



The Iranian Election on Twitter: The First Eighteen Days

About The Web Ecology Project

The Web Ecology Project is an interdisciplinary group of researchers based in Cambridge, Massachusetts. Its members are affiliated with the Berkman Center for Internet & Society, the Center for Future Civic Media, Harvard University, and the Massachusetts Institute of Technology. Our mission is to build tools to better understand the flows of culture and the formation of communities in the complex social ecology of the web.

For press inquiries, please visit our website (<http://webecologyproject.org>), or contact us via email at contact@webecologyproject.org or phone to Program Director Tim Hwang at (973) 960-4955.

This report is a publication of the Web Ecology Project. Contributing researchers to this report were: Jonathan Beilin, Matt Blake, Mac Cowell, David Fisher, Sam Gilbert, Russell Hanson, Tim Hwang, Alex Leavitt, Greg Marra, Rob Mason, Colin McSwiggen, Dharmishta Rood, Aaron Shaw, Frank Tobia, and Seth Woodworth. We also gratefully acknowledge Ethan Zuckerman (<http://www.ethanzuckerman.com/blog/2009/04/13/studying-twitter-and-the-moldovan-protests>), for his creation of some of the initial code on which this work builds.

This paper is dedicated to our mentor and hero, Charlie Nesson.

The Web Ecology Project releases this paper under a Creative Commons Attribution Share Alike 3.0 license (<http://creativecommons.org/licenses/by-sa/3.0/>).

Summary

The political events that transpired during and after the 2009 Iranian election garnered global attention, particularly due to the purported importance of Twitter. As a platform-independent service for communication, Twitter has become a preferred vehicle to broadcast unfolding events in Iran both within the country and to an international audience. The use of Twitter to communicate about these events is a striking instance of the memetic spread of behaviors and ideas online. Our report serves as an initial assessment of the Twitter web ecology -- that is, users interacting with their technological environment -- to sketch out the broad anatomy of discourse on Twitter by providing a novel perspective with quantitative data. Our servers continue to collect data, and we plan to publish a follow-up report that comprehensively explores the structure of the events from additional perspectives and depths.

Key Findings

- * From 7 June 2009 until the time of publication (26 June 2009), we have recorded 2,024,166 tweets about the election in Iran.
- * Approximately 480,000 users have contributed to this conversation alone.
- * 59.3% of users tweet just once, and these users contribute 14.1% of the total number.
- * The top 10% of users in our study account for 65.5% of total tweets.
- * 1 in 4 tweets about Iran is a retweet of another user's content.

Introduction and Existing Research

The conversation on Twitter about Iran offers a valuable opportunity to better understand the complex social ecology of the web, particularly in times of crisis. How has this conversation evolved? Who are the influential individuals and communities driving the conversation? How does the conversation interface with events and media flow outside of Twitter?

Since Iran's election (12 June 2009), countless press articles have been published that speculate on the role that Twitter has played in Iran. As an initial step toward contextualizing our data, we surveyed related published materials available online, reviewing articles from established media outlets, consultants, and researchers.

Recent analyses -- such as those published by Marc Ambinder (*The Atlantic*), Jack Shafer (*Slate*), Anne-Marie Corley (*Technology Review*), Joshua Kucera (*True Slant*), Evgeny Morozov (*Foreign Policy Magazine*), Daniel Drezner (*Foreign Policy Magazine*), Lev Grossman (*TIME Magazine*), and Clay Shirky (via the *TED* blog) -- have criticized the earlier hasty proclamation of a "Twitter Revolution," seeking instead to clarify the nature and scope of the phenomenon through more careful analysis. Most question the extent to which Twitter has been used as an organizing tool for Iranian citizens and emphasize instead the importance of Twitter as a system for publicizing events in Iran to the rest of the world. These analyses mostly provide qualitative evidence, and little quantitative data has been provided to support the claims.

Those quantitative analyses that do exist are typically based on very limited data sets.

For example, Maximillian Forte of *Open Anthropology* makes claims based on "a sample of 1,280 tweets" ranging from 13 June to 17 June, while Noam Cohen, from the *New York Times*, briefly investigates only one user, mousavi1388, from 16 June. Some of the quantitative analysis published to blogs seems to reflect the better data released from sources such as *Sysomos*, which has released a robust report that profiles users in terms of account creation and message location (both in Iran and abroad). Still, the team at *Sysomos* and even Ben Parr of *Mashable* provides results for only one hashtag, #IranElection, and one term, "iran."

As Andrew Sullivan of *The Atlantic* comments, after reposting two messages from Twitter, "Those are recent tweets which probably tells you more about the mood than hard facts. But mood matters." The proliferation of qualitative opinion regarding the Twitter-Iran issue has been helpful thus far in conveying the "mood" of the conversation, but this paper reveals some of those "hard facts" that give a fuller picture of the situation. With our report, we encourage researchers to further pursue qualitative analysis supported by quantitative data.

Data Set Description

In our research, we collected and analyzed a comprehensive body of tweets relevant to the Iran election. We gathered all tweets that use the terms listed below, either as hashtags (preceded by a # symbol) or as words (terms without a # symbol, but not including terms of two or more spaced words), between the dates of 7 June 2009 and 26 June 2009, our date of publication. Our data set begins with a tweet that occurs 5 days

before the Iran election date on 12 June 2009, to encapsulate the pre-election coverage. These tweets before the actual election date equate to 11,572 messages, or 0.57% of the entire data set.

Total tweets accumulated in this study, by term (some tweets contain multiple terms):

ahmadinejad - 1765 tweets
 basij - 3295 tweets
 gr88 - 151038 tweets
 iran - 903193 tweets
 iranelection - 857401 tweets
 iranian - 9929 tweets
 khameni - 1409 tweets
 mousavi - 16970 tweets
 mousavi1388 - 325 tweets
 neda - 97872 tweets
 rafsanjani - 77 tweets
 tehran - 85019 tweets

Our complete set of data contains the text of 2,024,166 tweets; the username and respective userID of each message; and the time when the user posted each tweet. Due to time constraints for this publication, no information was curated with regard to user account data.

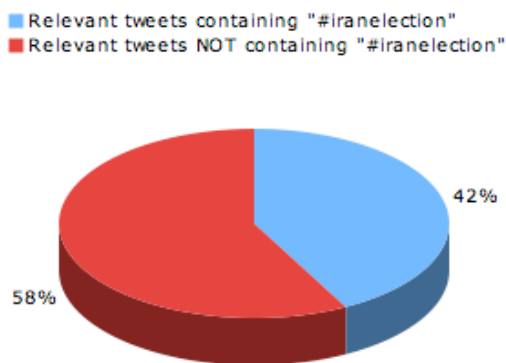


Figure 1.

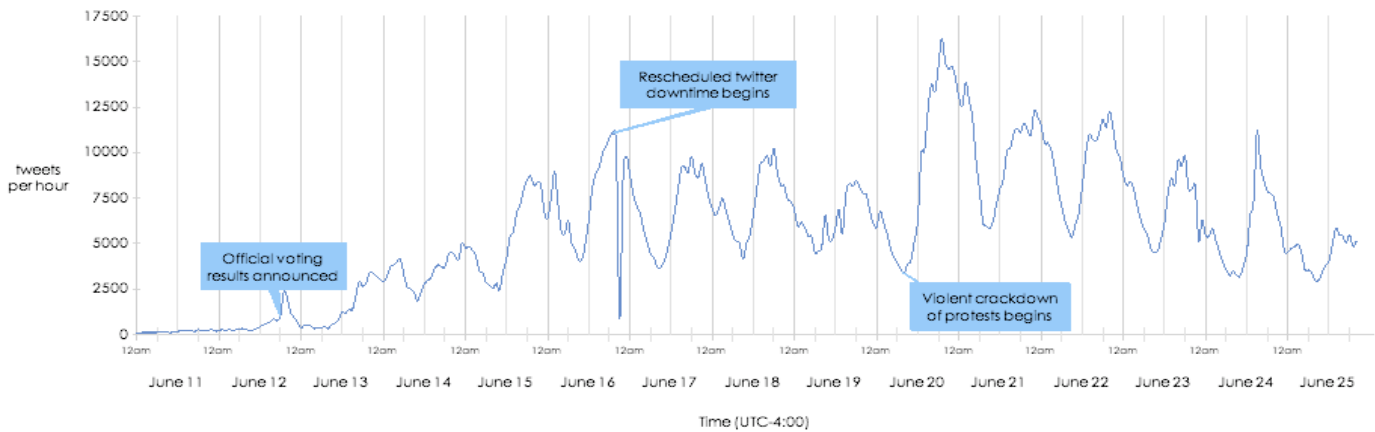
Overview of the Conversation

Early in the time period researched, a consensus among Twitter users appeared to emerge to tag a tweet with #iranelection, valuing said tweet as a relevant message related to the political events in Iran. However, our dataset clarifies that limiting the examination to tweets with only one, specific hashtag creates an incomplete understanding of the discourse as a whole.

Our method, which captures all tweets containing #iranelection as well as other related hashtags (#neda, #mousavi, #gr88, etc.) and relevant keywords (Ahmadinejad, Rafsanjani, etc.), provides access to a much larger portion of the conversation. As *Figure 1.* shows, the number of tweets using hashtags other than #iranelection amount to 1,166,765 messages, or 57.6% of the total set accumulated in our study (a significant portion of the discourse that other studies ignore when focusing solely on #iranelection). Among the total accumulation of messages, 104,127 tweets (about 5%) contain multiple (at least two) hashtags within the 140 character limit.

We can use the aggregate of collected information to illustrate the broad contours of the conversation. As seen in *Figure 2.*, the rate at which users post relevant tweets gradually increased as the events in Iran and the use of Twitter provoked attention, spiking dramatically in relation to political events inside Iran (eg., suppression of protests, as detailed in the graph), as well as in relation to news events and incidents particular to the Web.

Figure 2.



Description of the User Population

There are at least 479,780 users who have contributed to the Iran election conversation and, based on this statistic, each user broadcasts a mean of 4.22 tweets. As with most trends on Twitter, participation in the Iranian election conversation is unequally distributed. As Figure 3. (based on a Lorenz curve) shows, 59.3% of users who have contributed to the Iran election conversation account for only 14.1% of the conversation; in relation to the percentage, these users have tweeted about the events only *once*. On the other hand, the most active 10% of users, all of whom have tweeted at least 6 times, account for 65.5% of all relevant tweets. The most active 1% of users (all of whom have tweeted about the election at least 58 times) account for 32.9% of relevant tweets.

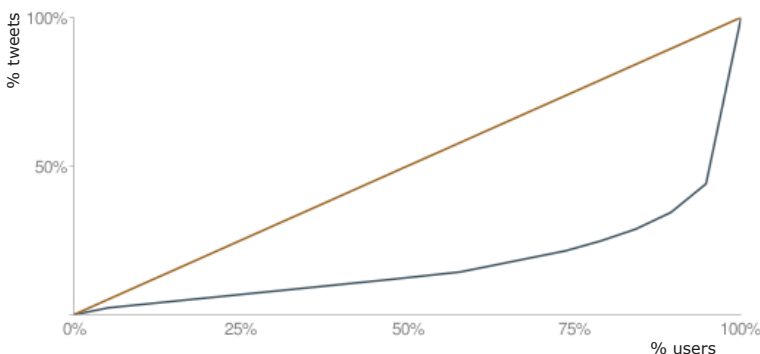


Figure 3.

A user’s relative contribution to the Iran election conversation may have little to do with the number of tweets they put out -- the loudest tweeter, in other words, is not always the most authoritative -- but we infer, for example, that those users tweeting once or twice are likely not central players in this conversation or reporting on these events very closely. We have taken a closer look at the highly active users below.

Prominent and Influential Users

Using our data set, we have attempted to identify influential users, whom we consider those users impacting the direction of the conversation, both in volume of original tweets and level of popularity among other users (both followers and additional onlookers). Appendix A lists 100 users that have contributed the largest quantity of tweets, who make up just .02% of the user base. Their tweets constitute 4.04% of our database. Some of the names listed in Appendix A, such as @oxfordgirl, may be familiar to those following the conversation; other users whose content is regularly retweeted,

such as @mousavi1388, are notably absent. Again, the loudest users are not always the most influential.

In Appendix B, we list 100 of the most retweeted users to provide another perspective on user influence. Many of the users from Appendix A, namely @oxfordgirl and @WOTN, reappear in Appendix B, but other users that do not tweet enough to merit a place in Appendix A, such as @mousavi1388 and @persiankiwi, emerge as influential players in the evolving conversation by composing relevant content that other users copy and retweet.

<http://twitspam.org/?p=1403> - 1845
<http://twitpic.com/7ki6e> - 1785
<http://digg.com/d1uPU9> - 1509
<http://iran.twazzup.com> - 1219
<http://tinyurl.com/m7w4pg> - 1197
<http://bit.ly/qmZhc> - 1196
<http://twitpic.com/7c85l> - 1193
<http://iran.twazzup.com/> - 1178
<http://emsenn.com/iran.php> - 1126
<http://bit.ly/15ROVX> - 1117
<http://bit.ly/16NJm8> - 1108
<http://www.youtube.com/watch?v=npdISZUtdmU> - 1108

Future Approaches

While this paper has provided a statistical overview of the coverage prevalent on Twitter regarding the Iranian election, our research has yet to qualitatively analyze the content of the conversation. We did, however, track trends in multiple occurrences of text strings as a preparatory step for future qualitative approaches. For example, in Appendix C, we list the 21 unique IP addresses (of proxy servers) shared among the users curated in our study. Below, we also provide a list of the most popular URLs between users (we have retained the shortened URLs, as they represent the actual text within the message).

<http://helpiranelection.com/> - 229486
<http://301.to/2iu> - 7995
<http://twitition.com/csfeo> - 6645
<http://iran.greenthumbnails.com> - 5976
<http://301.to/23o> - 3823
<http://bit.ly/xwcZY> - 2318
<http://tinyurl.com/nzxco5> - 2218
<http://gr88.tumblr.com/> - 2059

APPENDIX A**Top 100 Tweeters***username - # tweets*

DominiqueRdr - 2817
erections - 2391
Flowersophy - 2263
oxfordgirl - 2172
Dputamadre - 1400
Tymlee - 1286
WOTN - 1285
Katrinskaya - 1233
iran88 - 1230
MoraJamesLaw - 1194
anotherside - 1097
schachin - 1097
christmasfairie - 1087
ShakeyGoat - 1074
sp4rrowh4wk - 1058
zozizz - 1054
AdrienneVergara - 1042
Rezaliteit - 1023
iran_09 - 1001
thetilo - 990
ThinkIran - 972
m47713 - 958
scarletphlox - 941
irancomment - 929
ahuramazda - 921
PulseSearch - 911
loreleisigma - 901
magnolia_tree - 895
IranRiggedElect - 890
tweetstoday - 868
IranRevol - 845
iranrevolution - 834
SashaKane - 821
StopAhmadi2 - 811
oli2be - 808
MitraJoon - 798
Elizrael - 787
veganswines - 780
dreadedcandiru - 763
Mwolda - 746
eruanne - 740

jilevin - 732
trekkerGuy - 730
hardknoxfirst - 726
Winston80 - 720
SimplyDishing - 720
AlixandraLove - 718
Unstrung - 710
Hawkeye0071 - 688
Tajavioletta - 687
haverholm - 684
eaghili - 677
henksijgers - 670
JoanneMichele - 667
akhormani - 666
huichan - 665
IRANWWP - 658
DINESCU - 655
Dancinlor - 652
ruairi1338 - 635
jkslouth - 633
thefatherland - 624
iranfreeelection - 620
TerrelliC - 614
holakoozadeh - 611
greentips1388 - 608
sTavasoli - 607
MsVFAB - 603
jurassicpork59 - 602
Deskprotestor - 597
WeStandAs1 - 596
flemingcb - 595
tollwut - 593
eforsaith - 592
NoExpAffiliates - 590
B2020 - 590
mumke - 581
sagenshi - 579
arrested - 578
pmoallemian - 577
rookatpost - 574
tfsalomon - 574
IranLiveUpdates - 573
Sarah_onweb - 570
geologybabe - 567
BarbRad - 562

metabolica - 562
void00110000 - 551
motoko_nl - 551
politags - 550
reemiireem - 549
IranDemokratia - 549
Cally8 - 547
PruebaError - 544
butterflywind - 542
pilotwoman - 537
GhibliBlog - 533
AtlantaJJ - 532
IranTweet - 531
nihonmama - 531

APPENDIX B**Top 100 Retweeted Users***RT @username - # tweets retweeted*

RT @persiankiwi - 12584
 RT @StopAhmadi - 7144
 RT @oxfordgirl - 7085
 RT @BreakingNews - 5907
 RT @cnnbrk - 3828
 RT @mashable - 3354
 RT @IranRiggedElect - 2948
 RT @TehranBureau - 2945
 RT @Change_for_Iran - 2354
 RT @AnnCurry - 2291
 RT @mousavi1388 - 2283
 RT @stephenfry - 2206
 RT @ProtesterHelp - 2070
 RT @Alyssa_Milano - 1951
 RT @iran09 - 1923
 RT @jimsciuttoABC - 1838
 RT @lotfan - 1819
 RT @LaraABCNews - 1813
 RT @Jason_Pollock - 1313
 RT @IranElection09 - 1298
 RT @tweetmeme - 1272
 RT @austinheap - 1200
 RT @madyar - 1185
 RT @iranbaan - 1073
 RT @allahpundit - 1020
 RT @judyrey - 1013
 RT @IranNewsNow - 1008
 RT @zaibatsu - 956
 RT @nytimeskristof - 929
 RT @WOTN - 925
 RT @Fingertipnews - 869
 RT @TIME - 814
 RT @naseemfaqihi - 721
 RT @nytimes - 687
 RT @TimOBrienNYT - 686
 RT @whitehouse - 678
 RT @andersoncooper - 666
 RT @SashaKane - 646
 RT @nprnews - 620
 RT @cbn2 - 607
 RT: @persiankiwi - 605
 RT: @StopAhmadi - 572
 RT @iran88 - 549
 RT @huffingtonpost - 548
 RT @TheOnion - 539
 RT @shelisrael - 529
 RT @rkref - 504
 RT @dailydish - 498
 RT @TEDchris - 498
 RT @amadril - 494

RT @[username removed] - 472
 RT @timoreilly - 469
 RT @[username removed] - 452
 RT @[username removed] - 452
 RT @jaketapper - 445
 RT @katriord - 444
 RT @Katrinskaya - 444
 RT @octavianasrCNN - 439
 RT @saeedjabbar - 430
 RT @cnn - 417
 RT @flowersophy - 414
 RT @EileenLeft - 411
 RT @rainnwilson - 410
 RT: @iranbaan - 406
 RT @JasonBradbury - 406
 RT @GreatDismal - 402
 RT @GuyKawasaki - 399
 RT @Iran - 394
 RT @LilyMazahery - 381
 RT @Dputamadre - 381
 RT @SusanneUre - 378
 RT @LIFE - 370
 RT @BreakingTweets - 369
 RT @Uncucumbered - 368
 RT @PeterSantilli - 362
 RT @Elizrael - 362
 RT @jadi - 359
 RT @Tymlee - 344
 RT @CNNSaeed - 340
 RT @guardiannews - 331
 RT @dcb23 - 329
 RT @neodin - 324
 RT @michellemalkin - 320
 RT @jstrevino - 319
 RT @dominiquerdr - 314
 RT @HuffPolitics - 313
 RT @johnperrybarlow - 309
 RT @aplusk - 308
 RT @shahrzadmo - 307
 RT @Twitter_Tips - 303
 RT: @lotfan - 303
 RT @amnesty - 299
 RT @parhamdoustdar - 299
 RT @GregMitch - 297
 RT @azarnoush - 294
 RT @jeffjarvis - 289
 RT @Amysco - 286
 RT @iran - 283
 RT @LoriMoreno - 281
 RT @freedomist - 280

APPENDIX C**List of Unique IP (proxies)***IP address - # times retweeted*

218.128.112.18 - 4410
 148.233.239.24 - 2596
 128.112.139.28 - 235
 148.233.239.23 - 19
 24.238.221.163 - 4
 148.233.238.24 - 3
 131.252.214.101 - 2
 217.218.155.110 - 2
 218.28.192.10 - 2
 12.197.240.25 - 1
 24.131.125.239 - 1
 24.166.140.255 - 1
 24.19.212.141 - 1
 113.253.14.210 - 1
 124.29.215.27 - 1
 148.233.289.240 - 1
 168.143.162.100 - 1
 174.129.170.183 - 1
 193.136.191.26 - 1
 194.225.234.20 - 1
 216.24.170.159 - 1