence Metrics

- **Indegree Influence:**
  - Number of followers of a user
  - User’s *popularity*

- **Retweet Influence:**
  - Number of times a user’s message was forwarded
  - *Content value* of user’s tweet

- **Mention Influence:**
  - Number of times a user was mentioned in tweets
  - Ability to *engage* others in conversations
Twitter Dataset

• Gather data for 80M users

• Find the largest connected component

• Ignore users with private account settings, invalid screen names and inactive users

Final Dataset:
• ≈54M users with focus on ≈6M users
• 2 billion follow links, 1.7 billion tweets
• [http://twitter.mpi-sws.org/](http://twitter.mpi-sws.org/)
Measuring User Influence

• Sort users by each influence measure
  \[ I_1 \geq I_2 \geq I_3 \geq \ldots \geq I_n \] (number of followers)
  \[ R_1 \geq R_2 \geq R_3 \geq \ldots \geq R_n \] (number of retweets)
  \[ M_1 \geq M_2 \geq M_3 \geq \ldots \geq M_n \] (number of mentions)

Most Influential  Least Influential

• Users with same influence value get same rank

• Increasing rank indicates less influential user
Finding Top Influentials

What kind of users rank high for a given measure?

**Indegree Influence:**
- Public Figures and News Sources
  - Eg: Barack Obama, CNN

**Retweet Influence:**
- Content Aggregators, Experts and News Sources
  - Eg: Mashable, NY Times

**Mention Influence:**
- Celebrities
  - Eg: Roger Federer, Will Smith

![Venn diagram of the top-100 influentials across measures: The chart is normalized so that the total is 100%.]
Finding Top Influentials (contd.)

• How a user’s rank varies across different measures?

\[
\rho = 1 - \frac{6 \sum (x_i - y_i)^2}{N^3 - N}
\]

\( N = \) Size of dataset
\( X_i \) & \( Y_i \) are ranks of user i for influence measure \( X \) & \( Y \)

• Relative Influence Ranks

Why did the values decrease across columns?

Conclusion?

High Indegree User
\( \xrightarrow{\text{Initiates High Conversations}} \)
\( \xleftarrow{\text{Highly Spreads Messages}} \)
Does User Influence Vary across Topics?

Degree to which influence can vary?
- For each user and each topic, plot total retweets and mentions
- User X’s rank may vary across topics

Conclusion from the graph?
“Target top influencers than large number of ordinary users for effective marketing”
Does User Influence Vary across Topics? (contd.)

<table>
<thead>
<tr>
<th>Topics</th>
<th>Retweet Top 10%</th>
<th>Retweet 1%</th>
<th>Mentions Top 10%</th>
<th>Mentions 1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iran vs. Swine</td>
<td>0.54</td>
<td>0.62</td>
<td>0.59</td>
<td>0.68</td>
</tr>
<tr>
<td>Iran vs. Jackson</td>
<td>0.48</td>
<td>0.54</td>
<td>0.59</td>
<td>0.63</td>
</tr>
<tr>
<td>Swine vs. Jackson</td>
<td>0.55</td>
<td>0.50</td>
<td>0.80</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Xi = Influence measure M, user i, topic X
Yi = Influence measure M, user i, topic Y

Conclusions from the graph?
Rise and Fall of Influentials with Time

- Track rankings of top influencers over time
  - All-time Influentials

Choose Top Influentials → Track Influence over 8 months → Compute Probability

![Graphs showing retweet and mention probability over time for different user categories.](image)
Rise and Fall of Influentials with Time (contd.)

- Study users who increased influence in short time
  - Topical Influentials

Choose Top Influentials ➔ Track Influence over 8 months ➔ Compute Probability

![Graphs showing influence trends over time](image-url)
Who is the Best Choice?

- Your company wants to market a product on Twitter. Who is the best choice?

A user with –

a) High indegree, low mentions & retweets
b) Moderate indegree, mentions & retweets

c) Moderate indegree, high mentions & retweets
d) Low indegree, moderate mentions & retweets
Summary

• Justifying the title: “Measuring User Influence on Twitter: Million Follower Fallacy”

• User’s influence varies across topics and time

• Top influentials hold influence across diverse topics

• Ordinary users can gain influence by dedicated efforts on a single topic
References


• Summary of this paper: http://www.diplomacy.edu/blog/indegree-retweets-and-mentions-what-does-it-take-be-influential-twitter

• Goyal, A.; Bonchi, F.; and Lakshmanan, L. V. S. 2010. Learning Influence Probabilities In Social Networks. In ACM WSDM.
Thank You !!!