From Obscurity to Prominence in Minutes: Political Speech and Real-Time Search

Panagiotis Takis Metaxas
Wellesley College
Wellesley, MA02481, USA
pmetaxas@wellesley.edu

Eni Mustafaraj
Wellesley College
Wellesley, MA02481, USA
emustafa@wellesley.edu

ABSTRACT

Recently, all major search engines introduced a new feature: real-time search results, embedded in the first page of organic search results. The content appearing in these results is pulled within minutes of its generation from the so-called "real-time Web" such as Twitter, blogs, and news websites. In this paper, we argue that in the context of political speech, this feature provides disproportionate exposure to personal opinions, fabricated content, unverified events, lies and misrepresentations that otherwise would not find their way in the first page, giving them the opportunity to spread virally. To support our argument we provide concrete evidence from the recent Massachusetts (MA) senate race between Martha Coakley and Scott Brown, analyzing political community behavior on Twitter. In the process, we analyze the Twitter activity of those involved in exchanging messages, and we find that it is possible to predict their political orientation and detect attacks launched on Twitter, based on behavioral patterns of activity.

Keywords

Social Web, Real-Time Web, US elections, Twitter, Twitter-bomb, Google

1. INTRODUCTION

The web has become a primary source of information for most of decision-making situations. In particular, 55% of all American adults went online in 2008 to get involved in the political process or to get news and information about the election, up from 37% in 2004 [8]. Though just 1% of Americans used Twitter to post their thoughts about the campaign, the vast majority used search engines to be informed on any issue. It is well established that people trust search engine results, usually consulting only the first page of ranked results by the search engine. The belief that they are receiving trustworthy results is expressed through their consistent use of their favorite search engine.

Search engines, throughout their evolution, have struggled with the burden of having to deliver results that are both relevant to the query and trustworthy. And they have to wage an everyday war against spammers who use all kind of tricks to bypass the barriers and land in the first page of the search results [6]. Being in the first page is widely viewed as a strong indicator of reputation and popularity. It takes time to reach ranking levels that will allow a web site's link to appear in the first page of the search results. But spammers, scammers, or defamatory trouble-makers are in the business of reaping the rewards of ephemeral success. When they can trick a search engine to appear, even for a short time, in their first page search results, they are succeeding in their goal.

By incorporating real-time search results about timely popular queries in their first page of results, search engines have introduced a new opportunity for success to tricksters of all trades. This is especially troublesome in the context of political speech, where defamation of a candidate, once it catches the attention of the public, might have far-reaching consequences (e.g., the "Swift Boat campaign" [10] against Senator John Kerry).

In this paper, we argue that in the context of political speech, this feature provides disproportionate exposure to personal opinions, fabricated content, unverified events, lies and misrepresentations that otherwise would not find their way in the first page, giving them the opportunity to spread virally. To support our argument we provide concrete evidence from the recent Massachusetts senate race between Martha Coakley and Scott Brown, analyzing the activity of users on Twitter. In the process, we find that it is possible to predict their political orientation and detect attacks launched on Twitter, based on graph-theoretic properties, statistical properties and behavioral patterns of activity.

The rest of the paper is organized as follows: in Section 2 we show how Twitter messages (known as tweets) are displayed in the Google results page. In Section 3 we describe the data collected during the MA senate race in January 2010 and provide a detailed analysis of deriving the political orientation of users, community behavior and patterns of interaction. In Section 4, we discuss in detail spamming attacks from within Twitter. Section 5 summarizes our findings and offers some proposals to alleviate some of the concerns raised in the paper.

2. REAL-TIME SEARCH

Google announced on Dec 7, 2009 the introduction of real-time search [1], which provides fresh results relevant to a timely query. As people start generating new content, for example related to a sudden earthquake or a live event on TV, other people searching around the same time for such events on the Google search engine will get to see a box of latest results with the fresh content dynamically scrolling. The content appearing in these results is pulled from the so-called "real-time Web" such as Twitter, blogs, and news

Web Science Conf. 2010, April 26-27, 2010, Raleigh, NC, USA.

Table 1: Number and percent of tweets by message type. Repetitions is a separate category.

Type of tweet	Number	Percent
replies (@)	13,866	7.47
retweets (RT @)	75,407	40.63
other	96,311	51.91
TOTAL	185,584	100
repetitions	59,412	32.01

websites, within minutes of its generation.

Elections have always increased the public's interest in following the candidates. It is not surprising, then, that there was a similar public interest for the January 19, 2010, MA special election to replace the late Senator Ted Kennedy who died the year before. The election was contested mainly between two candidates, the Republican Scott Brown and the Democrat Martha Coakley. One of the ways that the public sought information about these two candidates was using search engines, as Figure 1 shows. Given the recent addition of the "real-time" content in Google's search results, the people who searched for "Martha Coakley" or "Scott Brown" saw the posts that were displayed around the same time on Twitter, like those shown in Figure 2 and 3. (Since the name "Coakley" is not common, searching for it brings in the same results as for searching for "Martha Coakley". This is not the case for "Brown" and so we included the whole string "Scott Brown" in our searches.)

3. TWEETS DURING THE 2010 MASS. SPECIAL ELECTION

During the period of January 13 to January 20, 2010, we monitored and collected the stream of more than 185,000 messages [1] containing the keywords "Coakley" and "Scott Brown" using the Streaming Twitter API [9]. About 41% of these messages (see Table 1 and Figure 5) were retweets, or messages that users had received and posted on their own account for their followers to see. A small percentage (7.47%) were replies, or messages directed towards another user. Interestingly, one out of three tweets was repetition of another identical message.

These messages were posted by almost 40,000 users in the period of 7 days, but not all users were equally active. The number of posts follows a power law-like distribution, as can be seen on Table 2. Based on their activity levels, we divided the users into three broad categories: Those who sent at least 100 messages, (there were 205 such users; we refer to them as *top200*); those who sent between 100 and 30 messages, (there were 765 such users; we refer to them as *topK*) and the remaining who sent less than 30 messages (we refer to them as the *low39K*).

3.1 Show me your friends, and I tell you who you are

For the *top200* users we also retrieved the friend and follower networks (Twitter API provides two social graph methods that allow to get the list of all followers or friends of a user. For privacy reasons, the list contains user IDs, instead of account names). Using graph-theoretic techniques, we drew their follower connections using a force-directed algorithm (see, e.g., [4]) and we found that the group clearly separated itself in two major components, as evident in Figure 4. The larger group is composed of 175 users leaning conservative, 29 users leaning liberal, one neutral, displayed as a light blue node at the top of the graph (who is on a mission to end the use of robo-calls by both candidates) and one spammer displayed as a light-colored node in the middle on the figure's right margin (who is likely trying to monitor the twitter trends over time). The figure reveals a number of other unconnected users, most of which did not have a Twitter account anymore at the time of this writing. We suspect that Tweeter deleted them as spammers, due to their unusual activity (high volume in a short period of time, no friends or followers). We report on the activity of most of them in Section 4.

The determination of the the political orientation of the *top200* group was done both manually (by reviewing the users' self-description or some of their messages) and automatically by searching for some obvious sentiment-revealing short phrases, such as "Go Scott Brown!"). Perhaps as a clear indication of the validity of the well-known proverb "show me your friends, and I tell you who you are" in social networks, the graph algorithm accurately guessed 98% of users' political orientation from the *top200* group. In a later paper, we report of the success of determining political orientation for all the users using a combination of graph theoretic and automatic mining methods.

Table 2: Number of messages posted by users follows a power law-like distribution.

Number of messages	Number of users
1	22482
2-3	9121
4-7	4090
8-15	2002
16-31	1093
32-63	524
64-127	227
128-255	88
256-511	36
512-1024	10
TOTAL	39,673

3.2 Whose message would you RT?

As we mentioned earlier, a large percentage of messages in our corpus were retweets of other messages. In Twitter vocabulary, a retweet is recognized by the initial phrase "RT @originalSender", where originalSender is the user name of the person who sent the original tweet. Interestingly, many of these messages were not simple RTs, but sequences of up to 7 RTs, as the table 3 shows. This fact made us wonder about the purpose of such activity. We formulated the following hypothesis:

[1]We have recently discovered about 50 thousand more tweets recorded during this period, but due to time constraints we have not included them in the analysis of this paper.

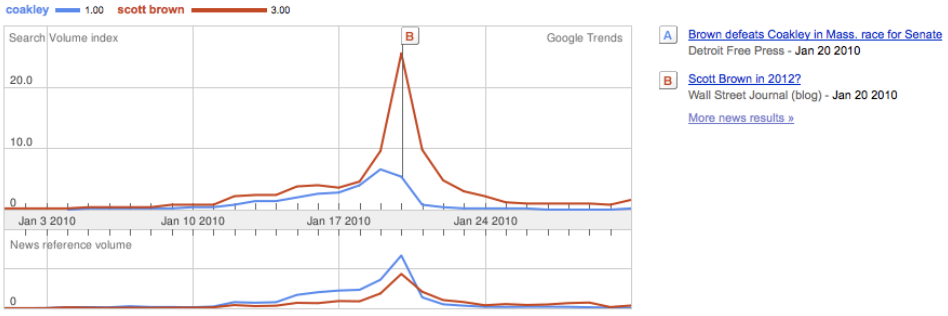

Figure 1: Google Trends for keywords "Coakley" and "Scott Brown" in January, 2010, shows a huge increase in searches during the week leading to the special election (January 19, 2010) and for a couple days after that.

Figure 2: Real-time results for the phrase "Scott Brown" displayed on the first page of Google's search results. Retrieved on January 15, 2010, four days before the election.

Figure 3: Real-time results for the phrase "Martha Coakley" displayed on the first page of Google's search results. Retrieved on January 15, 2010.

One is much more likely to retweet a message coming from an original sender with whom one agrees (shares political orientation).

One way to test this hypothesis was to test it on the fully-characterized *top200* group. Members of this group sent 10,008 RTs. The results, shown in Figure 6, shows this to be largely true: 96% of liberals and 99% of conservatives did so. The few messages that did not follow the overall trend are retweets with a negative commentary. By and large, users were very unlikely to retweet a message that they did not agree with. We should note, however, that the results are skewed by the fact that many users may not see the messages of users they do not follow, unless someone who does, retweets that message.

About 57% of the retweeted messages were between members of the *low39K* group which included those that sent a small number of messages overall. We are currently analyzing whether this hypothesis is able to distinguish as clearly the political orientation of members of the *low39K* group as well.

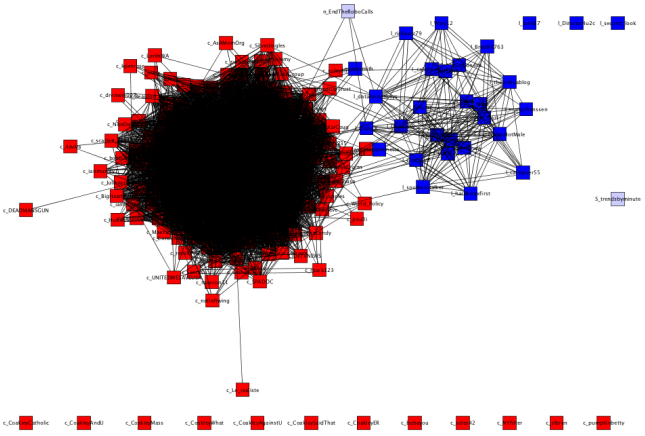

Figure 4: Two groups of users based on the followers graph. The graph is created using a force-directed drawing algorithm which draws nodes sharing many neighbors closer to those who do not.

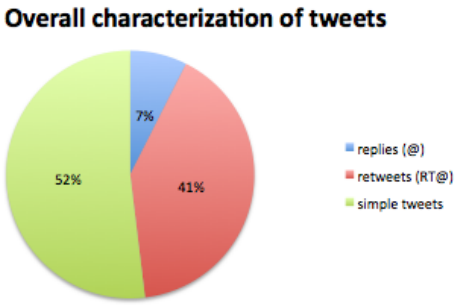

Figure 5: Overall characterization of corpus.

Table 3: Number of chain retweet messages.

RTs/msg	Number of messages
1	47730
2	21090
3	5349
4	939
5	149
6	47
7	18
TOTAL	75322

3.3 Repeating the same message

Since many of the users were aware of the fact that Google's search results were featuring Twitter trends, it made sense that they would repeat the same message in the hope that this message will show up in the first page of the search results. In fact, a surprisingly high number of tweets in our corpus, (one out of three tweets or 59,412 messages) are repetition of 16,453 different messages. Moreover, our data show that the *top200* group was far more likely to repeat messages (see Figure 7). We believe that this fact shows awareness of the new role that real-time web plays, since it

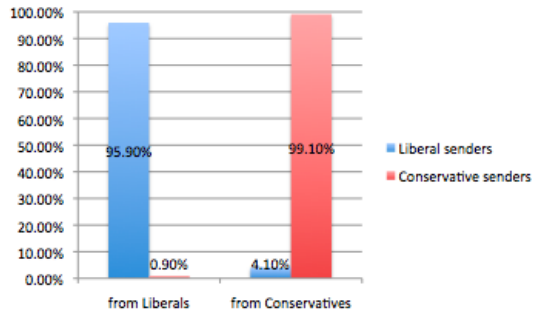

Figure 6: Both liberal-leaning and conservative-leaning users did not retweet messages they clearly did not agree with, though they retweeted 40% of all the messages.

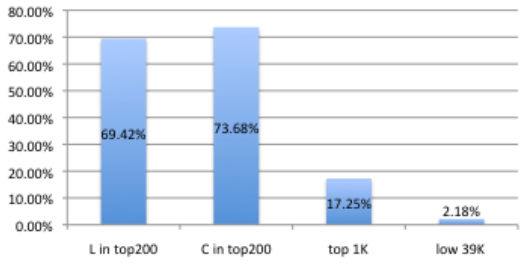

Figure 7: The members of the *top200* group, both liberals and conservatives, were far more likely to repeat a message (about 70 times) compared to the members of the other groups. This behavior reveals a highly motivated group who try to influence their followers and dominate search results on a topic.

does not make sense to bombard your followers, with whom you greatly agree, with the same message.

We discovered several threads of conversation that reveal the interest of the involved communities in following the real-time web, by discussing how certain phrases are trending in Google or Twitter, as well as encouraging others to google for a certain phrase they would like to see trending. Additionally, users are aware that by googling often for a person or topic, spikes in Google searches will attract media reports that attribute to such spikes a predictive power, noticed in previous political races [3]. Because metrics such as Google searches, number of views in YouTube, number of followers in Facebook or Twitter, or Twitter trending topics are being publicized as indicators that show advantage of one candidate over the other (because of greater public interest in them), we see a tendency from communities to skew these numbers toward their desired outcome.

3.4 Why would you reply?

If retweeting indicates agreement with the message, and repeating the same message multiple times indicates an effort to motivate the community and influence the Google search, what does it mean when you choose to reply to a message? We hypothesized that this direct engagement with the person who sent the message indicates that you are will-

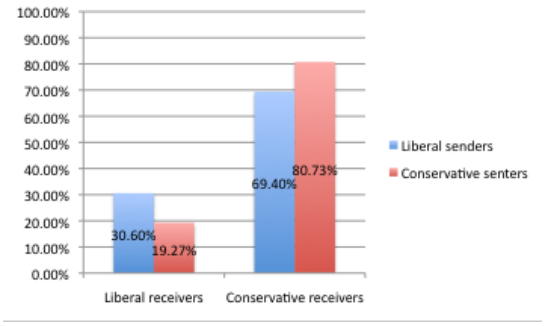

Figure 8: Despite their overall high activity, the *top200* users spent very little time replying to others. The majority of such messages were directed towards users of *topK* and *low39K* groups.

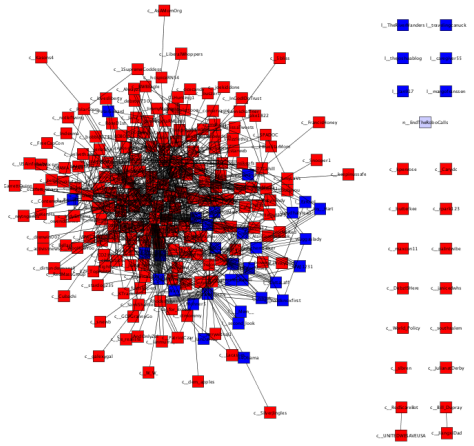

Figure 9: The reply activity of the *top200* users show a topology of closer engagement.

ing to be involved in an argument with the sender over some issue – in our case, the special elections.

While retweeting and repeating involves low levels of human activity (the press of a button or maybe the action of a computer program), truly replying requires time and energy. Not surprisingly, therefore, only 7.4% of all the messages were replies. Interestingly, the vast majority of the replies did not come from the *top200* users, despite their large message volume. Only 28.7% of replies were sent by the *top200* group, and a meager 7.4% of their replies were directed to members of the *top200* group. We present the following data (Figure 8) with the note that they are drawn from a very small part of our corpus (1016 messages).

Another way to visualize the reply-activity of the *top200* users is offered in Figure 9. This is also drawn with the force-directed algorithm. Note that the two groups are not separable based on their reply behavior. We observed, however, that a small number of *top200* accounts were responsible for many of the replies, in an attempt to flood the network with spam, as the next section 4 describes.

4. REPLYING AS A SPAM ATTACK

The common way in which spam works (independently of the distribution, by email, a web ad, or a tweet) is to provide a link to a website, that a user would likely not visit otherwise. Until recently, the best-known method of political spam on the Web involved the involuntary help of search engines. It has been widely reported in the news that, in 2006, political blogs had been actively trying to influence the US elections by pushing web pages carrying negative content to the top of the relevant search results of the major search engines. This practice of "gaming" the search engines was implemented with link bombing techniques (also known as Googlebombing), in which web site masters and bloggers use the anchor text to associate an obscure, negative term with a public entity [5]. In particular, during the 2006 US midterm congressional election, a concerted effort to manipulate ranking results in order to bring to public attention negative stories about Republican incumbents running for Congress took openly place under the solicitation of the liberal blog, MyDD.com (My Direct Democracy) [11]. Google took steps to curb such activity by promoting uncontroversial results in the first page, and it was found that political spammers were not very successful in the 2008 Congressional elections [7]

Thus, our search for spammers started with the analysis of tweets containing links. We extracted links and ranked them by their frequency in the corpus. Some of the links were expected, such as, `mybarackobama`, or the two campaign websites of the candidates, `brownforussenate` and `marthacoakley`. However, there were some unexpected links as well. One of them was `coakleysaidit`, which appeared 1088 times. Analyzing the content of the tweets containing this link, we discovered a concentrated spam attack. The tweets containing the links originated from 9 Twitter accounts, created within a 13 minutes interval, as shown in Table 4. The names of the accounts are related to the name of the website and are similar with each other. A domain lookup for `coakleysaidit` reveals that the website was also registered in the same day of their creation, January 15, 2010, using a service that hides the domain's owner identity.

It turns out that two months later, this web site was eventually signed. The group that signs the web site is a Republican group from Iowa that has been accused in the past of being behind several other attacks on Democratic candidates, including the "Swift Boat" attack [2].

An analysis of the spam attack shows that these 9 accounts sent 929 tweets addressed to 573 unique users in the course of 138 minutes. All tweets have the identical signature `@account Message URL`. Some examples of such tweets are shown in Table 5. We discovered that there are 10 unique text messages and 2 unique shortened URLs, both pointing to the same website. When treating all the volume of tweets as coming from one spammer, the median interval between two tweets is 1 second. Our assumption is that the attacker used an automatic script that randomly picked a user account, a text message, and a URL; packaged them in a tweet; and sent it by randomly choosing as sender one of the 9 spam accounts. While this seems as a good strategy to circumvent Twitter spamming detectors and may qualify as the first example of a *Twitter-bomb*, the attack was nevertheless discovered and all the spam accounts suspended.

The success of a Twitter-bomb relies on two factors: targeting users interested in the spam topic and relying on those users to spread the spam further. Especially the second factor is important, since spam accounts created only a few

Table 4: Accounts created for a spam attack

Account Name	Creation Time (EDT)	Nr. of tweets
CoakleySaidWhat	Jan 15 18:43:46 2010	28
CoakleyWhat	Jan 15 18:44:55 2010	127
CoakleySaidThat	Jan 15 18:46:12 2010	125
CoakleyAgainstU	Jan 15 18:48:21 2010	127
CoakleyCatholic	Jan 15 18:50:22 2010	127
CoakleyER	Jan 15 18:52:05 2010	127
CoakleyAG	Jan 15 18:53:17 2010	32
CoakleyMass	Jan 15 18:54:31 2010	109
CoakleyAndU	Jan 15 18:56:02 2010	127

hours before an attack have 0 followers, thus, no one would read their messages. The strategy used to find users interested in the topic, is a common spamming techqnique in Twitter: collect tweets that contain some desired keywords and find out the users who sent these tweets. Then, send a reply to these users and hope they will act upon it. There was a 4 hour interval between the creation of the accounts and the timestamp of sent messages and during that time, the attacker collected accounts that were tweeting about the senate race. In fact, 96% of the targeted accounts are also in our corpus posting in that time interval.

The attack was successful in terms of reaching the Twitter accounts of many users. We found 143 retweets in our corpus, the first after 5 minutes and the last after 24 hours of the attack. To estimate the audience of these messages, we calculated the set of all unique followers of the users that retweeted the original tweets. The audience size amounts to 61,732 Twitter users.

On the other hand, the effect of this attack could be seen as "preaching to the choir:" If the networks of friends and followers of the people following this campaign are as separate as the ones we observed in the *top200* group (Figure 4), far fewer undecided potential voters would have seen the message. But the attack would certainly have the effect of exciting the anti-Coakley conservatives.

While we cannot know how many of these users either read or acted upon these tweets (by clicking on the provided URL), the fact that a few minutes of work, using automated scripts and exploiting the open architecture of social networks such as Twitter, makes possible reaching a large audience for free (compared to TV and radio ads which cost several thousands of dollars), raises concerns about the deliberate exploitation of the medium.

Therefore, analyzing the signature of such spam attacks is important, because it helps in building mechanisms that will automatically detect such attacks in the future. An example is shown in Figure 10, which depicts the hourly rate of sent tweets during the 26 hours that include the attack timeline for the top 10 most active users. Accounts U5 to U10 belong to the spam attackers and it can be noticed that they have an identical signature (going from 0 to almost 60 tweets per hour). Thus, an averaged hourly sending rate would be a good distinguishing feature, though not sufficient. Currently, we are investigating a combination of features that take into account data on the source of the tweet (web, API, mobile web, etc.), the number of followers of the sender, the number of total tweets, the life of the account, etc.

Our experiments with Google real-time search has shown that, even though Google doesn't display tweets from users that have a spammer signature, it does display tweets from non-suspected users, even when these are retweets coming from spammers. Thus, simply suspending spamming attacks is not sufficient. There should be some mechanism that allows for retroactively deleting retweets of spam and some mechanism that labels some Twitter users as enablers of spam.

5. CONCLUSION

The introduction of real-time search results gives a search engine an aspect of social network communication, which recently has seen dramatic growth. But, by its current implementation by search engines, it also opens the door to exploitation and easy spamming. Currently, there is no way for the users to have any way of evaluating the trustworthiness of the real-time results, and the vast majority of the population that is not familiar with the way Twitter and blogs operate are likely to be fooled. In the political arena, it makes possible for a small fraction of the population to hijack the trustworthiness of a search engine and propagate their messages to a huge audience for free, with little effort, and without trace. We expect that, unless addressed by the search engines, this practice will intensify during the next Congressional elections in 2010.

6. ACKNOWLEDGEMENTS

Part of this research was funded by a Brachman-Hoffman grant.

7. REFERENCES

[1] A. Singhal. Relevance meets the real-time web. http://googleblog.blogspot.com/2009/12/relevance-meets-real-time-web.html, Dec., 7 2009.

[2] J. Hancock. Secrets of the American Future Fund. http://iowaindependent.com/4203/secrets-of-the-american-future-fund, 2008.

[3] R. Klein. Is Scott Brown closing the GOP technology gap? http://blogs.abcnews.com/thenote/2010/01/is-scott-brown-closing-the-gop-technology-gap.html, Jan., 18, 2010.

[4] S. G. Kobourov. *Force-Directed Drawing Algorithms.* CRC Press, R. Tamassia (ed.), Handbook of Graph Drawing and Visualization, 2010.

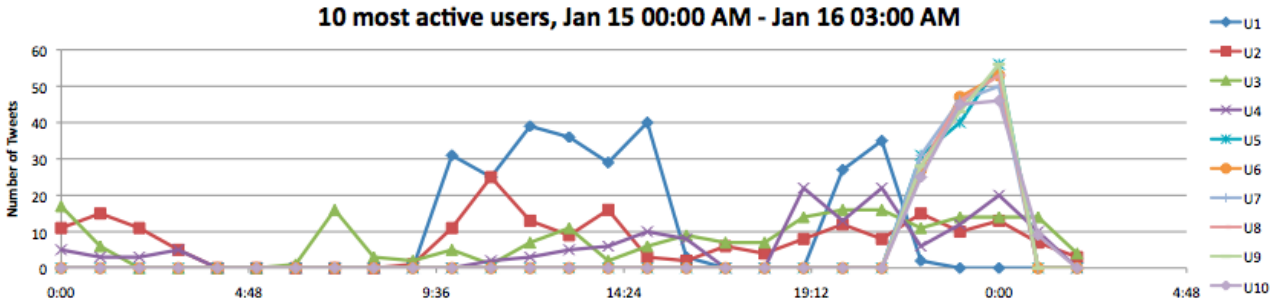

Figure 10: Top-10 users activity.

Table 5: Tweets from spamming accounts

@account	Message	URL
@theRQ	AG Coakley thinks Catholics shouldn't be in the ER, take action now!	http://bit.ly/8gDSp5
@Leann_az	Tell AG Coakley not to discriminate against Catholics in medicine!	http://bit.ly/8gDSp5
@mabvet	Catholics can practice medicine too! Tell AG Coakley today.	http://bit.ly/7yXbTd
@BrianD82	Sign the petition to AG Coakley today. We won't tolerate discrimination of any kind!	http://bit.ly/8gDSp5

[5] T. McNichol. Engineering google results to make a point. New York Times, January 22., 2004.

[6] P. T. Metaxas. On the evolution of search engine rankings. In *In the Proceedings of the 2009 WEBIST Conference*, March 2009.

[7] P. T. Metaxas and E. Mustafaraj. The battle for the 2008 us congressional elections on the web. In *In the Proceedings of the 2009 WebScience: Society On-Line Conference*, March 2009.

[8] The Pew Foundation. *The Internet's Role in Campaign 2008*. Published at http://www.pewinternet.org/Reports/2009/6–The-Internets-Role-in-Campaign-2008.aspx, New York, 2010.

[9] Twitter. Streaming API documentation. http://apiwiki.twitter.com/Streaming-API-Documentation, 2010.

[10] Wikipedia. Swift vets and pows for truth. http://en.wikipedia.org/wiki/Swift_Vets_and_POWs_for_Truth, Retrieved on March 25, 2010.

[11] T. Zeller Jr. Gaming the search engine, in a political season. New York Times, November 6., 2006.