Twitter Lab - Analysis Report

Our AnalyzeTweets method mostly focused on the time range and frequency that tweeters were posting, analyzed over the period of time that we were collecting tweets. There were two different tests. The first was if a tweeter regularly posted around the clock and didn't seem to sleep. This was checked by whether the tweeter posted tweets for more than 17 hours (the average time an average person goes without sleeping according to Wikipedia) in a day without sufficient time to sleep in between (e.g. 7 - 8 hours). In other words, if a tweeter posted once every 3 hours for 24 hours, then they were likely to be a bot. The second test was whether a tweeter posted a number of tweets in short succession, e.g. within a minute or two of each other. Our threshold was 6, if a tweeter posted more than 6 tweets in a row with less than a minute or two in between each tweet, then they might be a bot.

Finally, we also checked every username to see if it contained "bot" as a substring. This is a very obvious method, but the end result was that it turned up quite a few twitter bots.

The graph is the top ten tweeters from the list of 100 most-likely bots our program gave us. The percentages on the right are from the BotOrNot website, and show that our program is a success!

@CartoonPlaza	85%
@Movie_Msgs	71%
@CineSerieTVbot_	64%
@m1ttyc	60%
@MediaBot	59%
@mediabot	59%
@VineBotRT	52%
@thatsmye	50%
@sdwahine78	47%
@GrandLakeTheater	40%