

Strategies and Motivations behind Artificial Trending Topics in Twitter

ABSTRACT

This paper presents several case studies on how groups cooperate to manipulate Twitter's Trending Topics to their benefit, creating what we called Artificial Trending Topics. We show evidences that such topics are quite common and discuss some of the strategies used by these groups to reach their goals and their motivations to do so. We further discuss these different types of topics in the context of social capital. Finally, we show how the social network behind artificial topics are quantitatively different from networks behind organic topics, paving a way for the automatic classification of each type.

1. INTRODUCTION

Twitter is currently a popular microblogging service used for many different communication purposes (Huberman et al., 2009). One of the services offered by Twitter is the Trending Topics list, composed of a list of terms (e.g. keywords, hashtags, short phrases) automatically extracted from the real-time stream of public messages exchanged within the service and that are seen as *trending*. While Twitter does not disclose the exact algorithm used to define what is a trending topic, the main idea behind the service is to identify "topics that are immediately popular, rather than topics that have been popular for a while or on a daily basis, to help people discover the hottest emerging topics of discussion on Twitter" (<https://support.twitter.com/articles/101125> as seen on May 22nd 2012).

Unlike the user's stream where only information shared by those followed can reach the user, trending topics allows information outside the users' social network to reach them. This feature also appears in other social network sites, such as Google+ and Facebook. The increased reach of Trending Topics makes the space very valuable, as information listed there can be potentially read by millions of users. Such high value creates an incentive for users to try and manipulate the list. In this paper, we argue that this manipulation gives rise to *Artificial Trending Topics*, created by a coordinated group with the specific purpose of making a message appear as a trending topic, as opposed to Organic Trending Topics which are the result of the uncoordinated interactions of many users. This manipulation of the service reduces its utility, as these artificial topics displace attention with organic topics, which can be more useful to a wider audience.

We provide a definition of an Artificial Trending Topics and explore some of their properties and implications. By the analysis of a set of topics collected from Brazil's Trending Topics list over a period of time, we analyze several cases and report on motivations and strategies used by groups to artificially inflate desired terms. Furthermore, we show that there are quantitative differences in the social networks behind Artificial and Organic trending topic, which may be used to automatically detect artificial (and organic) topics.

Our report contributes to the understanding of how users organize and cooperate on Twitter. Several studies focus on Trending Topics. Cheong (2009) monitored Trending Topics in order to understand "Twitter's ecosystem". Becker et al. (2011) also focused on them as a way to identify events and non-events in Twitter, trying to understand how they reflect information movements. Zubiaga et al. (2011) proposed a taxonomy for the popular trending topics. However, as far as we know, there are no previous work analyzing the manipulation of Trending Topics and our work provides the first insights into this phenomena.

2. ARTIFICIAL TRENDING TOPICS

The concept of an Artificial Trending Topic was created when, by evaluating Brazilian's Trending Topics in early 2012, we observed different behaviors regarding how a topic trended. For most topics, when reading tweets containing one could see that content was indeed about the topic; moreover tweets were coming from many different users, from seemingly different groups. There didn't seem to be anything special about how the topic first appeared - it was often a casual use of the topic (often a hashtag) by some user that caught on. We considered these topic as trending organically, since it matched how Twitter defines a trending topic.

However, some topics were different. For these, the first occurrence of the topic was already in a context of incentivizing other users to use the topic with the specific purpose of making it trend. As we will describe in more details later, this often was a user explicitly asking his or her followers to use the topic and help him or her make it trend. Moreover, the topic's format was different from organic ones - they were not used to give context to a tweet, they contained a single and autonomous piece of information (i.e. they made sense even when stripped from the tweets). We called these topics Artificial Trending Topics, since they were artificially created with the specific purpose of making it trend and artificial boosting methods were applied to make it do so.

We define an Artificial Trending Topic as one that trends (i.e. appear in Twitter's Trending Topics list) after purposeful and coordinated actions of users specifically aiming at making the topic trend. Therefore, it is the motivation behind actions that classify a topic as artificial. In contrast, an *Organic Trending Topic* is one that trends without any explicit coordination and without an initial goal of making the topic trend. That is, an organic topic trends as a collateral effect of the topic being talked about, while an artificial one only trends because a group of users take actions for it to trend.

3. GOALS AND METHODOLOGY

In this paper we aim at showing characteristics of Artificial Trending Topics, focusing on the strategies users apply to try and make a topic "trend" and their motivations to do so. It is our goal to provide insights on how common artificial topics are becoming and the reasons and methods by which they are able to exist.

To classify a topic as organic or artificial, one must find where it started and evaluate the motivations behind it appearing. We consider a topic an Artificial Topic if (1) one or more users explicitly ask others to use a particular tag and (2) this tag wasn't being used before.

In order to so, over 5 days in January 2012 we collected at regular intervals the Trending Topics appearing in Brazil's Trending Topics list. In total, we collected 460 topics. For each topic, we performed a search and collected all tweets containing the topic, qualitatively analyzing their content to extract their causes and reasons for growth.

We classified each collected topic as Organic or Artificial, according to whether there were, at some point, coordinated efforts to make the topic trend. For each type, we randomly selected 20 topics and further collected, using NodeXL, the social networks behind them (users and follower network). The majority of tweets collected were primarily written in Portuguese and only a small part were written in English (e.g. topics that were tributes from Brazilian fans to an American artist). We will, however, present all data translated to English in this paper.

4. RESULTS

4.1 Artificial vs Organic Topics

Of the 460 unique topics collected, 276 (60%) were classified as artificial and 184 as organic. This is evidence of how common artificial topics are in Brazil’s Trending Topics.

An analysis of the purpose of these artificial topics showed that they can be further classified in four basic categories, as shown in Table 1. In this table, *Personalities* are topics related to famous people, mostly singers and actors, created by fans; *Protests* are mostly topics showing dissatisfaction with a government or organization action; *Promotions* are topics created by business companies to promote some product. We can see that the great majority of artificial topics are created to state a message about some idol or personality. A minority were created by companies trying to promote some product or users protesting about something (companies or the government).

The organic topics collected could be classified as shown in Table 2. Events are topics about something that is happening or has recently happened, such as an earthquake. Memes are typically jokes. Personalities are topics related to famous people, mostly singers and actors. We can see that organic topics related to personalities or idols are much less common than artificial ones.

Table 1: Artificial Topics by category.

	Number	Percentage
Personalities	242	87.6%
Protests	12	4.3%
Promotions	12	4.3%
Other	10	1.1%

Table 2: Organic Topics by category.

	Number	Percentage
Events	70	38%
Memes	70	38%
Personalities	44	24%

4.2 Structure of Artificial Topics

The majority of Artificial Topics are composed of tags (a word or sequence of words without spaces), that can contain the hash symbol or not but, unlike in Organic Topics, each tag typically contains a full statement, sometimes even a whole phrase.

One of the utilities of trending topics is to provide a way to quickly find users talking about that topic and join the conversation; in this sense, organic topics only hint at what they are talking about, as they try to be concise in order to provide space for an opinion or discussion. However, users behind Artificial Topics are mostly interested in spreading a message in the Trending Topics list, not in being found by other users. Therefore, the topic must contain a complete statement.

For example, in organic topics we often observed the tag “#restart”, which is the name of a popular teenage band in Brazil. In artificial topics, tags were more like “#restartisthebest” or “#brazilmissesluansantana” (Luan Santana is a popular singer). Most artificial topics do not use spaces. We believe this is because otherwise users would be tempted to modify or add to the statement, reducing its effectiveness to reach the Trending Topics list. A statement without spaces clearly shows that it is a single conceptual unit that should not be changed. Nonetheless, a few topics were seen containing spaces (e.g. “luan santana my smile”).

This use induce larger tags in artificial topics. In our sample, on average artificial topics contained tags 13.1 characters long and the largest observed tag was 32 characters long (but, as in the above example, multiple large tags can be used in the same tweet). On the other hand, tags in organic topics were on average 8.9 characters long and the largest tag was only 15 characters long.

4.3 Features

On average, 446 unique users were behind each artificial topic, with an average of 12 tweets per user. These numbers are strikingly different from organic topics, where on average 647 unique users per topic were observed and only 4 tweets per user. Organic topics result from many different users sparsely tweeting about them, whereas artificial topics result from fewer users tweeting a lot about them.

The social networks behind each type of topic are also quite different. Table 3 shows several measurements for the collected networks. We can see that networks behind artificial topics are denser, more clustered and have users that are more connected. These numbers are evidences of a strong community behind artificial topics. Figures 1 and 2 illustrates these features by showing prototypical examples of the networks behind each topic.

Table 3: Average values for different attributes of networks behind organic and artificial topics. Each value is the mean of 20 topics.

	Artificial Topics	Organic Topics
Density	9.35	2.21
Clustering Coefficient	0.22	0.08
Closeness	0.07	0.18

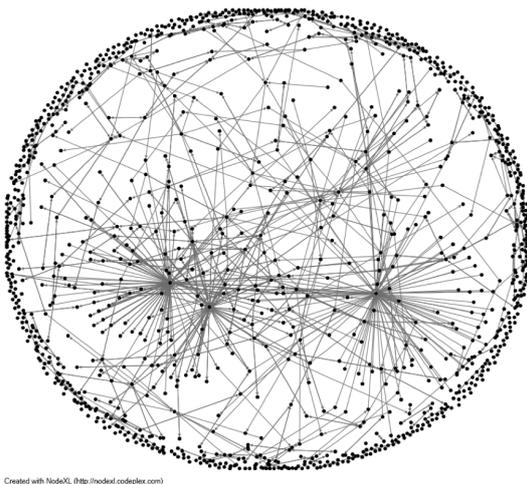


Figure 1: An example of a social network behind an organic topic. Nodes are users that tweeted using the topic tag and edges are following relations.

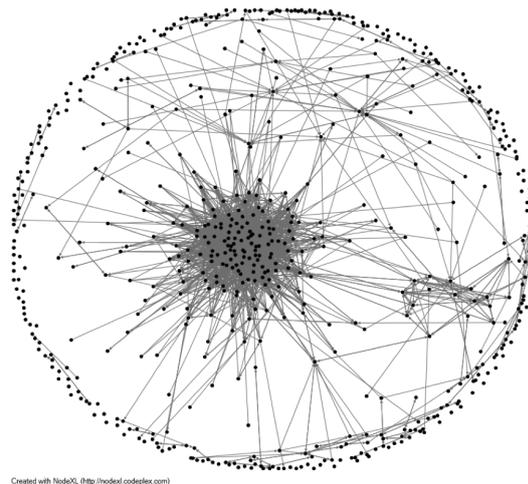


Figure 2: An example of a social network behind an artificial topic. Nodes are users that tweeted using the topic tag and edges are following relations.

4.4 Motivations

The main reason to create an artificial topic is to make a statement visible to a broader audience, beyond the small group where the statement originated.

These groups seem to be motivated by the desire to show strength and cohesion, and the ability to make a topic trend is seen as a proxy for these attributes. Topics protesting or supporting something all start with some

users urging others to help the topic trend so that they can show everyone that they are unhappy or happy about such thing.

This is even more evident in topics about personalities, started by fandoms (which are communities of fans (Monteiro, 2011)). As shown in the previous section, it is comparatively uncommon for topics about personalities or idols to trend organically. Even though some idols enjoy a large fan base, their fans' tweets only coordinate when some event is about to happen or has happened (e.g. an upcoming big concert); otherwise, they typically tweet at different times about different specific subjects, remaining under Twitter's radar.

By coordinating efforts, fandoms are able to "show the world" their support and, at the same time, to show the group's ability to organize and seem larger than it really is. Moreover, it is not uncommon to have competitions among different fandoms to see who can make their topic of choice trend higher in the list. Often, groups will try and trend messages bashing the other group or the other group's idol. To a large extent, the Trending Topics list is the main medium through which these groups communicate - it is their battleground.

Artificial Topics arise from the cooperation and engagement of a small group of users, clearly a product of bonding ties. Organic Topics, on the other hand, seem to rely mostly on bridging ties, as they rise from the interactions of different groups and otherwise unrelated users.

4.5 Strategies

Groups promoting artificial topics use several different strategies to artificially inflate the topic and avoid Twitter's spam detection algorithms. In what follows, we discuss the strategies we observed in our sample. All observed artificial topics were the result of more than one strategy being applied at some point - i.e. it is not the case that each artificial topic is the result of a single strategy. This list is not meant to be exhaustive but rather a glimpse at what was common at that particular moment; it is likely that strategies are modified and replaced to adapt to Twitter's changes in policy and algorithms.

A successful artificial topic starts with a single user or a small group of users (seeds) asking their followers to use a specific tag in their tweets. There does not seem to be deliberation around which tag is to be used, but it could be the case that many different tags are proposed by different sets of users and only those that resonate among a community end up being used and reaching the trending topics list. When asking to use the tag, these seed users explicitly mention that they want to make the topic trend and appear as a Trending Topic. No further explanation is usually given, an evidence that their target audience is, for the most part, already aware of what they are trying to accomplish and the strategies to do so.

These seeds also typically specify when to start the operation, so that users can prepare in advance. They will specify a day and a time, so that all efforts are concentrated at some point in time, in what seems to be a necessary condition for a topic to be considered as trending. When we examine qualitatively the tweets, users activate the network asking their connections to help by using words like "Let's go Troops! Let's show our strength" or "Come on, family, everybody using the tag!". This sort of wording shows that there is a community behind the topic and how showing the group's strength is an important motivation.

At this point, a group has negotiated and synchronized on a common tag to use when talking about the topic. Users then proceed to tweet using the tag in all of their tweets regarding the topic, starting at the specified time. A great care is taken to never post the tag by itself or sending the same exact tweet more than once - doing so seems to trigger Twitter's spam detection algorithm, compromising the operation. Some users often remember other users on the risks of taking these actions.

Users may also increase the frequency with which they will tweet about the topic, however often users have other audiences that may not care about the topic they are trying to trend. Therefore, there seems to be a limit on how often they will tweet about the topic. As a consequence, only synchronizing on a tag hardly makes up the volume necessary to make the topic trend. To improve the volume, many users rely on another strategy, which we call Tag Hitchhiking.

Tag Hitchhiking is the most common strategy to artificially inflate a topic, appearing in all collected cases. Users use the specified tag in almost all of their tweets, even (and foremost) in posts that have nothing to do with the tag. This strategy allows for the topic to seem organic, as the tweets are still organic in nature (i.e. they would probably have been tweeted even without the tag), without annoying too much other audiences.

Users also reinforce and incentivize other users to use the tag, by directly mentioning other users in the tweet. The majority of mentioned users are followers of the users mentioning them. Therefore, the targets are users that are known to be part of the group but hasn't tweeted the tag yet (by mentioning them, the target users are more likely to see the message). Mentioning users outside the group would probably be ineffective and likely to be considered spam by Twitter.

5. DISCUSSION

These cases, even if from a somewhat small sample, are evidence that there are indeed two types of trending topics and that the networks behind these topics influence how they trend. These networks seem to use different values embedded in their structure to allow the propagation of the topics they want.

Artificial topics, therefore, are a form of collective action, where the social network cooperates to reach benefits and resources that are harder to achieve without this cooperation. Artificial topics seem to rely on bonding ties, because these networks activate fewer nodes, requiring more engagement and cooperation to achieve their goals. Through the social structure, these users create and share the values they need (Bourdieu, 1983; Putnam, 2000) and bonding ties (Granovetter, 1973; Putnam, 2000, 2007) are the ties that provide more engagement for the actors in the social structure. These networks are based on mutual identification with some cause, idol or idea. People work together because they share the same values and the same vision. Homophily is a characteristic associated with bonding rather than bridging ties (Burt, 2000, 1992; Lin, 2001).

The increased difficulty in leveraging one's social network to reach the Trending Topics list create the opportunity for different groups to work together. In one case that appeared in our sample was the network created by fans of two Brazilian artists, *Luan Santana* and *Restart*. Fandoms of each artist cooperated to put a chosen tag that made a tribute for both artists in the Trending Topics list. 3 shows that there are two different groups acting, each highly clustered but otherwise with fewer connections among them. Upon closer analysis, each cluster can be associated to one fandom and both groups were actively cooperating to promote a tag of common interest.

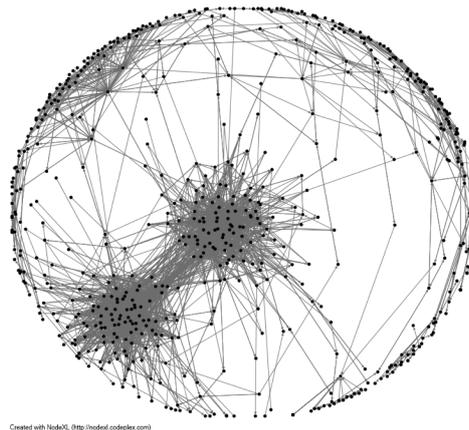


Figure 3: Social Network behind the topic “#restartluansantana”. Nodes are Twitter users, edges are drawn when the two users follow each other.

Interestingly, an artificial trending topic may also become organic. One example was the tag “CompreiMagazineLuiza-NaoEntregou” (I bought and Magazine Luiza - a store - didn’t delivered). This topic was created by a group of Luan Santana fans that became enraged that one of them had bought a refrigerator from the on line store and didn’t receive it. These users used their network, and their experience in manipulating Trending Topics, to share the story, cooperating to put the tag in the Trending Topic list. As it happens, other unrelated users, seeing the tag trending and having had problems with the same store, started to use the tag, which ended up being used organically. Although this topic was artificially created, it became organic because other users felt empathy or identification with the cause.

Influencing Trending Topics can have an impact beyond Twitter. Several newspapers in Brazil comment on topics that are trending in social networks on a daily basis, increasing the reach of a message beyond Twitter and the internet. In one observed case, a group of skaters was able to trend the tag “#QueremosUmaPistaDeSkate-EmCanoasPrefeito”, which demands a skate park in the city of Canoas, was answered personally by the Mayor’s official twitter account, promising to build a new skate park.

Two other topics also made it to the list “QueremosVodkaGratis-EmCanoas” (We Want Free Vodka in Canoas) and “CanoasMafiadTag” (something similar to “Canoas Tag Mafia”). The first topic was a joke about how easily the group got the attention of the Mayor and the second one, a statement about the group’s ability to trend topics. Both topics are about the value of making a topic trend and the value of the group itself. This shows how much users value to have influence and the ability of creating movements in Twitter.

Organic topics, on the other hand, seem to travel more bridging ties [23][24]. Rather than being represented by an organized movement to reach visibility, they emerge from discussions and conversations around the topic. They don’t need users to cooperate to bring them to the TTs. Thus, they use more bridging ties to reach more groups, in a more organic way. Their value is connected to the novelty and the hotness of the topic rather than a statement made by a group (similarly to what Wu & Huberman [29][30] demonstrated).

These topics are more naturally related to the essence of how Twitter perceives Trending Topics, as argued in the beginning of this paper. It is based on the perception of values the information creates for the group if tweeted rather than the group creating value for the information. Thus, information here is the type of social capital [18] [8], rather than cooperation and engagement.

While artificial trending topics are the result of a group strategy to obtain resources in a war for visibility and attention (Lanhan, 2006), organic trending topics are not. However, our study shows evidences that artificially created topics are, quite often, taking over organic topics in the Trending Topics list.

6. CONCLUSIONS

In (Recuero and Araujo, 2012) it was argued that there are two different types of Trending Topics in Twitter, one that emerges organically, from the spontaneous interactions of several users, and another where growth is due to the coordinated cooperation of a smaller group of users.

By analyzing a sample of trending topics, we collected more evidences of such division and further showed that each type of topic has strikingly different characteristics. Particularly, artificial topics use longer tags that represent complete statements and have stronger communities behind them. Furthermore, we reported on the motivations behind artificial topics, showing that the great majority of artificial topics seem to be organized by fandoms, in order to send “global” messages about their idols.

Finally, we reported on several strategies used by these groups to achieve their goal of artificially trend a topic, relating these strategies and motivations to forms of social capital. We argued that artificial topics are the product of bonding ties, while organic topics are the product of bridging ties.

The differences behind the two types of topics seem separable enough to allow for their automatic classification, potentially by feeding the many signals (network properties, tag features) to a machine learning algorithm. This could allow for a better management of the Trending Topics list, as artificial topics are becoming quite

common and disputing attention with organic topics, which are arguably more useful to a wider audience.

This work contributes to the understanding of how cooperation arises in Twitter to manipulate the service and the ways groups are able to create artificial topics. Future lines of research include understanding hierarchies in groups promoting trending topics and how the negotiation of a particular tag takes place. We also plan on better quantifying the observations, by looking at a larger and more detailed sample. Finally, we'd be interested in building automatic classification systems based on machine learning algorithms.

References

- Becker, H., Naaman, M. and Gravano, L. (2011), Beyond trending topics: Real-world event identification on twitter, in 'Proceedings of ICWSM 2011'.
- Bourdieu, P. (1983), The Forms of Capital, In: Handbook of theory and Research for Sociology of Education, edited by J. G. Richardson, Westport, CT.
- Burt, R. (2000), The network structure of social capital, in 'Research in Organizational Behavior', Vol. 22, JAI Press.
- Burt, R. S. (1992), Structural Holes: The Social Structure of Competition, Cambridge, MA: Harvard University Press.
- Cheong, M. (2009), "'what are you Tweeting about?': A survey of Trending Topics within the Twitter community", Technical report 2009/251. Clayton School of Information Technology, Monash University, Vol. 12.
- Granovetter, M. (1973), "The strength of weak ties", The American Journal of Sociology 78, Vol. 6, pp. 1360–1380.
- Huberman, B., Romero, D. and Wu, F. (2009), "Social Networks that matter: Twitter under the microscope", First Monday, Vol. 1.
- Lanhan, R. (2006), The Economics of Attention, Chicago: University Of Chicago Press.
- Lin, N. (2001), Social Capital, A Theory of Social Structure and Action, Cambridge.
- Monteiro, C. (2011), Familia #happyrock: recomeco em cores, in 'I Simposio em Tecnologias Digitais e Sociabilidade - Salvador, BA, Brazil'.
URL: <http://migre.me/7dt0V>
- Putnam, R. D. (2000), Bowling Alone, New York: Simon & Schuster.
- Putnam, R. D. (2007), "E pluribus unum: Diversity and community in the twenty-first century the 2006 johan skytte prize lecture", Scandinavian Political Studies, Vol. 30, pp. 137–174.
- Recuero, R. and Araujo, R. (2012), On the rise of artificial trending topics, in 'Proceedings of ACM Hypertext 2012'.
- Zubiaga, A., Fresno, V., Spina, D. and Martnez, R. (2011), Classifying trending topics: A typology of conversation triggers on twitter, in 'Proceedings of the CIKM11'.