Brandtology in Social Media

Hadi Amiri
Department of Computer Science
University of Maryland

hadi@umd.edu
www.umiacs.umd.edu/~hadi
Your Interest Based on Your Profiles!
TWITTER The Fastest Growing Social Platform

Twitter is now the fastest growing social platform increasing 40% between Q2 and Q4 2012. This means there are now 485m account holders and 288m active users.

FIND OUT MORE AT: globalwebindex.net
Twitter Active Users

33% asked friend’s opinion about a product
26% bought a product or service
29% posted a complaint about a brand / product
27% posted a pos comment about a brand / product
33% followed a group created by a brand
25% used an app created by a brand
30% shared photos from a brand
30% shared video created by a brand
779 K

599 K
Verizon Mentions

- Verizon
- @Verizon

Time (days)

#mentions
AT&T Mentions

- AT&T
- @AT&T
T-Mobile Mentions

- T-Mobile
- @T-Mobile

Time (days)

#mentions
Brands in SM

Figure 25: Brand and Reputation Monitoring of SMNs

Overall 465 respondents, LOB=107, EMEA=168

- Yes (employing in-house IT systems) 26.67%
- Yes (employing third-party service) 11.4%
- No 61.94%
Brands in SM

Figure 27a: Organizational Plans to Leverage Social Media Metrics Into Business Processes

Overall 459 respondents, LOB=107, EMEA=167

Preliminary consideration being given to this 28.76%

Yes 25.27%

No 45.79%
### Brands in SM

#### Figure 33: Top Business Processes Leveraging Social Media Data

Overall 300 respondents, LOB=79, EMEA=105

<table>
<thead>
<tr>
<th>Process</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Understanding customers</td>
<td>50%</td>
</tr>
<tr>
<td>Enhance the customer experience</td>
<td>43%</td>
</tr>
<tr>
<td>Innovation of services</td>
<td>31.3%</td>
</tr>
<tr>
<td>Implementation of social CRM</td>
<td>22.3%</td>
</tr>
<tr>
<td>Innovation of product and service delivery</td>
<td>21.3%</td>
</tr>
</tbody>
</table>
Customers in SM

- **Keep up with activities**: Facebook 52%, Twitter 41%, Pinterest 47%, YouTube 56%, Instagram 57%
- **Learn about product/service**: Facebook 35%, Twitter 41%, Pinterest 47%, YouTube 56%, Instagram 56%
- **Sweepstakes/Promotions**: Facebook 20%, Twitter 28%, Pinterest 36%, YouTube 48%
- **Provide helpful feedback**: Facebook 22%, Twitter 27%, Pinterest 23%, YouTube 22%
- **Join community of brand fans**: Facebook 27%, Twitter 26%, Pinterest 25%
- **Make purchases**: Facebook 15%, Twitter 21%, Pinterest 21%, YouTube 25%, Instagram 27%
- **To complain about product/service**: Facebook 9%, Twitter 12%, Pinterest 18%, YouTube 19%, Instagram 15%
Data Characteristics

Tomorrow!

2m, 2ma, 2mar, 2mara, 2maro, 2marrow, 2mor, 2mora, 2moro, 2morow, 2morr, 2morro, 2morrow, 2moz, 2mr, 2mro, 2mrrw, 2mrw, 2mw, tmmrw, tmo, tmoro, tmorrow, tmoz, tmr, tmro, tmrow, tmrrow, tmrrw, tmrw, tmrww, tmw, tomaro, tomarow, tomarro, tomarrow, tomm, tommarow, tommarow, tommarow, tommorow, tommorow, tommorw, tomorro, tomo, tomolo, tomoro, tomorow, tomarro, tommorw, tomoz, tomrw, tomz!

Credit: Alan Ritter
This Talk

- Brandtology in Social Media

**Who**

**What**
ORGSENSE System
Tweets

Michele Lim @pairow
Biz canteen sells cheap ramen for $3! Not as bad as I thought..

Joon Xin Ting @jinting
@stupidsmartalec @junksing not sure if nus/ntu biz is better but I don't think I’ve a high chance for being accepted at nus!

Mizael Poh @mizaelpoh
I found out earlier today that UTown has plugs underneath the tables outside of Starbucks! Awesome.... #swakoo

Alderick Teo @Alderick
staff at UTown Hwang's incredibly polite.

Madeline Sally @sallymadeline
Woah. Utown has such awesome facilities for meetings. :D

weiren. @thatredmoon
I'm at NUS High School Theatrette 4sq.com/HICxdl

Recent Images

User Network Structure
Tweet Chart

Keyword Chart

Location Chart

Sentiment Chart
This Talk

Fixed Keyword Crawler

Known Account Crawler

Org Key-user Crawler

Current Keywords

Keyword Miner

Classifier

Relevant Tweet Repository

User Miner

Sentiment Miner

Topic Miner

User Graph

Evolving topics

Emerging topics

Churn predictor

User Ranked List

Friend List Crawler
This Talk

• Topic Detection & Tracking

• Churn Prediction
Topic Detection - Cnt.

- **Input**
  - A dataset
  - Parameter $k$

- **Output**
  - $k$ topics
  - Each topic is a distribution over words in the dataset.
Topic Detection- Cnt.

• Different learning techniques
• Matrix factorization methods
  ▫ LU decomposition
  ▫ Singular Value Decomposition (SVD)
  ▫ Probabilistic Matrix Factorization (PMF)
  ▫ (Online) Non-negative Matrix Factorization (NMF)
  ▫ Etc.
NMF Approach

- Dataset: $m \times n$
- Topics: $m \times k$
- Topic Assignment: $k \times n$

$m$: # terms in the dataset
$n$: # docs in the dataset
$k$: # topics in the dataset

A sample doc vector
A sample topic vector
The topic assignment to docs
NMF Approach - Cnt.

\[ (D, X) = \arg\min_{D,X} \| S - D X \|_F^2 + \lambda \| X \|_1 \]

s.t. \( X \geq 0, D \geq 0, \|d_i\| = 1 \) for \( i = \{1, \ldots, k\} \)

\[ m: \# \text{ terms} \text{ in the dataset} \]
\[ n: \# \text{ docs} \text{ in the dataset} \]
\[ k: \# \text{ topics} \text{ in the dataset} \]
Topic Tracking

• Input
  ▫ a dataset of streaming (textual) contents
  ▫ Optional: initial set of topics

• Output
  ▫ topics in dataset
    • Each topic is a vector or distribution over words in the dataset.

Figure 4: Top 50 threads in the news cycle with highest volume for the period Aug. 1 – Oct. 31, 2008. Each thread consists of all news articles and blog posts containing a textual variant of a particular quoted phrase. (Phrase variants for the two largest threads in each week are shown as labels pointing to the corresponding thread.) The data is drawn as a stacked plot in which the thickness of the band corresponding to each thread indicates its volume over time. Interactive visualization is available at http://memetracker.org.

• Incremental Clustering for Topic Discovery

- Compute similarity btw each incoming tweet and each cluster center.
- If the maximum similarity value is greater than \( \tau \), assign the tweet to the cluster and update cluster center.
- Otherwise, generate a new cluster and cluster center.

- A better approach is to use Minhash or LSH

```plaintext
1: Input: tweet sets \( D \), topic cluster set \( C \), cluster center set \( Center \), and threshold \( \tau \).
2: Output: update topic clusters \( C \), and update cluster centers \( Center \).
3: Process:
4: if \( C = \emptyset \) then
5:     random select \( N \) tweets from \( D \) and add into \( C \) and \( Center \).
6: end if
7: initialize \( max, tmpC, tmpCenter \).
8: for \( d_i \in D \) do
9:     for \( center_j \in Center \) do
10:         compute Cosine Similarity \( sim \) between \( center_j \) and \( d_i \).
11:         if \( sim > max \) then
12:             \( max = sim, tmpC = C_j, tmpCenter = center_j \).
13:         end if
14:     end for
15:     if \( max > \tau \) then
16:         distribute \( d_i \) to cluster \( tmpC \), and update \( tmpCenter \).
17:     else
18:         new cluster and centroid and add to \( C \) and \( Center \).
19:     end if
20: end for
21: return \( C \) and \( Center \).
```
Input data matched with current topics

Non-active topics

K^{t-1} topics learned till t-1

Incoming data at t

S^t, D^{t-1}

\( R^*(S^t, D^{t-1}) < \eta \)

Residual error

Yes

Purging

D^{t-1}_{m \times k^{t-1}}

TL

N_{0}

Clustering

S^{em}

Input data with new topics

NMF, X-mean (Pelleg and Moor, 2000), etc.

Merge evolving and emerging topics

Evolving Topics

Emerging Topics

TL

TL

\[ D^t_{m \times k^t} = [D^{ev}_{m \times k^{t-1}}, D^{em}_{m \times k^t}] \]

Merging

Time

Topic Learning

\[(D, X) = \arg \min_{D,X} S - DX \|_F^2 + \lambda \| D - D^{t-1} \|_F^2 + \lambda \| X \|_1\]
Topic Tracking- Cnt.

- **Temporal Coherence constraint for topic learning:**
  - $D_{ev}$ to be a smooth evolution of $D^{t-1}$
  
  $$(D, X) = \arg \min_{D,X} \| S - DX \|_F^2 + \lambda \| D - D^{t-1} \|_F^2 + \lambda \| X \|_1$$
  
  s.t. $X \geq 0, D \geq 0, \|d_i\| = 1$ for $i = \{1, \ldots, k\}$

- **Can be solved efficiently** (Mairal et al., 2010)
  - **Space:** $O(n*m)$, given that $m >> k$
  - **Running time:** $O(n)$

### Table 5.4: Topic detection F1 performance with known and overall input. The higher values show better performances

<table>
<thead>
<tr>
<th>Organization</th>
<th>Baseline (NMF)</th>
<th>Optimization Framework</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>fixed-known</td>
<td>overall</td>
</tr>
<tr>
<td>NUS</td>
<td>39.30</td>
<td>40.03</td>
</tr>
<tr>
<td>DBS</td>
<td>64.67</td>
<td>61.42</td>
</tr>
<tr>
<td>StarHub</td>
<td>42.88</td>
<td>46.84</td>
</tr>
<tr>
<td>Average</td>
<td>48.95</td>
<td>49.43</td>
</tr>
</tbody>
</table>
Temporal Coherence constraint positively affect the performance

- Utilizing the past information about topics to make a better judgment about their current state.

**Figure 5.4:** Effect of the temporal continuity constraint on the performance of topic modeling for three organizations. We perform the experiments with $\mu = 0, 0.3, 0.6, 0.9$. 
Emerging Topic Detection for Organizations from Social Media. Chen et al., SIGIR 2013.
Summary

• Huge amount of contents are generated about brands on twitter.
• Finding the what (topics) is a challenging task:
  ▫ Dynamic Vocabulary
• Coherency constraints are important for more accurate topic detection and tracking.
• Early prediction of topics can help brand to do corrective actions before things go viral!
This Talk

- Topic Detection & Tracking
- Churn Prediction
Churn Prediction

- Churn happens when a customer leaves a brand or stop using its service.
- Churn rate indicates*:
  - customer response
  - the average time an individual remains a customer.

* Number of customers leaving a brand during a given period, divided by the average customer base during that period.
Churn Prediction - Problem

• Given a tweet about a brand, determine if it's churny or non-churny with respect to the brand!

• Sample churny tweets about Verizon:
  ▫ Cant wait to leave verizon for t-mobile! One more bill!!
  ▫ I cant take it anymore, the unlimited data isn't even worth it. My days with verizon are numbered.
  ▫ Verizon: I will change carriers as soon as contract is up.
  ▫ Verizon your customer service is horrible. this loyal customer will be gone #awfulcustomerservice.
Churn Prediction - Problem

• Given a tweet about a brand, determine if it's churny or non-churny with respect to the brand!

• Sample churny tweets about Verizon:
  ▫ Cant wait to leave verizon for t-mobile! One more bill!!
  ▫ I cant take it anymore, the unlimited data isnt even worth it. My days with verizon are numbered.
  ▫ Verizon: I will change carriers as soon as contract is up.
  ▫ Verizon your customer service is horrible. this loyal customer will be gone #awfulcustomerservice.
Isn’t It Easy?

Isn’t it enough to produce a list of churny keywords and rely on them alone to classify tweets as churny or non-churny.

* {leave, leaving, switch, switching, numbered, cancel, canceling, discontinue, give up, call off, through with, get rid, end contract, change to, changing, ...}
Isn’t it Sentiment Classification?

- Treat negative tweets about a brand as churny and positive ones as non-churny!

<table>
<thead>
<tr>
<th></th>
<th>Verizon</th>
<th></th>
<th></th>
<th>AT&amp;T</th>
<th></th>
<th></th>
<th>T-Mobile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P⁺</td>
<td>R⁺</td>
<td>F1⁺</td>
<td>P⁺</td>
<td>R⁺</td>
<td>F1⁺</td>
<td>P⁺</td>
</tr>
<tr>
<td>MetaMind</td>
<td>30.52</td>
<td>42.28</td>
<td>35.45</td>
<td>15.47</td>
<td>46.59</td>
<td>23.23</td>
<td>25.46</td>
</tr>
<tr>
<td>NRC</td>
<td>29.25</td>
<td>54.45</td>
<td>38.06</td>
<td>12.06</td>
<td>69.32</td>
<td>20.54</td>
<td>28.51</td>
</tr>
<tr>
<td>SemEvalSnt</td>
<td>28.49</td>
<td>54.76</td>
<td>37.48</td>
<td>11.86</td>
<td>65.91</td>
<td>20.10</td>
<td>28.58</td>
</tr>
<tr>
<td>TargetSnt</td>
<td>29.86</td>
<td>60.84</td>
<td><strong>40.06</strong></td>
<td>14.16</td>
<td>60.12</td>
<td><strong>23.51</strong></td>
<td>30.37</td>
</tr>
</tbody>
</table>

Table 4: F1 Performance of baselines for churn classification (reported for the churn class).
So, What makes it Challenging?

1. Comparative tweets that compare two or several brands!

I am leaving BRAND-1 for BRAND-2
So, What makes it Challenging?

2. Simple language constituents

Compare phrases like:

“switch to ATT” vs. “switch from ATT”
So, What makes it Challenging?

3. Negation effect!

“ATT is awesome, I’ll never leave them”
So, What makes it Challenging?

4. Churny keywords may not convey churnyness!

“I need a little ATT’s #help b4 leaving the states”
So, What makes it Challenging?

5. Subtle ways in using language as in

“debating if I should stay with ATT”
How to Solve the Problem?

• Demographic

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity ratio:</td>
</tr>
<tr>
<td>average No. of posts about brand/competitors per day</td>
</tr>
<tr>
<td>ratio of active days about brand/competitor</td>
</tr>
<tr>
<td>average time gap between posts about brand/competitor</td>
</tr>
<tr>
<td>ratio of urls in posts about brand/competitor</td>
</tr>
<tr>
<td>average No. of words in post about brand/competitor</td>
</tr>
<tr>
<td>Average of friends activity ratios</td>
</tr>
<tr>
<td># followers and friends</td>
</tr>
<tr>
<td>If user has bio information</td>
</tr>
<tr>
<td>If bio contains URL</td>
</tr>
</tbody>
</table>

• Content

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unigrams / Bigrams</td>
</tr>
<tr>
<td>Neighboring words of brand/competitors names</td>
</tr>
<tr>
<td>Syntactic and Comparative marker features</td>
</tr>
<tr>
<td>Sentiment features</td>
</tr>
<tr>
<td>Tense of tweet</td>
</tr>
<tr>
<td>News indicator features</td>
</tr>
</tbody>
</table>

• Context

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content features of user/friends/brand/competitors posts in thread (as defined in Table 4)</td>
</tr>
<tr>
<td># posts from user/friends/brand/competitors in thread</td>
</tr>
<tr>
<td># posts in thread</td>
</tr>
<tr>
<td>Reciprocity between user and brand/competitors posts</td>
</tr>
</tbody>
</table>

Target-dependent Churn Classification in Microblogs. Amiri and Daume III. AAAI 2015.
Handling Negation

I will leave Brand
(a) syntactic features: \{doj-leave-Brand, nsubj-leave-i, aux-leave-will\}.

I will never leave Brand
(b) syntactic features with negation effect: \{Neg-dobj-leave-Brand, Neg-nsubj-leave-i, Neg-aux-leave-will\}.
I want to switch from crappy t-mobile to verizon or att
I want to switch from crappy t-mobile to verizon or att
Content Representation

• Neural Net-based Language model
  ▫ Predicting the next word in a text given the previous words.
    • Given “Mary likes her coffee with milk and”,
      • A good LM may predict “sugar”, while a bad LM might predict "socks."
@VerizonSupport your customer service is HORRIBLE. This loyal customer will be gone #awfulcustomerservice

11:38 AM - 3 May 2014

Verizon Support @VerizonSupport · May 3
@micgold70 Please let us know what issues you are having with your services. We would like to help you. ^BJJ

AT&T Deals @ATTDeals · May 4
@micgold70 Sounds like you're reaching the end of your rope. Tweet me back about it! *KaylaC

Michelle Cosatino @micgold70 · May 4
@ATTDeals I will be visiting an AT&T Store as soon as I can verify I have good service at home

Michelle Cosatino @micgold70 · May 4
@VerizonSupport On the phone with 2 reps yesterday, I have ZERO coverage in my house. How generous of Verizon to have me pay for an extender?

AT&T Deals @ATTDeals · May 4
@micgold70 Hey Michelle, you're in luck. You can also check our coverage.
Data and Setting

• Twitter data about:
  ▫ Verizon
  ▫ T-Mobile
  ▫ AT&T

• Data and Paper available at
  ▫ www.umiacs.umd.edu/~hadi/chData
Performance

- Macro-Average Performance over Verizon, T-Mobile, AT&T

<table>
<thead>
<tr>
<th></th>
<th>BOW Features</th>
<th>hinge</th>
<th>logistic</th>
<th>BOW Features</th>
<th>hinge</th>
<th>logistic</th>
<th>RNN Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>unigram</td>
<td>65.30</td>
<td>64.30</td>
<td></td>
<td>62.97</td>
<td>63.73</td>
<td>tRep</td>
</tr>
<tr>
<td>(2)</td>
<td>unigram+Nb</td>
<td>73.63*</td>
<td>72.17</td>
<td></td>
<td>71.07</td>
<td>73.90*</td>
<td>twRep+NbRep</td>
</tr>
<tr>
<td>(3)</td>
<td>unigram+Dep</td>
<td>72.40</td>
<td>71.80</td>
<td>75.66*</td>
<td>75.43*</td>
<td>75.03*</td>
<td>twRep+DepRep</td>
</tr>
<tr>
<td>(4)</td>
<td>unigram+Cntx</td>
<td>74.27</td>
<td>73.20</td>
<td>75.47*</td>
<td>75.03*</td>
<td>75.03*</td>
<td>twRep+CntxRep</td>
</tr>
<tr>
<td>(5)</td>
<td>Cntt+Cntx</td>
<td>77.03</td>
<td>75.60</td>
<td></td>
<td>76.77</td>
<td>77.56*</td>
<td>CnttRep+CntxRep</td>
</tr>
<tr>
<td>(6)</td>
<td></td>
<td></td>
<td></td>
<td>hinge: 78.30*</td>
<td></td>
<td>logistic: 78.15*</td>
<td></td>
</tr>
</tbody>
</table>
## Context Effect

<table>
<thead>
<tr>
<th>Features</th>
<th>Hinge</th>
<th>Logistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>base: unigram+twRep</td>
<td>68.9</td>
<td>67.3</td>
</tr>
<tr>
<td>base+Context</td>
<td>75.7</td>
<td>75.4</td>
</tr>
<tr>
<td>base+Context—user</td>
<td>65.9</td>
<td>68.1</td>
</tr>
<tr>
<td></td>
<td>(-9.8)</td>
<td>(-7.3)</td>
</tr>
<tr>
<td>base+Context—friend</td>
<td>76.1</td>
<td>75.9</td>
</tr>
<tr>
<td></td>
<td>(+0.4)</td>
<td>(+0.5)</td>
</tr>
<tr>
<td>base+Context—compet</td>
<td>75.5</td>
<td>74.9</td>
</tr>
<tr>
<td></td>
<td>(-0.2)</td>
<td>(-0.5)</td>
</tr>
<tr>
<td>base+Context—brand</td>
<td>75.5</td>
<td>75.4</td>
</tr>
<tr>
<td></td>
<td>(-0.2)</td>
<td>(-.03)</td>
</tr>
</tbody>
</table>

Table 6: Ablation analysis of context churn indicators. Macro-average F1 Performance for all three brands. The performance is reported for the churn class.
Summary

• Churn prediction is a challenging task!
  ▫ Different from sentiment analysis.

• Demographic, content, and context features are important.
  ▫ Dependency path significantly improves the performance.

• Brand and competitor activities in discussion threads are useful for churn detection.

• This model is general and could be extended to any brand.
Acknowledgement

Hal Daumé III    Tat-seng Chua

Anqi Cui, Zhenpeng Li, me!, Zhenhua Zheng, Chen Yan
Thanks -- Questions?
NMF Approach - Cnt.

\[(D, X) = \arg \min_{D,X} \left\| S - DX \right\|_F^2 + \lambda \left\| X \right\|_1 \]

s.t. \(X \geq 0, D \geq 0, \|d_i\| = 1 \text{ for } i = \{1, \ldots, k\}\)

- Non-convex optimization problem.
  - many local optimum.
- But, if one of the variables, either \(D\) or \(X\), is known, optimization wrt the other will be convex.
  - Solution:
    - Iteratively optimize the objective function
    - Alternatively optimize wrt \(D\) and \(X\) while holding the other fixed!
Early Detection of Emerging Topics

- **User effect wrt the topics**
- **Tweet effect wrt topics**

\[ f_1 = \frac{|U^t|}{\sum_{x=0}^{t-1} \frac{1}{t-x+1} |U^x|}. \quad (6) \]

- **Tweet effect**

\[ f_2 = \frac{|T_{x}^{w^t}|}{\sum_{x=0}^{t-1} \frac{1}{t-x+1} |T_{x}^{w^x}|}. \quad (7) \]

- **Retweet rate**

\[ f_3 = \frac{|R_{w}^{t}|}{\sum_{x=0}^{t-1} \frac{1}{t-x+1} |R_{w}^{x}|}. \quad (8) \]

- **Overlap between org key users and top N influential topic users**

\[ f_4 = \frac{\#(k_{u_{tp}} \cap k_{u})}{\#k_{u_{tp}}}. \quad (9) \]

- **Overlap between org keywords and top N influential topic keywords**

\[ f_5 = \frac{\#(k_{w_{tp}} \cap k_{w})}{\#k_{w_{tp}}}. \quad (10) \]

- **Rate of increase of influence of the accumulated weight of tweets**

\[ f_6 = \frac{|A^t|}{\sum_{x=0}^{t-1} \frac{1}{t-x+1} |A^x|}. \quad (11) \]

where \( A = \sum_{w \in T_{w_{tp}}} a_{w_{tp}}(t_w) \).