Retweets indicate Agreement, Endorsement, Trust: A Meta-Analysis of Published Twitter Research

Panagiotis Metaxas and TwitterTrails Research Team

Computer Science Department, Wellesley College

Abstract

Since its introduction, Twitter has proven to be an increasingly important platform in Social Media research. A significant number of research papers studying Social Media these days collect and analyze data using the Twitter API. The collected data are used to perform mainly observational studies, while a few researchers have also started performing experimental studies on Twitter.

Arguably one of the most important features of Twitter is the support for "retweets" or messages re-posted verbatim by a user that were originated by someone else. (This does not include "modified tweets" that sometimes are confused as retweets.) Despite the fact that retweets are routinely studied, many important questions remain about their use and significance. Importantly, casual reading of the literature does not reveal an obvious answer to the question of why do people retweet, or what affects the rates of retweets observed in various corpora.

In this paper we present a meta-analysis of over 100 research publications examined for clues about fundamental questions regarding retweets. Starting our survey with relevant papers published between 2008 and 2013 in three major conference venues, AAAI ICWSM, IEEE SocialCom, and WWW, we expanded the coverage to cover other references found in these venues.

Our findings indicate that retweeting is, under specific conditions, a form of expression of agreement with the message, and endorsement –or even trust– of the message originator. The specific conditions are related to expression of opinions influenced by emotion and intention. The existence of emotion and intention that can be detected by the presence of hashtags, are responsible for the variability of retweet rates in a domain. While there have been additional claims by researchers about the possible reasons for retweeting in the past, most of them are not supported. Moreover, the technical changes introduced recently by Twitter make these additional claims irrelevant.

Keywords: Social Media, Twitter, Social Networks, Social Theorems

1. Introduction

Twitter is a real-time information network[1] that allows its users to write short messages ("tweets") up to 140 characters in length. Created in 2006, Twitter has become an enormously successful platform and many research papers on social media study phenomena related to its service. We chose to focus on Twitter in particular because, unlike other popular social media platforms, it has, in fact, grown into a real-time news source created by everyday users. Twitter is credited for its role in political events such as monitoring elections, the so-called "Arab Spring" [BHB13], and for drawing attention to news stories that were largely ignored by traditional news media such as the Wendy Davis[2] filibuster.

We will add more stuff related to the paper's findings. Why are we interested in the problem. Why others are interested.

1.1. How Twitter Works

Over the years Twitter has developed its own syntactic components. Users may choose to use hashtags (#) to tag a tweet and indicate that it is relevant to a particular topic or event (e.g. #election2012). Bursty popular hashtags will occasionally be featured in the "Trends" section on Twitter. Users may also decide to mention one another by adding "@[user_account]" to their tweets. Mentions direct a tweet at a specific user. If a user is mentioned in a tweet, the tweet will appear in the "Interactions" tab of the Twitter user's homepage. Multiple hashtags and user mentions can be used in a single tweet.

Users can choose to "follow" others as a way of being informed of the tweets of those they choose to follow. Following is not a symmetric action and the user being followed does not have to follow back or even agree on the follower's decision to follow them.

Users can also **retweet (RT)** to actively forward a message from another source to their own followers. Retweeting was not part of the original design of Twitter operations but has became popular quickly. Because of its later adoption date, tweet forwarding can be done in one of two ways, one that is supported by the Twitter API and another that is not: either by clicking a "Retweet" button provided by the Twitter client (and some other clients), or by manually typing "RT @[user]" or adding "via @[user]" in a new tweet. We call the tweets produced by typing " modified retweets" (MRTs) to distinguish them from the first kind. This distinction is important because an unmodified retweet is treated differently by the Twitter platform and is guaranteed to point to its originating source.

[1] Twitter is widely known as a micro-blogging service, and until Jan. 2010 was describing itself as "real-time short messaging service that works over multiple networks and devices". This changed in February, 2010 when it started describing itself as "real-time information network powered by people all around the world that lets you share and discover whats happening now."

[2] Texas filibuster on abortion bill rivets online, by Heather Kelly. June 26, 2013. http://www.cnn.com/2013/06/26/tech/social-media/texas-filibuster-twitter

1.2. Why focusing on Twitter

As a platform, Twitter has several characteristics that make it a convenient research platform. First, it has a wide and increasing base. It is reported[3] that on its 7th birthday, in March 2013, there were over 500 million Twitter accounts world wide sending half-a-billion tweets every day, or about 6 thousand tweets/sec. Second, compared to other social networks, it has a simple API that enables the collection of data related to specific keyword, specific account activity, and in real time (albeit one can only receive 1% of sampled data for free). And, third, it has been extensively studied by Social Media researchers since its creation, so there are many data points for comparison.

Nevertheless, no matter how successful, it is a valid question to ask ourselves whether it is interesting to study a specific platform such as Twitter. If Twitter does not exist in, say, 10 years, will this paper's findings matter?

Twitter is recording human communication that requires relatively little effort to produce and consume. While any particular social media platform may cease to exist or lose popularity in the future, the importance of human interaction through social media is unlikely to change. Humans are social animals and their desire to communicate with each other and comment on their social environments is one of their universal and unique characteristics. As recent and measurable evidence of this fact one can see the tremendous and continuous growth that social media have enjoyed since their creation. We have chosen to study the interaction of humans through social media in an abstract way, not a way specific to the particular social media platform. We are simply looking at the behavior as revealed through their interactions.

While there is nothing unique about the Twitter platform, its service makes it easy for people to say something. The effort in contributing to the general social dialog is far less than that of writing a comment a blog on a web site or a newspaper op-ed, and it has wider impact than talking person-to-person or via email. In addition, the effort to propagate a message sent by someone else is also remarkably small – giving rise to degrading characterizations of online social participation such as "slactivism" [Gla10].

2. On the Variability of Retweeting Rate

The ease in repeating something via RT can give insight into how people think and act in social media. Patterns of retweets offer the opportunity in measuring the opinion of individuals or groups, but also organizing them, making them aware of the extent that others share their opinion and attempting to influence others' opinions.

In this paper we study the use and significance of (unmodified) retweets on Twitter. We are looking for clues to answering the question "Why do people retweet?" It turns out that, while there have been many papers studying

[3] Twitter In Numbers, by Richard Holt, The Telegraph, 21 March 2013. http://www.telegraph.co.uk/technology/twitter/9945505/Twitter-in-numbers.html

retweets, there are not many that have tried to address this question directly. Most have attempted to provide an opinion in passing. Nevertheless, we wanted to know what is the collective knowledge of all this research. To focus our research question a bit further, we did a meta-analysis of the research corpus that has been published in the last several years on Twitter.

Retweeting was not part of the original design of Twitter, but was created through user initiative. Its adoption is a prime example of *use influencing design*. The term "Re-Tweet" can be traced back to a tweet sent[4] in March 2007, and retweeting was first mentioned as an act[5] shortly after, in April 2007. Use of the form "RT @user:" became a Twitter convention in approximately July 2008 [RBC+11] but it was not supported by the Twitter API[6] in the json format until August, 2009[7]. The introduction of the "retweet_count" field in json data, supported by the official retweet button was also introduced then but the button's utilization initially was slow. For example, [SHPC10] counted 2.99M retweets generated by the one-click retweet button on Twitter, which was 36.34% of the total of 8.24M retweets collected in a corpus of 74M tweets. So, by March 2010, about two thirds of retweets were created manually, while in a collection of 39M tweets in mid-summer of 2012 about one fifth of retweets are created manually by users.

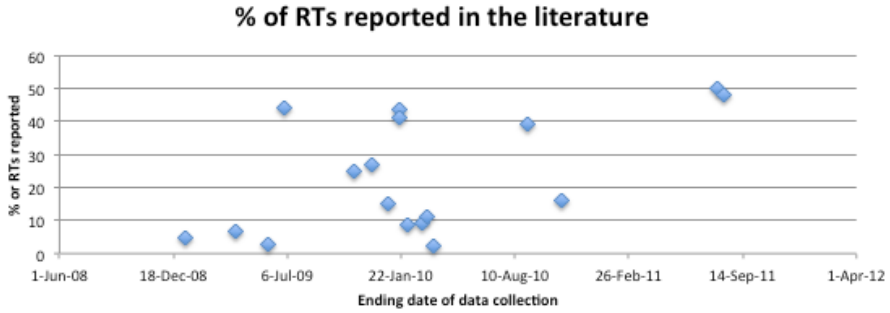

Figure 1: There seems to be little agreement in the literature on the percentage of retweets one might expect to observe in a corpus collected. Reported percentages range from a low of 3% to a high of 50%.

There is extensive work on **what** gets retweeted ([LG10], [POL11], [SHPC10], [PGS12], [ZJW+11], [NPS10], [RBC+11], [KA11], [HDD11], [AGK], [VL10], [HAN+11]), and on **who** is retweeted (and, in turn, who retweets others) ([BAH12], [LOTW13], [TAG13], [RAZ11], [Mur12], [TPT12], [LGA+11]). However, as we will show in the next section, there is no consensus in the literature

[4]https://twitter.com/derekpunsalan/status/8500871
[5]https://twitter.com/ericrice/status/31669791
[6]https://blog.twitter.com/2009/project-retweet-phase-one
[7]http://evhead.com/2009/11/why-retweet-works-way-it-does.html

on **why** people retweet. We will start by pointing to the fact that there is little agreement even for the reported percentage of retweeted content (see Fig. 1). Understanding the root of this disagreement turns out to be important.

Starting	**Ending coll. date**	**% of RTs**	**Citation**
1-Jan-09	7-Jan-09	4.9	[RBC+11]
1-Apr-09	7-Apr-09	6.8	[RBC+11]
1-Jan-09	1-Jun-09	3	[BGL10]
4-Jun-09	30-Jun-09	44	[NPS10]
1-Oct-09	1-Nov-09	25	[NPS10]
1-Aug-09	1-Dec-09	27	[NPS10]
1-Jun-09	31-Dec-09	15.07	[YL11]
13-Jan-10	20-Jan-10	43.5	[MM11]
13-Jan-10	20-Jan-10	41	[MM10]
11-Nov-09	1-Feb-10	8.46	[POL10]
1-Feb-10	1-Mar-10	9	[PGS12]
8-Jan-10	8-Mar-10	11.15	[SHPC10]
19-Mar-10	19-Mar-10	2.19	[SHPC10]
1-Aug-10	1-Sep-10	39	[MM11]
26-Oct-10	1-Nov-10	16	[MM11]
1-Nov-10	1-Aug-11	50	[MMFMH12]
9-Aug-11	11-Aug-11	48	[TPT12]

Table 1: Reported percentages of retweets in the literature, sorted by the ending collection date. There seems to be a no convergence on the percentage of retweets one might expect to observe, though it appears that the rate is increasing over time. Understanding the root of this discrepancy holds part of the answer the the question of why people retweet.

According to [PGS12], 9% of tweets are retweets, observed in a corpus collected in early 2010. This is significantly higher that the 3% reported by [BGL10] in a sample set from early 2009. One might assume that this reflects a growth in the use of retweets over the one-year period, but [YL11] in their late 2009 corpus found 15% retweets. Also in early 2010, [SHPC10] randomly collected 74M tweets (estimated to be 2% to 3% of all tweets created during the collection period) and found 8.24M retweets. These were selected through regular expression matching (retweets that have text markers like "RT@" or "retweet @") and accounted for 11.15% of their tweet corpus.

At the other end of the spectrum, [MM10] report that 41% of the tweets related to MA Special Senatorial Elections of 2010 were retweets and [MMFMH12] report 50% retweets related to the so-called "narco-tweets" (tweets informing of risk situations in drug war-torn Mexico) in 2011. It was noted that the existence of highly publicized events and the inclusion of URLs in the tweet seemed to affect the percentage of retweets reported: [NPS10] collected tweets in 3 different specific topical areas, and they found that highest proportion of retweets (44%) was in a set on the recent elections in Iran. [TPT12] used a dataset from the 2011 London Riots and found that 48% of the dataset were retweets, and that the retweets were over twice as likely to contain URL links compared to regular

tweets.

2.1. Insight: The role of hashtags

While the divergence of reported retweet rates may seem mystifying, things become clearer when one considers the role of hashtags in a tweet. Hashtags give the opportunity to experienced messengers of enabling the discovery of their message far beyond the community of their followers, through the main Twitter search facility and even promoting it to a trending topic. This, turns out, is an essential element in understanding the variations of retweet rates in the literature.

In fact, sorting the data in Table 1 according to the retweet rates and adding a column with the method used by the researchers in collecting the data (uniform random sampling versus collection through hashtags/keywords) we see a clear separation in the data of Table 2. Random sampling finds between 3% and 16% of retweets, while collecting through hashtags/keywords finds between 25% and 50% of retweet rates. See Fig. 2. This finding is in sync with the observation by [SHPC10] that "URLs and hashtags have strong relationship with retweetability." So, when collecting tweets based on hashtags, it should not be surprising to observe a higher percentage of RTs.

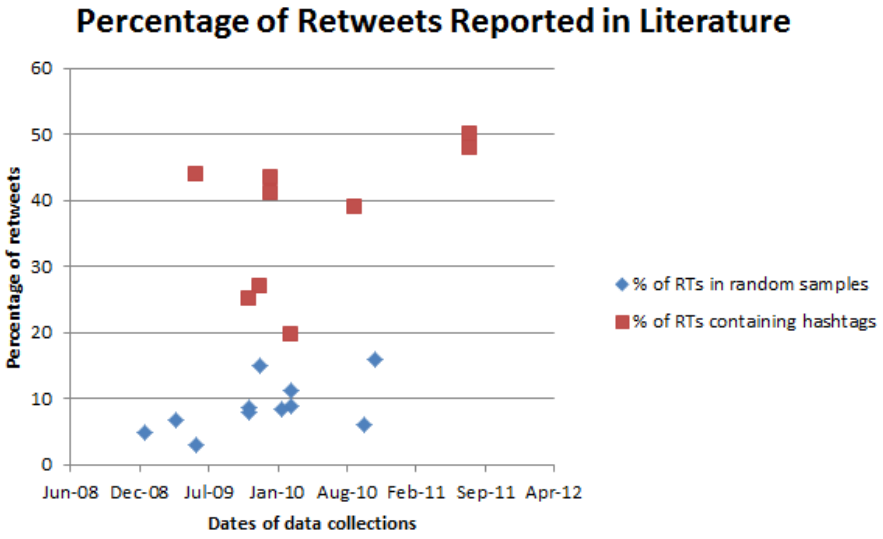

Figure 2: The mode used to collect Twitter data shows a clear separation between reported rates of retweets.

Moreover, one can observe in Table 2 that there is some correlation with emotional strength among the hashtags. Neutral (not emotionally charged) hashtags, such as those relating to a conference, record 25% retweet rates, general political (non-election) hashtags record between 27% and 39%, elections

% of RTs reported	Citation	Sampling method
50	[MMFMH12]	#MTYfollow
48	[TPT12]	#londonriot, #riotcleanup
44	[NPS10]	Iranian election (hashtags and keywords)
43.5	[MM11]	#MAsen10
41	[MM10]	Coakley, Scott Brown
39	[MM11]	#tcot, #p2
27	[NPS10]	health care reform (hashtags and keywords))
25	[NPS10]	ISCW (hashtags and keywords))
16	[MM11]	random collection
15.07	[YL11]	random collection
11.15	[SHPC10]	random collection
9	[PGS12]	random collection
8.46	[POL10]	random collection
6.8	[RBC+11]	random collection
4.9	[RBC+11]	random collection
3	[BGL10]	random collection
2.19	[SHPC10]	random collection

Table 2: Number of Retweets reported in the literature, sorted by the decreasing percentage of retweets found in the corpus. A column describing the method used for collecting the data is added.

range between 41% and 44%, and issues related to public safety contain between 48% and 50% retweet rates.

While we do not mean to imply that there is a clear distinction between various levels of emotion and intention, there is some correlation in the sense that

> *intentional hashtags, corresponding to more emotionally charged issues, result in higher retweet rates than those with less;*
> *and the latter result to higher rates than random collection of tweets (i.e., that may or may not contain hashtags)*

2.2. Newer Data on Retweet Rates

Wanting to verify that the reported retweet rates were still operating beyond the dates of reporting literature, we looked into more recent collection of tweets, of our own and of other researchers. What we found is that the trend that was reported in the literature has increased and tweets collected through

3. Reasons for Retweeting

So, it appears that messages with a topical marker (a hashtag or a rare keyword) are retweeted much more often that random messages. Can we tell something more about those messages? We will examine the literature for answers to why people retweet. There are many claims that have been made in

% RTs	End date	Hashtags	Users	Tweets	Retweets
73.7	12-Aug-13	Trending HT: #ScienceSaysSo	23201	32879	24223
73.3	16-May-13	Trending HT: #DoYourJobGOP	6462	20738	15191
61.1	13-Jul-13	Zimmerman trial verdit	1176298	2276175	1390800
60.0	2-Nov-12	Jobs Report	25499	37954	22766
57.7	6-Nov-12	MA Senatorial election 2012	43327	135319	78087
51.4	19-Aug-13	Obama's new puppy announcement	5168	5783	2970
50.6	22-Oct-12	2012 3rd Presidential Debate	852927	1668534	844264
48.6	22-Jun-13	MA Senatorial election 2013	38633	84210	40941
47.5	16-Oct-12	2012 2nd Presidential Debate	762101	1297076	615968
46.4	3-Oct-12	2012 1st Presidential Debate	1010941	1994906	925850
46.1	5-Jun-12	Wisconsin congressm. recall	223854	684449	315288
43.4	4-Aug-13	12th Doctor Announcement	211851	393203	170707
38.9	15-Feb-13	Trending HT: #ThoughtsInClass	60591	94365	36704
37.9	14-Dec-12	Guns-related	3725216	9495789	3599077
36.2	24-Sep-13	German elections 2013	467930	1327380	480375
34.8	14-Feb-13	Valentintes Day	819524	997464	346698
27.6	20-Aug-13	Misha Collins' Birthday	33175	68140	18814
27.2	12-Aug-13	Whitey Bulger trial verdit	16265	30332	8259
20.9	20-Feb-13	CBS Survivor (2013)	51589	124737	26040
20.7	6-Mar-13	Twitter Event: #AskThorin	2551	8716	1803
20.6	12-Jun-12	random(gardenhose) collection	12756587	39801489	8187704

Table 3: Newer data. See Fig. 3

the past about why people retweet, many of them made *en passé*, based on "common sense" but without strong data support. In the following subsections we will group related claims and we will characterize them according to the evidence cited by the authors.

3.1. Straightforward and Outdated Reasons

The first set of claims can be characterized as straightforward (in the sense of being self-evident) and outdated (in the sense that technical changes in Twitter do not support them anymore). As an example of straightforward claim is the claim that people retweet to broadcast and appropriate information. As an example of outdated claim is the claim that people retweet to appropriate the information.

According to [RAZ11]: "[Retweeting] was first a social appropriation done by manually copying someone's tweet and prefixing it with the letters "RT", so that other users would know that that was a copied content. Any addition to the information from the user retweeting normally (but not exclusively) appears before the "RT". Since then, Twitter has implemented a dedicated functionality that mimics this behavior to some extent."

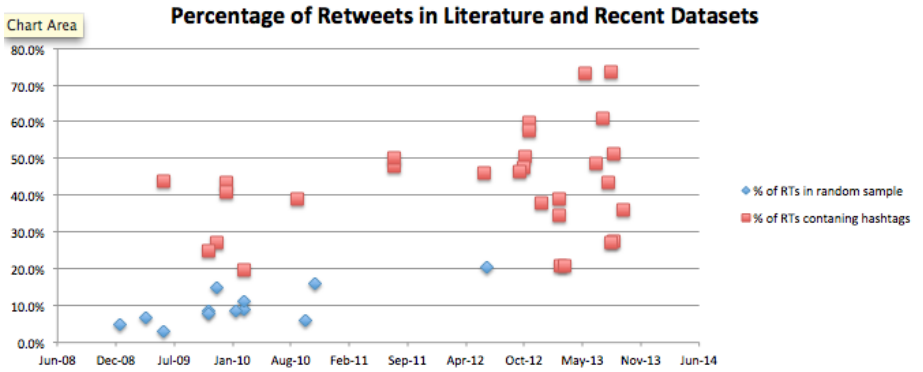

Figure 3: Recent data collected shows an increased trend of RTs, yet the rates of randomly selected data and hashtag-selected data are separable.

3.1.1. Straightforward: Retweeting as a form of broadcast, promotion

Many researchers elaborate further in the category of straightforward reasons by mentioning a "desire to promote issue" [PB12], [BGL10], "dissemination of breaking news" [TAG13], defining retweet as "a mechanism for information diffusion" [LGRC12], or noting that retweeting enables "users to propagate information across multiple hops in the network through word-of-mouth" [RBC+11]. Reflecting the information related to popular events (such as the latest Royal wedding in the UK, a Japanese earthquake, and the Super Bowl in the US) [TAH+13] point to the importance of timing in the retweets: "During these events, breaking news are often retweeted not long after being posted." In fact, [ZJW+11] found that the largest proportion of retweets they observed was event-oriented in the "World" category, suggesting a large number of retweets pertain to world news events that others are likely to want to know about.

3.1.2. Straightforward: Retweeting expresses interest

Another straightforward characterization refers to retweets as reflecting something that is "interesting": For example, [KA11] mention that the usual purpose of retweeting is "when users find a message particularly interesting" and describe retweet as "a measure of interestingness." However, one should note that "interesting" is a weak characterization since it understates emotional relevance. Indeed, it is mentioned as self-evident that "retweet shows a strong interest on the topic of the tweet" by the user ([WWB+13], [PGS12]) or by the community that the user participates [HDD11].

3.1.3. Outdated: Retweeting as an act of personal curation or appropriation

Several early papers mention personal curation or appropriation as a reason for retweeting, that is, an action by the user to save tweets for later [BGL10], [SHPC10]. Simimarly, [AGK] mentions that "a retweet allows Twitter users to

rebroadcast specific tweets of interest by incorporating all or part of the original tweet into their own."

While these may have been reasons for retweeting in the early days of Twitter, it does not appear that this practice is still prominent. Appropriation is not supported by the new API which would not mark a user's modified tweet as a retweet. And the introduction of a "favorite" button for a tweet is now playing the role of saving tweets for later.

3.1.4. Outdated: Retweeting as a conversational or social act

Another set of now outdated claims is related to the view of retweets as a conversational or social act. Quite a few of early papers have made this observation. For example, [BGL10] views retweets as a social assertion; by retweeting, one is establishing oneself as a community member related to the topic, to the originator, or both. [RAZ11] see retweets as a win-win social situation, claiming that "retweets play an important part in gathering social capital." Finally, [NPS10] see "retweets are a means of participating in a diffuse conversation."

However, these claims really refer to modified tweets, a distinction that did not exist in the early days of Twitter. They are based on early observations where retweeting was of very limited practice (e.g., only 3% of the [BGL10] corpus were retweets) and involved modifying a tweet by including both the username of the originator as well as the retweeting user, thus giving visibility to the retweeterWhile this practice was exciting to researchers since it was making it possible to study information diffusion through retweet trees, the Twitter API does not allow it anymore since now only the originator gets credit in the retweeting process.

4. Retweeting as a form of agreeing, endorsing, trusting

The most strongly supported claim in the literature about the intention behind retweets is that they indicate some form of endorsement of the opinion expressed, the originator of the opinion, or both. This becomes apparent when one realizes that these claims come from authors that examine corpuses collected using hashtags and keywords. As we saw, such collection indicate a higher percentage of retweets and are focusing on emotionally charged discussion. We will examine these claims in some detail in this section.

4.1. Supported: Retweets indicate agreement with the topic or the opinion

One of the more consistently supported claims is that retweeting suggests agreement of the retweeter with the topic of information or the opinion expressed. [BGL10] report that retweets show agreement publicly and contribute in getting a specific topic to trend[8]. This claim finds strong support in political

[8]We should note that this highly cited work collected data using a survey with an opportunity sample of early Twitter users

discussions, in particular during elections: [MM10] show how retweet networks[9] can detect political orientation of users using force-directed graph-drawing algorithms ([**?**]) and community-detection algorithm (such as the [**?**] algorithm). Their results were supported by [CGR+11] who repeated the experiment during later elections. Essentially the same claim is supported by [TPT12] who note that "Twitter users retweeted to show support for their beliefs in others' commentaries." Similarly, [PB12] find that retweets provide for lightweight protest in political or confrontational events.

In addition to the studies above, several other researchers offer support to this claim based on their experience with their corpuses. [AGK] writes that retweeting "commonly serves as a way of saying 'me too,' in response to a users tweet." [KLPM10] argues that retweeting "empowers users to spread information of their choice beyond the reach of the original tweets followers," and [TPT12] write that retweets are a way to incorporate information into your own existing belief system. All of these statements effectively accept the claim about user agreement with the topic. Moreover, [VL10] argue that "The person who retweets has specifically chosen to retweet that tweet which can be seen as an endorsement of that particular piece of information. Retweeting can be thought of as a vote for the quality, novelty or timeliness of a piece of information." While [VL10] use the word "endorsement", we include their claim in the "agreement" category, with the disclaimer that often these two terms, along with the term "trust" are used in literature as roughly equivalent terms.

4.2. Supported: Retweeting as endorsement of the author

A stronger statement about the role of retweets is that they are effectively an endorsement of the originators, not just in agreement with their opinion of the topic being discussed. It is, of course a fine line to make the distinction between supporting someone's opinion with endorsing the individual, but when it comes to political opinions, there seems to be little doubt about the strength of the relationship. [CGR+11] study is aligning twitter users politically, based on what they mention and what they retweet. They find that "Retweets act as a form of endorsement, allowing individuals to rebroadcast content generated by other users, thus raising the contents visibility." One can see the findings of [MM10] and [TPT12] as supporting the same claim. Recently, [WGB13] studying tweets in both Arabic and English found that "Retweeting signifies endorsement. Using simple retweet information we could label users as either Islamist or secular with accuracy similar to inter-judge agreement." In even stronger terms, [CGVMA11] base the effectiveness of their research on the claims that "retweets are endorsements in which a user propagates a message posted by another user to their list followers." Finally, it can be argued that when people retweet a celebrity's (e.g. Justin Bieber's) messages, sometimes indiscriminately, retweeting count as an endorsement to the celebrity.

[9] graphs depicting users as nodes and retweets as directed edges from originator to retreater

Other researchers support the endorsement claim as well, but based on their experience with their data, not by conducting specific experiments. [BB12] write "Retweeting can also be interpreted as an implicit endorsement for message and sender, unless additional commentary is added by the retweeter during retweeting."

4.2.1. On the claim that "Retweeting does not mean endorsement"

Of particular interest is the statement encountered in certain Twitter users' profile summary that "retweet does not mean endorsement." It seems that this statement contradicts the research presented in this section so we should examine it carefully. First we analyze its semantics and then we present quantitative results from our research.

One has to consider the reason for people placing such a clarification statement in their profile. The need for such clarification can be seen as a concern that one's practices may be mistaken by others. It is reasonable to argue that the mere fact of including this disclaimer is an implicit admission by those using it that others may mistake their intentions. Effectively it is an admission that, apparently, for most people retweet *is* endorsement.

There are relatively few users making this claim, and we have found that the largest groups are journalists or bloggers. In particular, we examined two sets of profiles, the first collected with keywords related to the Snowden revelations that were initially promoted by news organizations, and the second with keywords related to guns. Our collections were reasonably large with 140,053 and 3,736,618 profiles respectively. In the first collection, we found that the statement appeared in less than 1.5% of the profiles and in the "guns" collection in less than 0.2% of the profiles. Therefore it is a tiny percentage of users that feel the need to make such a statement. Nevertheless, we found that the statement was overrepresented among journalists and bloggers: about 45% of them containing it (38% journalist and 6% bloggers). Another interesting observation is that among those profiles who make this statement, almost 40% of them also make the statement "Tweets represent personal opinions!" It is not clear if these people do not see the oxymoron of their claims or they try to protect themselves from defending what they retweet. Several journalists have been challenged for this claim (e.g., see Sreenivasan's commentary at the Washington Post[10]) and have even gotten in trouble in a few circumstances for using it [**?**] since its inclusion in one's profile is not an effective disclaimer (many people will likely miss it).

4.2.2. Special case of endorsement: Retweeting as indication of influence of originator

Another way to examine the concept that retweet is indicative of endorsement is to look at the relationship from the other end. A number of authors use retweets as an indication of the influence tweet originators have on those

[10] http://wapo.st/19OIcfy

retweeting their messages. [YW10] define as the degree of influence the number of retweets an originator gets: "The more frequently the users messages are retweeted by others, the more influential this user is." Further, [YW10] comment that users with stable retweet counts are more likely to "keep their influence stable". [TC12] agree with this definition, arguing that "the retweets that a user incurs is a suitable indication of their influence." Along the same lines, [GWT11] "use retweeting behavior as an indicator of influence in this community and count how many tweets are being retweeted by other members of this community."

4.3. Supported: Retweeting as an expression of various forms of trust of originator

Some authors see retweets as an expression of trust in the person being retweeted. In particular, [TAG13] define trust as a psychological attitude of A towards B with respect to some possible desirable behavior. In particular, they write that trust is "the idea that an interaction between two trusted parties will have an expected outcome" and they find that the higher the frequency of retweets, the higher the trust.

Other authors see retweets as trust in the validity of information being retweeted. More specifically, [MMFMH12], [MM13] argue that retweets indicate trust in the validity of safety-related information. This claim is verified by real-life behavior of users who reportedly check Twitter before getting out of their house for safety information put forth by the anonymous "curators" or prominent Twitter users who have strengthen their reputation over long periods of retweeting times (months or years). The authors note that the higher the frequency of retweets, the higher the trust in the originator of the information.

The above views on trust are supported by [AEG+10] who "make the assumption that when a user propagates information from some other user, there must be some element of trust between the two users." These authors claim that when a "node propagates information from another then it suggests that the propagator trusts the information. [...] a repeated propagation makes the conclusion stronger." In fact, retweeting is used in [AEG+10] as a way to compute trusting communities. Their main experimental result is that the behavioral trust graphs do indeed represent trust (at least as captured by retweets). They derive their results based on the likelihood (14.4%) of users to retweet repetitively within such communities, compared to random model's 3%).

In a more specific domain of trust, [LGA+11] analyzed information flows during the 2011 Egyptian/Tunisian revolutions and found that journalists tended to retweet other journalists more than bloggers and others. They interpret this behavior as an indication that those journalists trust the information provided by other journalists more than other people.

5. On the role of MTs (modified tweets)

6. Conclusions

In this paper we conducted a meta-analysis of the research literature on the use of retweets to answer the question "why people retweet".

Our conclusions are as follows: Retweeting does indicate a level of endorsement of the message and/or the originator. While the exact sense may be difficult to determine automatically, there is evidence that the stronger the emotions surrounding the message, the stronger the endorsement. We have found evidence in the literature that this endorsement may range from topic agreement to originator endorsement/expression of trust.

Disclaimers:

1. When we refer to "retweets" we mean propagation of unedited messages to one's follower network. Modified tweets (those that add the address of the retweeter or alter the original message in some way) are not included in our conclusions.
2. Our conclusion describes all retweets related to a subject that are typically retrieved through a hashtag or keyword search. It does not imply that every retweet is subject to the characterization of endorsement, since a particular retweet may be simple the result of unintentional actions such as confusion, accident, hack, sarcasm, etc.

We believe that our findings are an important contribution to social media research since the meaning of retweets has been the subject of dispute for as long as retweets exist. An important insight into our conlusion comes from the use of hashtags in a retweet, inserted by the originator with the clear intention of spreading of the message beyond the immediate community of followers. The retweeter effectively decides that wants to participate in this propagation and as a result we observe dramatically higher rates or retweets in messages containing a hashtag or related keywords.

References

[AEG+10] Sibel Adali, Robert Escriva, Mark K Goldberg, Mykola Hayvanovych, Malik Magdon-Ismail, Boleslaw K Szymanski, William A Wallace, and Gregory Williams. Measuring behavioral trust in social networks. In *Intelligence and Security Informatics (ISI), 2010 IEEE International Conference on*, pages 150–152. IEEE, 2010.

[AGK] Rio Akasaka, Patrick Grafe, and Makoto Kondo. me too2.0: An analysis of viral retweets on the twittersphere.

[BAH12] Roja Bandari, Sitaram Asur, and Bernardo A Huberman. The pulse of news in social media: Forecasting popularity. In *ICWSM*, 2012.

[BB12] Axel Bruns and Jean Burgess. Researching news discussion on twitter: New methodologies. *Journalism Studies*, 13(5-6):801–814, 2012.

[BGL10] Danah Boyd, Scott Golder, and Gilad Lotan. Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on*, pages 1–10. IEEE, 2010.

[BHB13] Axel Bruns, Tim Highfield, and Jean Burgess. The arab spring and social media audiences: English and arabic twitter users and their networks. *American Behavioral Scientist*, 57(7):871–898, 2013.

[Bru12a] Axel Bruns. Ad hoc innovation by users of social networks: The case of twitter. *ZSI Discussion Paper*, 16(2012):1–13, 2012.

[Bru12b] Axel Bruns. How long is a tweet? mapping dynamic conversation networks on twitter using gawk and gephi. *Information, Communication & Society*, 15(9):1323–1351, 2012.

[CGR+11] Michael D Conover, Bruno Gonçalves, Jacob Ratkiewicz, Alessandro Flammini, and Filippo Menczer. Predicting the political alignment of twitter users. In *Privacy, security, risk and trust (passat), 2011 ieee third international conference on and 2011 ieee third international conference on social computing (socialcom)*, pages 192–199. IEEE, 2011.

[CGVMA11] Pedro Henrique Calais Guerra, Adriano Veloso, Wagner Meira, Jr., and Virgílio Almeida. From bias to opinion: a transfer-learning approach to real-time sentiment analysis. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '11, pages 150–158, New York, NY, USA, 2011. ACM.

[DCLS+10] Munmun De Choudhury, Yu-Ru Lin, Hari Sundaram, K Selcuk Candan, Lexing Xie, and Aisling Kelliher. How does the data sampling strategy impact the discovery of information diffusion in social media? *ICWSM*, 10:34–41, 2010.

[Gla10] Malcom Gladwell. Small change: Why the revolution will not be tweeted. *The New Yorker*, October 4, 2010.

[GTPZ10] Chris Grier, Kurt Thomas, Vern Paxson, and Michael Zhang. @ spam: the underground on 140 characters or less. In *Proceedings of the 17th ACM conference on Computer and communications security*, pages 27–37. ACM, 2010.

[GWT11] Anatoliy Gruzd, Barry Wellman, and Yuri Takhteyev. Imagining twitter as an imagined community. *American Behavioral Scientist*, 55(10):1294–1318, 2011.

[HAN+11] Lars Kai Hansen, Adam Arvidsson, Finn Årup Nielsen, Elanor Colleoni, and Michael Etter. Good friends, bad news-affect and virality in twitter. In *Future Information Technology*, pages 34–43. Springer, 2011.

[HDD11] Liangjie Hong, Ovidiu Dan, and Brian D Davison. Predicting popular messages in twitter. In *Proceedings of the 20th international conference companion on World wide web*, pages 57–58. ACM, 2011.

[HLZ12] Alfred Hermida, Seth C Lewis, and Rodrigo Zamith. Sourcing the arab spring: A case study of andy carvins sources during the tunisian and egyptian revolutions. In *International Symposium on Online Journalism, Austin, TX, April*, pages 20–21, 2012.

[KA11] Nasir Naveed Thomas Gottron Jérôme Kunegis and Arifah Che Alhadi. Bad news travel fast: A content-based analysis of interestingness on twitter. 2011.

[KLPM10] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pages 591–600. ACM, 2010.

[LG10] Kristina Lerman and Rumi Ghosh. Information contagion: An empirical study of the spread of news on digg and twitter social networks. *ICWSM*, 10:90–97, 2010.

[LGA+11] Gilad Lotan, Erhardt Graeff, Mike Ananny, Devin Gaffney, Ian Pearce, and Danah Boyd. The revolutions were tweeted: Information flows during the 2011 tunisian and egyptian revolutions. *International Journal of Communication*, 5:1375–1405, 2011.

[LGRC12] Janette Lehmann, Bruno Gonçalves, José J Ramasco, and Ciro Cattuto. Dynamical classes of collective attention in twitter. In *Proceedings of the 21st international conference on World Wide Web*, pages 251–260. ACM, 2012.

[LOTW13] Zhunchen Luo, Miles Osborne, Jintao Tang, and Ting Wang. Who will retweet me? finding retweeters in twitter. 2013.

[MM10] Eni Mustafaraj and P Takis Metaxas. From obscurity to prominence in minutes: Political speech and real-time search. 2010.

[MM11] Eni Mustafaraj and Panagiotis Takis Metaxas. What edited retweets reveal about online political discourse. In *Analyzing Microtext*, 2011.

[MM13] P Takis Metaxas and Eni Mustafaraj. The rise and the fall of a citizen reporter. 2013.

[MMFMH12] Eni Mustafaraj, Panagiotis Takis Metaxas, Samantha Finn, and Andrés Monroy-Hernández. Hiding in plain sight: A tale of trust and mistrust inside a community of citizen reporters. In *ICWSM*, 2012.

[MPC10] Marcelo Mendoza, Barbara Poblete, and Carlos Castillo. Twitter under crisis: Can we trust what we rt? In *Proceedings of the first workshop on social media analytics*, pages 71–79. ACM, 2010.

[MSL+12] Richard McCreadie, Ian Soboroff, Jimmy Lin, Craig Macdonald, Iadh Ounis, and Dean McCullough. On building a reusable twitter corpus. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, pages 1113–1114. ACM, 2012.

[Mur12] Dhiraj Murthy. Towards a sociological understanding of social media: Theorizing twitter. *Sociology*, 46(6):1059–1073, 2012.

[NPS10] Meenakshi Nagarajan, Hemant Purohit, and Amit P Sheth. A qualitative examination of topical tweet and retweet practices. In *ICWSM*, 2010.

[OKM+12] John ODonovan, Byungkyu Kang, Greg Meyer, Tobias Hollerer, and Sibel Adalii. Credibility in context: An analysis of feature distributions in twitter. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*, pages 293–301. IEEE, 2012.

[PB12] Thomas Poell and Erik Borra. Twitter, youtube, and flickr as platforms of alternative journalism: The social media account of the 2010 toronto g20 protests. *Journalism*, 13(6):695–713, 2012.

[PGS12] René Pfitzner, Antonios Garas, and Frank Schweitzer. Emotional divergence influences information spreading in twitter. In *ICWSM*, 2012.

[POL10] Sasa Petrovic, Miles Osborne, and Victor Lavrenko. The edinburgh twitter corpus. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media*, pages 25–26, 2010.

[POL11] Sasa Petrovic, Miles Osborne, and Victor Lavrenko. Rt to win! predicting message propagation in twitter. In *ICWSM*, 2011.

[RAZ11] Raquel Recuero, Ricardo Araujo, and Gabriela Zago. How does social capital affect retweets? In *ICWSM*, 2011.

[RBC+11] Tiago Rodrigues, Fabrício Benevenuto, Meeyoung Cha, Krishna Gummadi, and Virgílio Almeida. On word-of-mouth based discovery of the web. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, pages 381–396. ACM, 2011.

[RBK12] Srijith Ravikumar, Raju Balakrishnan, and Subbarao Kambhampati. Ranking tweets considering trust and relevance. In *Proceedings of the Ninth International Workshop on Information Integration on the Web*, page 4. ACM, 2012.

[RGAH11] Daniel M Romero, Wojciech Galuba, Sitaram Asur, and Bernardo A Huberman. Influence and passivity in social media. In *Machine learning and knowledge discovery in databases*, pages 18–33. Springer, 2011.

[Ros12] Evan TR Rosenman. Retweetsbut not just retweets: Quantifying and predicting influence on twitter. 2012.

[SHPC10] Bongwon Suh, Lichan Hong, Peter Pirolli, and Ed H Chi. Want to be retweeted? large scale analytics on factors impacting retweet in twitter network. In *Social Computing (SocialCom), 2010 IEEE Second International Conference on*, pages 177–184. IEEE, 2010.

[SHT+13] Kazutoshi Sasahara, Yoshito Hirata, Masashi Toyoda, Masaru Kitsuregawa, and Kazuyuki Aihara. Quantifying collective attention from tweet stream. *PloS one*, 8(4):e61823, 2013.

[SMH10] Oshani Seneviratne and Andres Monroy-Hernandez. Remix culture on the web: A survey of content reuse on different user-generated content websites. *Web Science (Raleigh, NC*, 2010.

[TAG13] Mozhgan Tavakolifard, Kevin C Almeroth, and Jon Atle Gulla. Does social contact matter?: modelling the hidden web of trust

underlying twitter. In *Proceedings of the 22nd international conference on World Wide Web companion*, pages 981–988. International World Wide Web Conferences Steering Committee, 2013.

[TAH+13] Ke Tao, Fabian Abel, Claudia Hauff, Geert-Jan Houben, and Ujwal Gadiraju. Groundhog day: near-duplicate detection on twitter. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1273–1284. International World Wide Web Conferences Steering Committee, 2013.

[TC12] Ramine Tinati and Leslie Carr. Refluence: A real-time and historic visualization application for twitter conversations. 2012.

[TPT12] Emma Tonkin, Heather D Pfeiffer, and Greg Tourte. Twitter, information sharing and the london riots? *Bulletin of the American Society for Information Science and Technology*, 38(2):49–57, 2012.

[TSS+13] Fujio Toriumi, Takeshi Sakaki, Kosuke Shinoda, Kazuhiro Kazama, Satoshi Kurihara, and Itsuki Noda. Information sharing on twitter during the 2011 catastrophic earthquake. In *Proceedings of the 22nd international conference on World Wide Web companion*, pages 1025–1028. International World Wide Web Conferences Steering Committee, 2013.

[TSSW10] Andranik Tumasjan, Timm Oliver Sprenger, Philipp G Sandner, and Isabell M Welpe. Predicting elections with twitter: What 140 characters reveal about political sentiment. *ICWSM*, 10:178–185, 2010.

[VL10] Diederik Van Liere. How far does a tweet travel?: Information brokers in the twitterverse. In *Proceedings of the International Workshop on Modeling Social Media*, page 6. ACM, 2010.

[WGB13] Ingmar Weber, Venkata R. Kiran Garimella, and Alaa Batayneh. Secular vs. islamist polarization in egypt on twitter. In *The 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining ASONAM 2013*. IEEE/ACM, 2013.

[WLP+12] Claudia Wagner, Vera Liao, Peter Pirolli, Les Nelson, and Markus Strohmaier. It's not in their tweets: Modeling topical expertise of twitter users. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*, pages 91–100. IEEE, 2012.

[WTKM11] Shaomei Wu, Chenhao Tan, Jon M Kleinberg, and Michael W Macy. Does bad news go away faster? In *ICWSM*, 2011.

[WWB+13] Beidou Wang, Can Wang, Jiajun Bu, Chun Chen, Wei Vivian Zhang, Deng Cai, and Xiaofei He. Whom to mention: expand the diffusion of tweets by@ recommendation on micro-blogging systems. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1331–1340. International World Wide Web Conferences Steering Committee, 2013.

[YL11] Jaewon Yang and Jure Leskovec. Patterns of temporal variation in online media. In *Proceedings of the fourth ACM international conference on Web search and data mining*, pages 177–186. ACM, 2011.

[YW10] Shaozhi Ye and S Felix Wu. Measuring message propagation and social influence on twitter. com. In *Social informatics*, pages 216–231. Springer, 2010.

[Zar09a] D Zarrella. State of the twittersphere. *Retrieved December*, 15:2009, 2009.

[Zar09b] Dan Zarrella. Science of retweets. *Retrieved December*, 15, 2009.

[ZJW+11] Wayne Xin Zhao, Jing Jiang, Jianshu Weng, Jing He, Ee-Peng Lim, Hongfei Yan, and Xiaoming Li. Comparing twitter and traditional media using topic models. In *Advances in Information Retrieval*, pages 338–349. Springer, 2011.