Aether

Real-time Recovery and Visualization of Deleted Tweets

Pehr Hovey
Media Arts & Technology
Spring 2010
“I don't think I've ever felt so awkward in my life. I WANT TO CRY”
Approach

Melding art & technology

Provoke thought and discussion

Make system to be re-used in future projects
Permanence

Once you publish something on the internet, you can’t truly *take it back*

You can *delete* your post, but you never know who already has *your* data
Observation

The act of deletion adds a new dimension to the data, inviting interpretation.

No one really knows who is watching.

Not a lot of research done on erasing online identity.
Aether

Client-server software for real-time recovery and visualization of deleted tweets
Twitter
Twitter

Microblogging since 2006

“What are you doing, *right now*?”

Send short status updates (<140 char)

Follow interesting people, hope they follow back
Twitter

75 million registered accounts
15 million “active” in a month
50 million tweets per day in 2010
Twitter API

REST API: retrieving specific information
Search API: find tweets by keyword and location
Streaming API: real-time continuous TCP stream

*Aether* uses Streaming API & REST API
Twitter API

Data in XML, JSON or RSS Atom formats
Tweets come with User data inside

REST API is rate-limited
Stream API is downsampled
The Tweet (JSON)

{  
  "id": 13094530000,  
  "text": "Ok back not getting too bad. Hopefully I'll get some sleep.",

  "created_at": "Thu Apr 29 22:53:35 +0000 2010",
  "source": "<a href="http://ubertwitter.com">UberTwitter</a>",

  "coordinates": {  
    "type": "Point",
    "coordinates": [  
      34.05,  
      -118.25  
    ]
  }
}
The User (JSON)

"user": {
    "screen_name": "fifinoir",
    "name": "fiona",
    "id": 17502304,

    "followers_count": 63,
    "friends_count": 127,
    "statuses_count": 7038,

    "lang": "en",
    "created_at": "Thu Nov 20 02:36:19 +0000 2008",
    "time_zone": "Edinburgh",
    "utc_offset": 0,

    "geo_enabled": true,
    "location": "Anywhere you are!!",
}

Using Twitter API

**Twambulance**: Real-world impacts of a Tweet

1) Continuously watch the SB County fire dispatch feed
   @sbcFireDispatch

2) Look for tweets about Isla Vista:
   971 Embarcadero Del Mar *** Isla Vista *** Medical Emergency

3) Turn on microphone at the window
   and record the ambulance sounds
Deleted Tweets
Recovering Deleted Tweets

Save all tweets from streaming API

Twitter sends deleted status ID

Check archive for each id

Saved 22,297 234 recovered
Recovering Deleted Tweets

Found! @M2pics out for 326s:

I'm so doing this tomorrow morning RT @mashable HOW TO: Spring Clean Your Twitter Account http://bit.ly/aWc7jz #social-media #twitter

This user has deleted 1/1 we have seen
TwiXer Sample Stream Grab:
- Consume Stream
- Parse JSON
- Skip or put in cache

Retrieve: (deletes)
- See if we have Tweet
- Retry if not
- Create Event record

React:
- Store in history buffer
- Maintain list of clients
- Send events over OSC

Figures are for 30 minute run At mid-day

57613: 31/sec
22297: 12/sec
2572: 1.3/sec
234: 0.13/sec
234: 1 every 7 sec

Visualization

OSC Visualization (tweets)
Visualization
Visualizing Deleted Tweets

Present metrics to help viewers ponder
“Why do people delete their tweets?”

Two approaches:
User-centric: Who deletes their tweets?
Content-centric: What was deleted?
Things to look at

Server data rates:
Dynamic graph of estimated server throughput

Compare tweets per second and deletes per second
Things to look at

User details:

Who is this person?
What do they say about themselves?

Danielle Russo from New York, NY, US pressed delete 48 minutes ago

"I’m just a young girl livin’ in this crazy mixed up world3"
Things to look at

Context:
Other tweets sent around same time
May provide insight into reason for deletion

54m: Interesting: According to @Time, Iceland and Norway have the wc
1h: Interesting: According to @Times Iceland and Norway have the w
2h: @WaikikiBeachWlk @thelotuspond BTW is that Ruth Chris 45 Anni

7m: Decided it was best not to say anything... Going to be a long day.
33m: Awake.
10h: Thanks again for tonight @keifel!
Things to look at

**Lifetime:**
How long was the tweet publicly available before being recalled?

Hints at how many people may have seen it
Things to look at

Trajectory:
The path in time taken by the tweet
Things to look at

Recent Tweets

Most recent several tweets for comparison
Stamped with profile picture

@DonnieWahlberg I agree
THIS SITE GET 100 FREE TWITTER FOLLOWERS! http://tinyurl.com/2fha2rf.
@AtheistRE ...seems to be his bailiwick. Of course the Remnant Faithful daily prays fo
I JUST HAVE FREE 100 TWITTER FOLLOWERS! http://tinyurl.com/29dnlrp
Sometimes people dont realize how we feel inside. they just assume & say things
Things to look at

Word frequencies:

Which words are most frequently deleted?

How does this compare to typical tweets?

i
to the a and is love me of it its on be my you for have lol that would all ap at bowersox but com dont follow much new now other te thing this up want whole apart audition back blog building call cant de destroying didnt feel final finale girl glee just like lovefest made need no off oh or photo que retweet see should show shows
Results
Results

(Using sample stream – 5% of all tweets)

Skip up to 50% of the data (on purpose)
Save 40,000 tweets per hour
Check 4,600 deletions per hour
Recover 5-7% of deletions
~16% were deleted due to typo
Future Work

Revisit Database:
Store tweets in persistent database to increase lookback time
Need to solve performance issues

HTTP REST or Streaming API:
Allow others to access deleted tweet data
Aether becomes a public filter of the Twitter Stream

@Tweet_Morgue:
Twitter account that collects deleted tweets.
Could also send them back to their owner
“@FlyJeffOnDeck lol did u delete ur tweet about ur hard earned money?”

“@topmodeljasmin Yea I deleted a few... cuz I have a few stalkers who will go on my profile later today!!”

This was later deleted too