

Wright State University
CORE Scholar

Kno.e.sis Publications

The Ohio Center of Excellence in Knowledge-
Enabled Computing (Kno.e.sis)

2010

A Qualitative Examination of Topical Tweet and Retweet Practices

Meenakshi Nagarajan
Wright State University - Main Campus

Hemant Purohit
Wright State University - Main Campus

Amit P. Sheth
Wright State University - Main Campus, amit@sc.edu

Follow this and additional works at: <https://corescholar.libraries.wright.edu/knoesis>



Part of the [Bioinformatics Commons](#), [Communication Technology and New Media Commons](#), [Databases and Information Systems Commons](#), [OS and Networks Commons](#), and the [Science and Technology Studies Commons](#)

Repository Citation

Nagarajan, M., Purohit, H., & Sheth, A. P. (2010). A Qualitative Examination of Topical Tweet and Retweet Practices. .
<https://corescholar.libraries.wright.edu/knoesis/553>

This Presentation is brought to you for free and open access by the The Ohio Center of Excellence in Knowledge-Enabled Computing (Kno.e.sis) at CORE Scholar. It has been accepted for inclusion in Kno.e.sis Publications by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.

A Qualitative Examination of Topical Tweet and Retweet Practices



Meenakshi Nagarajan, Hemant Purohit, Amit Sheth
Kno.e.sis Center, Wright State University, Dayton, OH



Overview

This study focuses on the **retweet** behavior on Twitter surrounding real-world events. The larger goal is to observe the three-dimensional dynamic of people, content & network interactions for information diffusion on Twitter.

Twitter's popularity in harnessing real-time traffic

Just minutes after President Obama's address to Congress on healthcare, Twitter showed an avalanche of tweets about the outburst from Joe Wilson. Twitter's influence was also apparent following the terrorists attack in Mumbai and in the civil reaction to the Iranian elections.

Data and Tools

Analyzed 1,677,978 topical tweets from 3 real-world events: Iran Election (IE), Health Care Reform debate (HCR), International Semantic Web Conference (ISWC)

Event	#Tweets	Date Range	Data Source (top3)	#Unique Posters
HCR	1163687	Aug13-Dec22'09	USA, Canada, Mexico	223274
IE	508959	Jun4-Jun30'09	Iran, USA, Canada	142831
ISWC	5332	Oct 19-Nov 8'09	USA, Canada, UK	2437

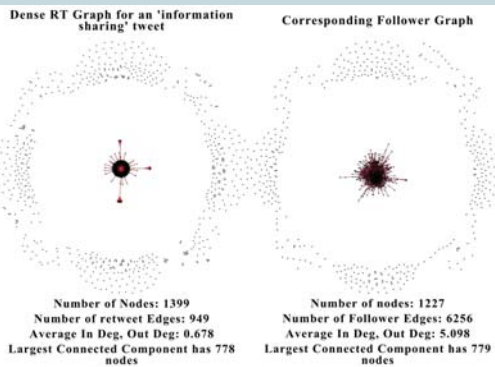
Dataset Features: varied social significance, attracting different populations, spanning different time periods and lengths of time

Patterns in Retweet (RT) Networks of Viral Tweets

• Study of top 10 viral tweets in each event. Extracted using volume indicators and captured variants of tweets using string similarity techniques.
Example: "Twtition: Google Earth to update satellite images of Tehran #Iranelection <http://twtition.com/csfeo@patrickaltoft>", and "Just signed petition 'Google Earth to update satellite images of Tehran' - <http://301.to/23o>"

• **RT Network:** NODE = Unique authors who posted the tweets; EDGE= Retweet Relationship (directed edge from A to B, if B retweets A)

• **Follower Network:** NODE = Unique authors who posted the tweets; EDGE= Follower Relationship (directed edge from A to B, if B follows A)



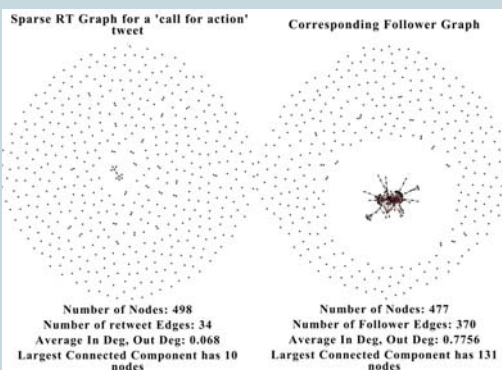
Dense RT Networks (dense author attribution)

Tweets sharing information (e.g., those containing hyperlinks to informative posts, videos, images) generated a dense retweet/attribution network. Among popular tweets of this type across the 3 events, **79%** of tweets contained author attribution information.

❖ Examples of viral Tweets that generated dense RT networks

• Information sharing

"Powerful video from @MoveOn and R.E.M. about the real lives at stake in the health care debate."
"Iran Election Crisis: 10 Incredible YouTube Videos - <http://bit.ly/vPDLo>."



Sparse RT Networks (sparse author attribution)

Among popular tweets of the type "call for action", "crowd-sourcing" or "collective group identity-making" across the 3 events, only **5%** of tweets contained author attribution information.

❖ Examples of viral Tweets that generated sparse RT networks

• Call for some sort of social action

"show support for democracy in Iran add green overlay to your Twitter avatar with 1-click".

• Collective group identity-making

"Join @MarkUdall and @BennetForCO to support an up-or-down vote on the public option".

• Crowdsourcing

"Tell John Boehner that you are one of millions of Americans who supports a public option".

Possible factors influencing sparse author attribution

- Tweets that make a "call for action" rarely credit a person. Users do not feel compelled to pass on credit to a person who acted as a messenger.
- Lack of familiarity among users in large communities might have played a role in sparse author attribution.
- Viral tweets may come from various sources. A user might not have seen the tweet from his network at all, leading to lack of attribution.
- Users are trying to make space for their content and therefore loose attribution information.

Macro-level Summaries

• A majority of users care about their voice being heard!
48% tweets in the HCR, 68% in the Iran Election and 52% of tweets in the ISWC datasets contained hashtags

• Proportion of Tweet Types

Tweet type	HCR	IE	ISWC
% directed conversations	12%	8%	23%
% retweets	27%	44%	24%
Others	61%	48%	53%

• Proportion of Poster Types

Most active posters: News and marketer profiles
Most mentioned and retweeted authors: Individual users

Conclusions

• The content being tweeted plays a key role in what an explicit retweet network will look like and in many cases, whether it will be traceable at all.

• The observed relationship between the tweet type and its retweet pattern has implications on link-based diffusion models and provenance studies.

• Findings after submission: A quantitative study of language properties of 300+ viral tweets indicates significant correlations between impersonal pronouns and verb usages in tweets (typical of making a call for action) and the sparse attribution networks they generate.