

NetrootsNation09 was the 2009 annual meeting of Netroots Nation

I collected 1,500 tweets from 3:29 p.m. on August 15 through 3:28 p.m. August 16 This was the end of the meeting, and the point was to look at 'bot' retweeting that might show up there. Dave Karpf raised the question about bot retweeting.

There are only a couple of micromessages about this; too few to pay attention to.

Dave:

I did a bit of research. Here are three pieces -- the first a summary by mashable -- about users of twitter. The second, mostactiveusers, is the 'full scoop' they have about bots. The third is a very interesting summary about users.

<http://mashable.com/2009/08/06/twitter-bots/>

<http://sysomos.com/insidetwitter/mostactiveusers/>

<http://sysomos.com/insidetwitter/>

I also did some search on #nn09. I got 1,500 tweets beginning 3:29 p.m. on the 15th and ending 3:28 p.m. on the 16th -- 1500 tweets in 24 hours. I found just over 400 (out of the 1,500) that had RT in them. That is a very high ratio; you netrooters are two or three standard deviations above the mean on retweets. [[nn09.xml](#), [nn09.txt](#)]

[I was curious about retweeting. I know they say it is low. But the total population may not be a good comparison. So I looked at the tweets #welovethenhs. The British are making fun of American ignorance about the NHS. So this is pretty political. It is also a stream -- not just independent tweets. In a stream one would expect more retweets, I think.

I have 15,000 tweets, which is days 2, 3, and 4 of the 'movement.' When I sort looking for RT at the beginning of the message I found 2,274. That is 15%, and that is about one-half the retweeting at #nn09 for the 1,500 tweets I have.

It is nice to have 'data' to play with when you know almost nothing.]

I sorted by name looking for a name that was sending many of the same messages. There were, naturally, many people sending many messages; you appeared twice. But I found very, very few sending the same message more than once. Then I sorted by message for tweets beginning with RT. That is not the entire population of retweets but it was most of them. Five was the most RT's of the same message, but it was by five different names. Almost uniformly the same message RTxxx was sent multiple times but by different names, which is the point of the procedure I believe.

None of this accounts for your experience. But I think it says that was an exception rather than a general pattern -- at least for the time period for which I have tweets.

I also searched using the twitter search routine and found 5 tweets by/from MikeConnery that referred to you. There were lots about his winning the poker championship. None of them were from scantily clad young ladies. They were probably earlier than my search could go back. Twitter is a bit better with its own search routine than it is with software reading their API. But that leads me to ask how you were finding these tweets that I did not find.

I am sorry I did not think to capture the entire 'collection' of #nn09. It would have been interesting to look at

the conversation that happened. There was a 'lot' of communication about poker in my little collection.

One answer is: we probably can do the research without much difficulty if we collect the tweets. The only real problem with that is knowing when to start. I am about a day late on all the collections I have going at this time.

Another answer is that while bots are big tweeters the biggest ones are at work for some company or other -- though one could certainly imagine political candidates paying for some service to do that kind of tweeting. Even if that was the case it should be easy to find as long as you know how to search -- they would be originating from the same place and have identical messages.

I have the file if you would like to look at it. It is a txt file that easily reads into excel. But it is the last moments and probably would not be very interesting to you.

My collection does not go as far back as your panel, but here is the retweet for something you wrote that was retweeted "@davekarpf "we have been called to take our country back, and now it is time for us to take it forward,.-Darcy burner" by democrat2theend at 5:42 on the 15th.

G. R. Boynton, Professor New Media and Politics

Analyses of the 2008 YouTube Campaigns

On Sun, Aug 16, 2009 at 11:47 AM "[Dave Karpf](mailto:davekarpf@gmail.com)" <davekarpf@gmail.com> wrote:

Interesting followup for us to be thinking about:

My panel at #nn09 was from 1:30 to 3:00PM yesterday. It generated some twitter posts and retweets, including two messages in particular: " Kerbel: Right is better using internet to drive media (@ [davekarpf](#) agrees. #nn09 #iar #p2 " and " RT @ [MikeConnery](#) : @ [DaveKarpf](#) saying good stuff re effectiveness and structure of blogosphere at Academic panel. Look him up. #NN09 "

Since then, those messages have been retweeted about once an hour for the past 18 hours straight. The retweets are coming from several users whose profile photos of are the skantly-clad female variety. I see only two possible explanations: (1) that I misread the audience yesterday and they have turned into a legion of adoring fans or (2) spambots. I think we can go with Occam's Razor and settle on thesis two for the time being.

The two questions I have are (1) why are they doing this? and (2) what biases is it introducing into the dataset? I imagine what the spambots are doing is echoing any trending topic, hoping that it leads to a bunch of "adds" from people searching or following that trend. Whether people actually indiscriminately follow new posters in this manner is, I think, a worthwhile question to examine. I'd guess that it's a tremendously low percentage, following the same logic of email spam (send out gigantic volume, get .001% replies, profit due to the near-zero cost environment), but it's a question worth playing with. The second question is a much bigger deal. This suggests that we should see trending topics go "viral" and stay active for an artificially long time. With enough data, we might be able to isolate the "spam echo" effect... not sure exactly how yet, but it certainly seems doable.

Anyway, the spambot issue is something that needs to be incorporated into any data analysis *somehow*... the fact that I picked up these echoes after only a few retweets on a topic that was only briefly trending indicates that threshold level for picking up this sort of echo effect is far lower than I would have otherwise guessed. It's magnifying some types of data, and it'll be worthwhile to figure out which ones.

Let's make sure to chat at APSA in a few weeks. Hope you're enjoying the summer.
-D

On Sat, Aug 15, 2009 at 6:43 PM, Dave Karpf <davekarpf@gmail.com> wrote:

Yep, here now. It's the Netroots Nation conference, been a great time.
-D

Sent from my iPhone

On Aug 15, 2009, at 6:07 PM, Bob Boynton <bob-boynton@uiowa.edu> wrote:

Hi:

I was checking out trending tweets, and #nn09 showed up -- briefly I think. Maybe you are/were there.

GRB

G. R. Boynton, Professor New Media and Politics
[Analyses of the 2008 YouTube Campaigns](#)

--

Dave Karpf, PhD

Postdoctoral Research Associate
Taubman Center for Public Policy
Brown University

www.davidkarpf.com
davekarpf@gmail.com