Are Trending Topics Useful for Marketing?

Visibility of Trending Topics vs Traditional Advertisement

Juan Miguel Carrascosa
Universidad Carlos III de Madrid
jcarrasc@it.uc3m.es

Roberto González
Universidad Carlos III de Madrid
rgonza1@it.uc3m.es

Rubén Cuevas
Universidad Carlos III de Madrid
rcuevas@it.uc3m.es

Arturo Azcorra
Universidad Carlos III de Madrid
Institute IMDEA Networks
azcorra@it.uc3m.es

ABSTRACT

Trending Topics seem to be a powerful tool to be used in marketing and advertisement contexts, however there is not any rigorous analysis that demonstrates this. In this paper we present a first effort in this direction. We use a dataset including more than 110K Trending Topics from 35 countries collected over a period of 3 months as basis to characterize the visibility offered by Local Trending Topics. Furthermore, by using metrics that rely on the exposure time of Trending Topics and the penetration of Twitter, we compare the visibility provided by Trending Topics and traditional advertisement channels such as newspapers' ads or radio-stations' commercials for several countries. Our study confirms that Trending Topics offer a comparable visibility to the aforementioned traditional advertisement channels in those countries where we have conducted our comparison study. Then, we conclude that Trending Topics can be useful in marketing and advertisement contexts at least in the analyzed countries.

Categories and Subject Descriptors

[**Information Systems**]: *Social Advertising*; [**Networks**]: *Online Social Networks*; [**Human-Centered Computing**]: *Social Media*.

Keywords

Trending Topics; Twitter; Marketing; Visibility

1. INTRODUCTION

Online Social Networks (OSNs) in general and Twitter in particular have changed the way in which people communicate, but also have a significant impact on the public image of celebrities or politicians and are being used by important companies with marketing and/or advertisement purposes [25]. In particular, Twitter has its own business web page [9] and marketing on Twitter has become a business itself [20, 21]. Twitter offers a functionality, that among other uses, is of high relevance in this context named *Trending Topics* (TTs) which are officially described as: "*the hottest emerging topics (or the "most breaking" breaking news), rather than the most popular ones*" [19]. As acknowledged by experts in the field of marketing, surprise is one of the most powerful marketing tools [6]. TTs hold by definition this surprise component and marketing experts have been exploiting it. For instance, TV and radio-station shows have started to announce *hashtags*[1] so that all tweets regarding the show can be aggregated using a hashtag which eventually may become Trending Topic. If that happens it is reported as a big success. Trending Topics have been also used with marketing purposes in politics. For instance, in the last public debate for the Spanish presidency in 2011, one of the candidates became TT as a result of an orchestrated operation by his party supporters. This was used as an unequivocal proof by his party and by several media that he had won the debate [24]. In addition, some social movements such as the "occupy" movements augmented their visibility among the population after becoming TT [14]. Furthermore, the commercial interest of Trending Topics for companies is reflected by the *Promoted Trending Topics* service offered by Twitter [12]. These are a special type of TTs that can be purchased in slots of 24 hours for around $200K [22]. This service is regularly used by companies in the context of advertisement and marketing campaigns.

Finally, another symptom of the relevance of TTs is the recent movement made by Facebook to implement its own Trending Topics service that is currently available for users in United States [4].

However, to the best of the authors knowledge, this (seemingly) common idea that TTs are a useful tool in marketing contexts is not supported by any scientific or technical work. We believe that a solid scientific basis is required to allow experts in different disciplines to make informed decisions regarding the actual impact that TTs may have in market-

[1]A hashtag is a special type of word that starts by the symbol #. It is a common practice that people tweeting about a common topic use a common hashtag to identify it.

COSN'13, October 7–8, 2013, Boston, Massachusetts, USA.
ACM 978-1-4503-2084-9/13/10...$15.00.
http://dx.doi.org/10.1145/2512938.2512948.

ing, advertisement, and related contexts. This paper constitutes a first effort in that direction in which we perform a thorough analysis of the actual *visibility* provided by TTs.

In particular, we study the visibility of World Wide Trending Topics (WW-TT), but more interestingly from the point of view of marketing[2] we analyze the visibility provided by the Local Trending Topics (Local-TT) from 35 different countries. Toward this end, we first define and implement a high resolution measurement methodology that leverages the Twitter API to collect the list of TTs with a resolution of dozens of seconds. Using this methodology we have collected 3 WW-TTs datasets between Sep 2011 and May 2013 that all together include more than 80K TTs. Using these datasets we demonstrate that the resolution provided by our methodology enables the detection of any change in the visibility of TTs. Identifying these changes is of high importance in the aforementioned marketing or advertisement contexts.
Furthermore, we use the same methodology to collect a dataset including more than 110K Local-TTs from 35 countries over a period of 3 months in 2013. We use this dataset to compare the visibility offered by TTs across these countries. In order to perform a complete comparison we define three metrics. The first one helps us to compare the *net-visibility* (i.e., the actual time of exposure) of TTs whereas the other two metrics named *potential-visibility* and *potential-online-visibility* take into account the penetration of Twitter among the population and the population with Internet access in a country, respectively. These metrics give an insight on the fraction of the population (or "online population") that the Local TTs are able to reach in a country. In addition, we use the aforementioned metrics to compare the visibility offered by TTs and traditional advertisement channels such as newspapers' ads and radio-stations' commercials for several countries with rather different demographics and cultural backgrounds. Finally, we analyze the variability offered by TTs visibility within a country and, for 3 selected countries (Ireland, New Zealand and UK) we present a more detailed analysis of the visibility: (*i*) using a novel and efficient methodology we classify the TTs of a country in different semantic categories and study which categories are more likely to become TT and which ones offer higher visibility periods; (*ii*) we study whether TTs visibility follows a diurnal pattern as Internet traffic [40] and many other online services do.

In summary, the main contributions of this paper are twofold: First, a measurement methodology that allows to monitor the visibility of TTs and its evolution over time. Second, a methodology to properly characterize the visibility of TTs within a country that permits to perform meaningful comparative analyses with other countries or with traditional advertisement channels. The utilization of these methodologies led to the following insights:
- Our results show that the median visibility of TTs is higher than that offered by radio-stations' commercials and newspapers' ads in 4 and 9 out of 10 studied countries, respectively. Hence, we conclude that (at least for the studied countries) TTs can be considered a useful tool in marketing and advertisement contexts.
- However, there is a strong variability on the visibility that TTs offer in different countries and also across Trending Topics within a country. In addition, the penetration of traditional media and TTs varies substantially accross countries. Therefore, we cannot generalize the previous conclusion for all the TTs in every country.
- Our detailed examination of few countries reveals that "Hashtags" present a higher visibility than other "non-hashtaged" TTs related to "Sport Events" or "Celebrities". Furthermore, the exposure time of TTs presents a clear diurnal pattern for most of the studied countries. Specifically, TTs provide longer visibility periods during night hours when fewer users are connected.

The rest of the paper is organized as follows: Section 2 describes our measurement methodology and our datasets. Section 3 details our methodology to evaluate the visibility of TTs within a country while we devote Section 4 to put in context our analysis doing a comparison with traditional advertisement channels. Section 5 dissects the visibility of TTs within a country from a semantic perspective. Finally, we summarize the related work in Section 6 and Section 7 concludes the paper.

2. MEASUREMENT METHODOLOGY, METRICS AND DATASETS

In this section we describe our large scale measurement methodology to collect information for thousands of Trending Topics over a period of several months. Additionally, we define temporal metrics to be used in the rest of the paper. We also discuss the basic filtering techniques applied to produce meaningful datasets and finally we summarize the datasets used to conduct our analysis.

2.1 Measurement Methodology

Twitter provides different APIs to access the information available in the system [17]. In our methodology we leverage two of these APIs, namely the REST and Streaming APIs. We query the REST API to obtain the list of 10 TTs at a given instant and for a given location (e.g., a country). Since the maximum number of queries allowed by Twitter to the REST API is 150 per hour, we are able to collect the list of TTs every 24 seconds for a given location. This guarantees a fine grain time resolution in the sampling of the Trending Topics list. Furthermore, we query the Streaming API to retrieve the tweets associated to a given Trending Topic. The Streaming API offers a best effort service in which the system provides as many tweets as it can (depending on the load) including the term (i.e., Trending Topic) requested in the query. In particular, our tool uses the Streaming API to collect tweets associated to the 20 most recent TTs at any moment.
Using multiple instances of our tool we are able to collect data from World Wide (WW) Trending Topics as well as Local Trending Topics from 35 different countries in parallel.

2.2 Temporal Metrics

The visibility of a TT is basically defined by the time that it is shown to users that we refer to as *exposure time*. We use the following meaningful metrics to capture the temporal characteristics of TTs:

- *number of active periods*, this metric counts the number of times that a given topic has become TT. We refer to each one of those active periods as an *instance*.

[2]Marketing experts are interested on studying different regional markets.

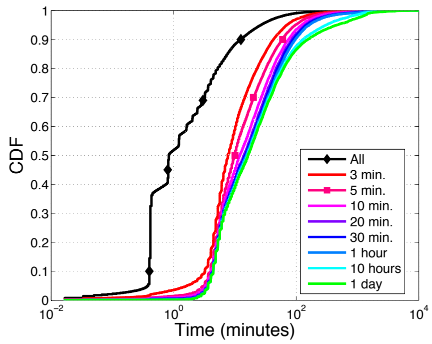


Figure 1: *Active time* of WW-TT instances without oscillations filtering (*All*) and with oscilations filtering (filtering times from 3 min to 1 day).

- *total active time*, this metric captures the total time a topic has been TT across one or multiple active periods, i.e., the total exposure time.
- *age*, this metric measures the total time between the first instant and the last instant a topic is a TT across one or multiple active periods.

To clarify these concepts, let us consider the following simple example: a topic that has been Trending Topic on Jan 1^{st} 2013 between 9 AM and 9:30 AM, on Jan 1^{st} between 6 PM and 6:20 PM and on Jan 2^{nd} between 8:50 AM and 9 AM. Then, the *number of active periods* for this Trending Topic is 3 (or in other words this TT has 3 associated instances), the *age* is 24 hours (from 9 AM Jan 1^{st} to 9 AM Jan 2^{nd}) and the *total active time* is 60 min (30, 20 and 10 minutes in the first, second and third active periods, respectively).

Previous studies have considered the volume of tweets [26, 35] to analyze TTs using the Search or the Streaming API. Although, this metric does not capture the visibility of TTs as well as those presented above, it could be an interesting complementary metric for our study. Unfortunately, as demonstrated by Morstatter et al. [38], the volume of tweets obtained from the Streaming API is not a reliable metric. In particular, that study shows that due to the best effort nature of the Streaming API in those peak hours where the number of tweets associated to a topic is higher the API provides the lower number of tweets[3]. In short, using the volume of tweets as a metric may lead to wrong results and thus we do not use it for our analysis.

2.3 Data Filtering

As described before, our methodology allows to gather the list of the 10 TTs for a given location (e.g., WW or Local TTs for a country) every 24 seconds. Unexpectedly, there is a high variability in the composition of this list in a time scale of few minutes (or even seconds). We conjecture that this high variability is due to those topics that are ranked by Twitter Trending Topics selection algorithm around the 10^{th} position that enter and leave the Top 10 list frequently. The curve labeled as *"All"* in Figure 1 shows the distribution of the *active time* for each WW-TT instance in our dataset.

[3]Note that this observation also applies to the search API since it provides a subset of the tweets provided by the Streaming API [23].

	Period	TT Instances	Unique TTs
WW-TT-2011	09/07/2011 - 11/30/2011	31251	13964
WW-TT-2012	12/01/2011 - 02/25/2012	80856	43985
WW-TT-2013	02/20/2013 - 05/20/2013	67221	29326
Local-TT-2013	02/20/2013 - 05/20/2013	713012	112196

Table 1: Basic statistics of Datasets.

We observe that half of the instances present an *active time* lower than 1 minute. Therefore, the Trending Topics selection algorithm works in intervals of seconds. Note that previous works considered that the list of TTs was updated in intervals of 5 minutes [35] or 20 minutes [26].

This real time selection of TTs produces a phenomenon that we refer to as *oscillations*. This occurs when a topic enters and leaves the Trending Topic list several times in a short period of time (e.g., a few minutes). However, *oscillations* are unlikely to be observed by users since neither the web interface of Twitter nor Twitter API-based applications refresh the Trending Topic information as frequently as our measurement tool. Therefore, in order to better approximate the user experience we would like to process the collected data in order to filter these short-term *oscillations*. For this purpose, we consider that a topic that presents one or more *oscillations* within a period of X minutes has been a Trending Topic during the whole X minutes period. Figure 1 shows the CDF of the *active time* of single instances of TTs after applying the described technique for X = 3, 5, 10, 20, 30, 60, 600 and 1440 minutes. The result suggests that a value of X = 5 min suffices to eliminate most of the short-term *oscillations* (i.e., those in the order of seconds or few minutes) and do not merge those long-term *oscillations* (i.e., those in the order of tens of minutes). Therefore, we filter out the *oscillations* using this value. We have repeated the experiments described along the paper with other values of X (3 and 7 minutes) obtaining similar results.

2.4 Datasets

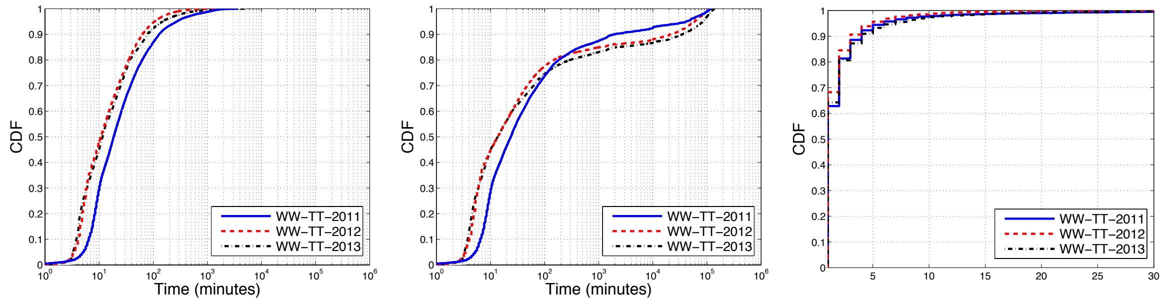
Using the measurement methodology and data filtering technique described in this section we collected the following datasets:
WW-TT: This dataset is formed by 3 traces including all the WW-TTs in 3 different periods of approximately 3 months each.
Local-TT: This dataset was collected in parallel to our most recent WW-TT trace. It includes the Local Trending Topics for 35 countries over a period of 3 months.

The specific dates of data collection along with the number of TTs included in each trace are shown in Table 1.

2.5 Accuracy of the measurement methodology

The final goal of our measurement methodology is to accurately collect the visibility offered by TTs at any moment, expressed through the previously defined temporal metrics. Hence, the proposed methodology should be able to discover any change on the visibility of TTs.

Figure 2 presents the distribution of the number of active periods, total active time and age across TTs within our three Worldwide datasets. We observe that TTs within WW-TT-2012 and WW-TT-2013 show a similar visibility that is significantly different from that shown by TTs within WW-TT-2011. In particular, Figure 2(a) reveals that the



(a) Total active time (b) Age (c) Number of active periods

Figure 2: CDF of the temporal metrics of TTs within our WW-TT datasets.

median value for the total active time halves, from 20 to 10 minutes, between WW-TT-2011 and WW-TT-2012 and then remains stable in WW-TT-2013. This result suggests that the TTs selection algorithm was modified to severely reduce the visibility of TTs in December 2011, most likely during the large system upgrade process carried out by Twitter on that month [18]. However, to the best of the authors knowledge, this modification on the Trending Topics selection mechanism was not publicly announced by Twitter despite the implications that it might have.

In order to corroborate the previous observation, we have calculated the distribution of the total active time for each individual month in our Worldwide datasets but December 2011 (for being the month where the modification took place) and performed a Kolmogorov-Smirnov test [37] for each pair of distributions. The obtained results show that the distributions of Sep'11, Oct'11 and Nov'11 are similar between them and so are the distributions of Jan'12, Feb'12 and those from 2013. Specifically, the parameter K of the test varies between 0.06 and 0.15 in all cases. However, when we compare any of the first three months to any of the other months the Kolmogorov-Smirnov test concludes that the distributions are significantly different, in particular, K varies between 0.27 and 0.32.

Moreover, Figure 2(b) shows the distribution of TTs age for our three WW-TT datasets. Again, we observe that the distribution for this metric is similar for WW-TT-2012 and WW-TT-2013 and different from WW-TT-2011. This confirms the reported change in TTs visibility. In particular, the modification in the TTs selection algorithm in Dec 2011 yielded around 80% of TTs (i.e., those that have one or two close active periods) to present a lower age in our WW-TT-2012 and WW-TT-2013 than in the WW-TT-2011 dataset. However, this trend is reversed for the 20% TTs presenting a longer Age (i.e., those with several associated active periods). This suggests that the TTs selection algorithm implemented since Dec 2011, in addition to shorten the active time of TTs instances, also requires that the period of time with a relative reduced volume of tweets for a topic to become TT again to be longer. It is worth to mention that we have performed equivalent Kolmogorov-Smirnov tests for this metric as for the total active time obtaining similar results.

Finally, Figure 2(c) shows the distribution of the number of active periods (or instances) for our WW-TT datasets. The results indicate that this distribution is similar for the three datasets. Then, the modification of TTs selection algorithm in Dec 2011 has not affected the ability of TTs to achieve this status multiple times, however as noted before the time between TTs instances has increased.

In summary, the results presented in this subsection confirm that the proposed measurement methodology is capable of accurately capture the visibility associated to TTs as well as identifying any change it may suffer along time.

3. METHODOLOGY TO CHARACTERIZE THE VISIBILITY OF LOCAL TTS

In this section we present a methodology to characterize the visibility of Local TTs in a country and compare it with that offered by TTs in other countries. For this purpose we define three meaningful metrics named *net-visibility*, *potential-visibility* and *potential-online-visibility*.

3.1 A first look at TTs visibility within a country

Let us use the temporal metrics defined in Section 2.2 to make a first comparison of the visibility granted by TTs across different countries.

Figure 3 shows the distribution of the *total active time* for TTs in each one of the 35 countries of our Local-TT-2013 dataset. Each distribution is represented in the form of a boxplot where the box shows the 25, 50 and 75 percentiles of the distribution and the whiskers indicate the 5 and 95 percentiles, respectively. Note that any boxplot used in the rest of the paper presents this same information unless otherwise stated.

We observe that there is an important variability in the *total active time* for the TTs within a country. We will address this issue in Section 5. Of more interest for this section is the significant difference among the distribution of *total active time* for different countries[4]. In particular, the median value of the *total active time* varies around 2 order of magnitude between 20 min in US and 1000 min in New Zealand (NZ). This observation suggests the presence of well differentiated groups of countries with respect to the visibility provided by TTs.

In order to find these groups we leverage standard clustering techniques. Specifically, we use the following 9 input variables to our clustering algorithm: 25, 50 and 75 per-

[4]Note that we also observe a significant variability for the *age* and the *number of active periods* across countries but we do not present the results due to space limitations.

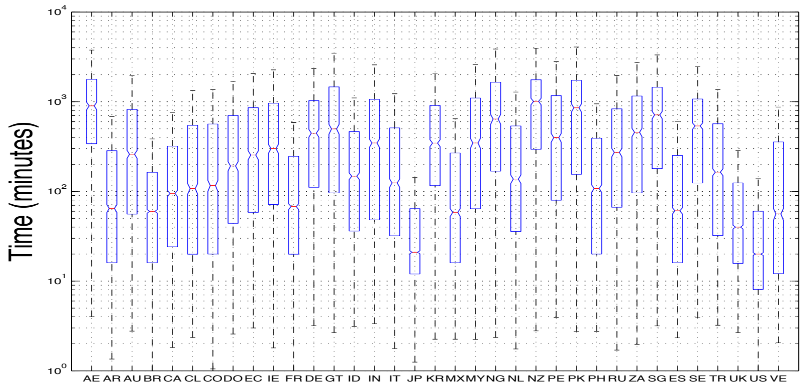


Figure 3: Distribution of Total Active Time for the TTs in each one of the 35 countries of our Local-TT-2013 dataset.

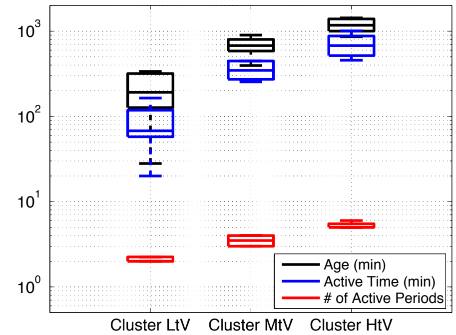


Figure 4: Summary of the distribution of temporal metrics for the HtV, MtV and LtV clusters.

centiles of the *total active time*, the *age* and the *number of active periods* for the TTs of a given country. We use the EM clustering algorithm since it provides as output the optimum number of clusters[5]. This clustering process results in 3 distinct clusters[6]. Figure 4 shows the distribution of the median value of the three temporal metrics (*total active time*, *age* and *number of active periods*) for the countries within each cluster in the form of boxplots. We can observe that the clustering algorithm produced meaningful results since the clusters are perfectly separated and thus represent three different groups that we refer to as: *High temporal Visibility* (HtV), *Medium temporal Visibility* (MtV) and *Low temporal Visibility* (LtV). In particular, the median values for the *total active time* of the HtV, MtV and LtV groups are 700, 350 and 70 min, respectively.

Note that the temporal metrics and, specifically, the *total active time* of a TT captures the *net-visibility* associated to that TT. This is the total time that the TT is visible (or exposed). In the next subsection we develop further the concept of *net-visibility*.

3.2 Net-Visibility

We define a normalized version of the *total active time* to represent the *net-visibility* associated to a TT. We refer to this metric as *net-visibility* (NV) and express it as follows:

$$NV = \frac{log(total\ active\ time)}{log(max(total\ active\ time))}\ \alpha \in [0,1] \qquad (1)$$

where the *max(total active time)* is the duration of our measurement period that is the maximum active time that a TT may have in our dataset. Moreover, the list of TTs shares the bandwidth of the medium (e.g., PC or tablet screen) with other elements like the timeline or the recommendation of users to follow. Then, it is likely that some users do not pay attention to the Trending Topics while browsing through the Twitter interface. The aim of the parameter α in the previous expression is capturing this behaviour.

[5]The EM algorithm follows a cross-validation approach to find the optimum number of clusters [3]. Furthermore, we double-check the correlation between variables to eliminate redundant information in the clustering process.

[6]We have repeated the clustering exercise using EM and different number of seeds and for all cases we always obtain the same optimum number of clusters.

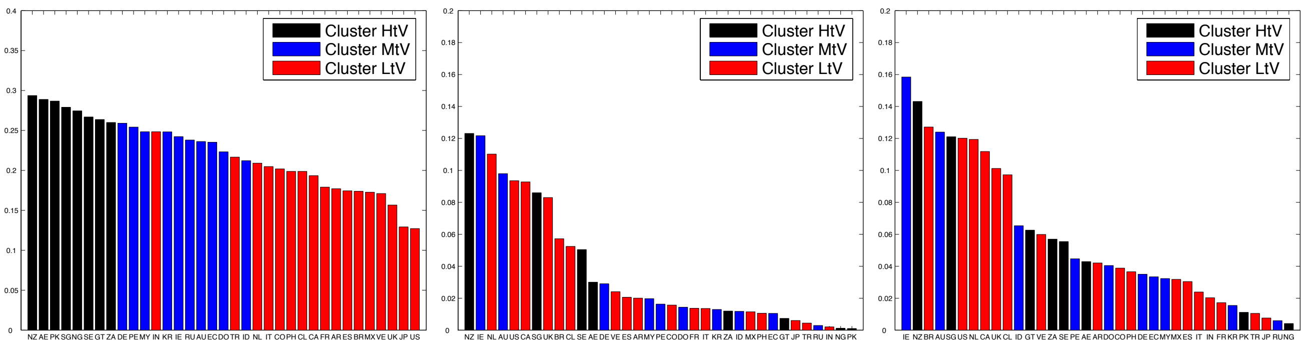
This phenomenon has been well studied in the area of online advertisement where it is refereed to as *banner blindness* [27, 33]. In a recent study, S. Heinz et al. [32] analyze the banner blindness among users who browse a web with exploratory purposes, i.e., not looking for a specific piece of information. This browsing behaviour represents well the typical browsing pattern of Twitter users. The authors quantify the banner blindness through a normalized metric of the *recognition* that captures whether a user reminds or not one (or more) banner(s) that was (were) shown during the browsing session. The value of this metric ranges between 0 (no recognition) and 1 (full recognition). The obtained results indicate that the average recognition for users performing an explorative browsing is 0.51. Given the similarity between the described scenario and ours, we will consider a value of $\alpha = 0.51$ along the paper.

Note that the *net-visibility* for a country is computed as the median of the *net-visibility* of all TTs of that country. We have computed the *net-visibility* for the 35 countries included in our Local-TT-2013 dataset. Figure 5(a) presents a ranking of countries based on their *net-visibility* (from highest to lowest). The results indicate that, as expected, countries within the HtV class present the highest *net-visibility*. Although *net-visibility* is definitely an interesting metric, it does not properly characterize the actual potential visibility offered by a TT since it does not take into account the penetration of Twitter in a country. For instance, the actual visibility granted by TTs in a country with 10K Twitter users and a *net-visibility* of 0.9 may be lower than in a country with 100M Twitter users and a *net-visibility* of 0.1. In the latter case the TTs would be visible for a shorter period of time but are (potentially) exposed to a much larger number of users.

3.3 Potential-Visibility & Potential-Online Visibility

To properly characterize the potential visibility offered by a TT we need to consider both the *net-visibility* and the penetration of Twitter in the country. Toward this end, we have defined a normalized metric that considers these two aspects. We refer to this metric as *potential-visibility* (PV) and it is expressed as follows:

$$PV = NV \frac{\#Twitter\ users}{country\ population} \in [0,1]. \qquad (2)$$



(a) Net-Visibility (b) Potential-Visibility (c) Potential-Online-Visibility

Figure 5: Trending Topics' visibility metrics for the 35 countries in our Local-TT dataset.

where, the fraction term represents the penetration of Twitter in a country. In particular, the #Twitter users is calculated as the % of registered Twitter users in a country (as obtained from our previous work [31]) multiplied by the most recent value of overall registered users reported by Twitter (554M) [15]. Furthermore, the population of each country is obtained from the Mundial Bank statistics [16]. The *potential-visibility* for a country is computed as the median of the *potential-visibility* for the TTs of that country.

We have defined a second valuable metric, the *potential-online-visibility* (PoV). This is a normalized metric that considers the penetration of Twitter among the Internet users of a country rather than among the whole country population. The number of Internet users for a country is also obtained from the Mundial Bank statistics. The expression for the PoV for a TT is the following:

$$PoV = NV \frac{\#Twitter\ users}{\#Internet\ users} \in [0, 1]. \tag{3}$$

Differently from the *potential-visibility*, that characterizes the capacity of TTs to reach the population of a country, this metric captures the capacity of TTs to reach the Internet users of that country. Then, a person or company interested on having online presence would be more interested in this second metric[7]. Furthermore, it is worth noting that by definition the *potential-online-visibility* $\geq$ *potential-visibility* and the equality happens only if the Internet penetration in a country is 100% (i.e., all the citizens from a country have Internet access). Again, the *potential-online-visibility* for a country can be computed as the median of the *potential-online-visibility* for its TTs.

Figures 5(b) and 5(c) present the sorted list (from highest to lowest) of the 35 studied countries based on their *potential-visibility* and *potential-online-visibility*, respectively. These figures allow to easily identify those countries in which TTs have potential to reach a larger portion of the population (Figure 5(b)) and/or the online population (Figure 5(c)). We believe that these metrics are of high interest to evaluate the usefulness of TTs in marketing and advertisement contexts.

As we guessed, the *potential-visibility* (and the *potential-online-visibility*) depicts a quite different picture than the *net-visibility*. For instance, Ireland (IE) that is ranked 14^{th} based on the *net-visibility* occupies the 2^{nd} position based on the *potential-visibility* (1^{st} based on the *potential-online-visibility*). This occurs because despite IE has a medium *net-visibility*, it shows a high Twitter penetration and thus the potential of TTs to reach a higher portion of the population is higher than in most of other countries. We observe the opposite effect for Nigeria (NG) that has the 5^{th} highest *net-visibility*, but due to the low penetration of Twitter in the country, it shows the 2^{nd} lowest *potential-visibility* (the lowest *potential-online-visibility*).

Finally, it is worth to mention that we observe slight variations between the ranking of *potential-visibility* and *potential-online-visibility* metrics for most of the countries. This variability is dictated by the different penetration of Internet in different countries.

4. TRENDING TOPICS VS. TRADITIONAL ADVERTISEMENT CHANNELS

In this section we first introduce the most common metric used to measure the visibility of ads in traditional media and discuss why it is not appropriate to assess the visibility of Trending Topics. Afterwards, we leverage the methodology and metrics described in the previous section to make a comparison of the potential visibility offered by TTs and ads in traditional media.

4.1 Background on assessment of visibility in Traditional Advertisement Channels

There is a standard metric used to measure the visibility achieved by ads in traditional media (e.g., radio-stations, TV channels or newspapers). This metric is named Gross Rating Point (GRP) [29,30] and is expressed as follows:

$$GRP = frequency * reach \tag{4}$$

Where the *reach* and the *frequency* are defined as:

- The *reach* is the ratio between the number of individuals within the target audience (e.g., men over 50) that use the specific media (e.g., a specific radio-station or TV chan-

[7]Note that in many cases advertisement campaigns have a specific target audience. Our metrics can be adapted to those cases. In particular, we would need to change the penetration value considering the estimated number of Twitter users belonging to the target audience and the size of the target audience in the numerator and denominator, respectively.

nel) and the total number of individuals within the target audience.

- The *frequency* is the ratio between the number of views (listenings) of an ad and the number of people who viewed (listened to) that ad. In other words, it indicates the average number of views (listenings) of an ad per user.

On the one hand, the *reach* used in the GRP is exactly the same metric as the *penetration* we use to compute our PV. On the other hand, marketing companies rely on the information provided by audiometers to compute the *frequency* for ads in TV-channels or radio-stations. These are devices installed in houses that monitor the watching (listening) activity of TV (radio-station) users. In the case of newspapers this metric is estimated based on the *Readership*. This is, the number of daily readers of a newspaper. Unfortunately, the frequency is a metric rather difficult to measure for alternative advertisement channels such as Trending Topics. Indeed, there is a controversial debate regarding the suitability of GRP for advertisement in online media [1, 5].

Our PV metric considers the time of exposure of an ad, that is an objective metric (similarly to the frequency), but it can be accurately measured for both traditional advertisement channels (e.g., radio-stations' commercials or newspapers' ads) and alternative ads channels such as TTs. Hence, our PV (contrary to GRP) allows comparing the visibility of traditional and new types of advertisement channels.

4.2 Visibility of Trending Topics vs. Newspapers' ads and Radio-stations' commercials

In this subsection we apply the metrics defined in Section 3 to traditional advertisement channels such as newspapers' ads and radio-stations' commercials and compare their visibility to that offered by TTs for 10 selected countries: Canada (CA), Colombia (CO), Ireland (IE), France (FR), Germany (DE), Guatemala (GT), New Zealand (NZ), Spain (ES), United Kingdom (UK) and United States (US).

Let us focus first on newspapers' ads. We consider full-page ads for our analysis and thus α is equal to 1 because the ad uses all the bandwidth of the medium. For comparison purposes we assume that an ad appears in a newspaper every day over a period equivalent to the duration of our Local-TT-2013 dataset (90 days). Finally, R. Pieters and M. Wedel [39] report that the average time that readers dedicate to an ad in newspapers is 17.26 seconds. In particular, their results are obtained from an experimental study in which they use eye-tracking techniques on a population of slightly more than 3600 users. Using these values we can estimate the average total active time associated to newspapers' ads that would be equal to 17.26 (sec/day)* 90(days) = 25 min and 53 sec. Moreover, the information regarding newpapers' readership is typically available. In particular we have collected that information for some of the most popular newspapers in the countries under consideration[8]. The described data allows us to estimate the *net-visibility* and the *potential-visibility*[9] for popular newspapers of the studied countries.

Now, we consider the example of radio-stations' commercials. Again, α is 1 because radio stations' commercials use all the bandwidth of the medium. We consider the traditional duration of radio-stations' commercials of 60 seconds for our analysis. Note that slots of 15 or 30 seconds are typically offered by radio-stations as well [7, 11]. Furthermore, radio-stations' advertisement campaigns vary between few weeks and few months depending on their goal. Then, for comparison purposes we consider the duration of our dataset (90 days) that is included in this range. Finally, the advertiser has to define a schedule for the ad. This is, the number of used slots per day and time-frames associated to those slots (morning, afternoon, evening or night). To this end, advertisement companies indicate that an ad should be listened at least 3 or 4 times by a person in order to be sure that he/she got the message [10, 13]. Hence, they use this reference value to define the most suitable schedule for each specific campaign. In this paper, we consider an aggressive campaign in which the ad is played three times in every time-frame (12 times a day) so that the probability of people listening to it 4 times is high.

We can use the previous data to estimate the *total active time* associated to a radio-station's commercials as 60 (sec/commercial) *12 (commercials/day)*90 (days) = 1080 minutes. Furthermore, the audience of some of the most popular radio-stations in the considered countries is publicly available[10]. Hence, with the described data we can compute our visibility metrics for those radio-stations.

The computed *net-visibility* for radio-stations' commercials and newspapers' ads is 0.5927 and 0.2760, respectively. Comparing these results with the median *net-visibility* of TTs for the 35 countries shown in Figure 5(a) we observe that radio-stations' commercials present a significantly higher *net-visibility* that TTs in all the 35 countries. Furthermore, TTs offer a slightly higher *net-visibility* than newspapers' ads in only 3 countries: New Zealand (NZ), Arab Emirates (AE) and Pakistan (PK). Hence, we conclude that ads in traditional media enjoy longer exposure times than Trending Topics.

However, as indicated in Section 3 the *potential-visibility* is a more accurate metric since it takes into account the penetration of the specific media in the country. Figure 6 shows the *potential-visibility* associated to popular radio-stations' commercials and newspapers' ads as well as the median *potential-visibility* of TTs for the 10 considered countries. We observe that the *potential-visibility* depicts a different picture than the *net-visibility* due to the different penetration of Twitter, newspapers and radio-stations in these countries. In particular, radio-stations' commercials, Trending Topics and newspapers' ads show the highest *potential-visibility* in 5 countries (IE, FR, DE, ES and UK), 4 countries (CA, CO, NZ and US) and 1 country (GT), respectively. Moreover, in all countries, excepting Guatemala, Trending Topics show a higher *potential-visibility* than newspapers' ads. These results, indicate that despite having a lower exposure time, the higher penetration of Twitter compared to traditional media makes that Trending Topics have a higher potential visibility than radio-stations' commercials in several countries and newspapers' ads in almost every considered country.

[8] The references to the sources from where we obtained the data for the different newspapers' readership are available in our Technical Report [28].

[9] The *potential-online-visibility* does not make sense in this case since we are not considering online media.

[10] Again, the references to the sources from where we collect the data are available in our Technical Report [28].

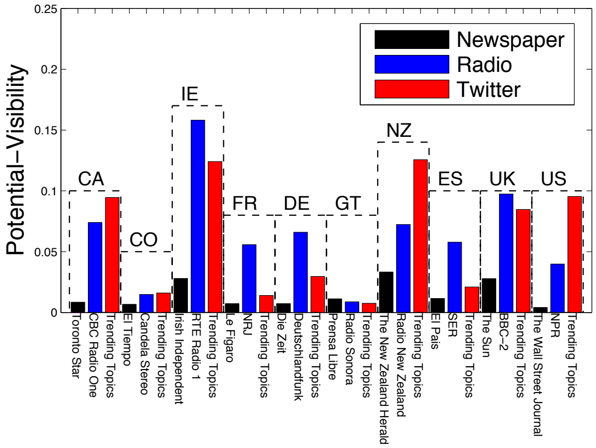


Figure 6: Potential-visibility for TTs, radio-stations' commercials and newspapers' ads for the 10 considered countries. The x-axis presents the names of the studied media for each advertisement channel and country.

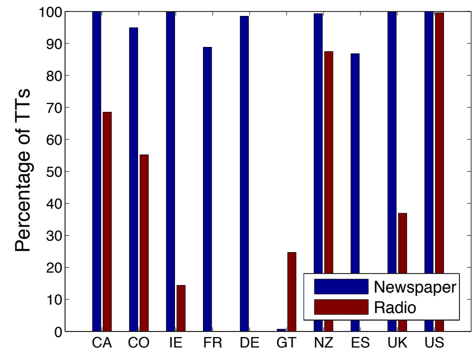


Figure 7: Percentage of TTs with higher *potential-visibility* than newspapers' ads and radio-stations' commercials for the 10 considered countries.

Therefore, we conclude that Trending Topics offer a visibility comparable to other traditional ad channels for the analyzed countries. This confirms that Trending Topics are a useful tool for marketing and advertisement purposes.
However, several considerations should be taken into account with respect to our results. First, Trending Topics should not be considered as a substitute to traditional ads channels. Instead, they should be considered a complementary tool in advertisement and marketing contexts. In particular, in traditional ad channels the advertiser buys several slots and it has the certainty that its product would be shown to the audience during those slots. However, the same advertiser may lunch a marketing campaign in Twitter but it has not the guarantee that its product will become Trending Topic. In fact, the definition of strategies to help companies to generate Trending Topics is still an unsolved matter and requires further research[11]. Second, our analysis has been conducted under certain assumptions. For instance, we have only considered popular radio-stations and newspapers in each country with a higher penetration than the average newspapers or radio-stations in those countries. Furthermore we have considered values that represent realistic advertisement campaigns in newspapers and radio-stations, but other type of campaigns are possible and may lead to different visibility results. Finally, some other subtle aspects such as how the ad support (e.g., audio vs text vs images) affects the attention of the user have not been considered.

5. ANALYSIS OF THE VARIABILITY OF TTS VISIBILITY WITHIN A COUNTRY

As Figure 3 revealed, there exists a notable difference, over an order of magnitude, in the *total active time* across Local TTs in a country. Hence, distinct TTs within a country enjoy rather different visibility. In this section we dig into this difference. In particular, we conduct the following analyses: *(i)* we present a methodology whose aim is to unveil which type of TTs are likely to provide a higher visibility and *(ii)* we study whether the visibility offered by TTs at different times of the day presents an identifiable daily pattern. Due to space limitation, we present the obtained results for three selected countries. Specifically, we have chosen one country from each one of the temporal-visibility groups defined in Section 3 to guarantee the diversity in our selection: New Zeland (NZ) from the HtV group, Ireland (IE) from the MtV group and UK from the LtV group. Note that we will refer to results for other countries when warranted.

Before going into those analyses, we would like to briefly extend our comparison between TTs and traditional advertisement channels. In the previous section we have used the median value of the different visibility metrics of TTs within a country to perform the comparison study. However, due to the high variability of TTs visibility in a country, we would like to present more statistically meaningful results. To this end, we have computed the percentage of TTs that present a higher *potential-visibility* than radio-stations' commercials and newspapers' ads for each one of the 10 countries analyzed in Section 4. Figure 7 shows the obtained results. First, at least 85% TTs present a higher *potential-visibility* than newspapers' ads in all countries but Guatemala in which due to the high penetration of the considered newspaper only the top 1% most visible Trending Topics would achieve a higher visibility than ads in that newspaper. Second, in the case of radio-stations' commercials we observe a high variability in the results. For instance, in FR, GE and ES the visibility of commercials in the considered radio-stations' is higher than for any TT whereas in US we observe the opposite effect, 99% TTs enjoy more visibility than commercials in the considered radio-station. This variability is dictated by the interplay of the penetration of different media as well as the associated *net-visibility*.
In summary, these results confirm the conclusion from Section 4: Trending Topics offer a visibility comparable to traditional ad channels and then they are useful as a tool in marketing and advertisement contexts. However, the high variability observed in the visibility of TTs across (and within) countries requires to conduct an individual analysis for each specific case to obtain accurate results.

Finally, we would like to highlight that in order to study the variability of TTs visibility within a country we use the

[11]Note that companies have the option of purchasing Promoted Trending Topics that follow a similar business model (pay-per-slot) as traditional ad channels.

CLASS	CATEGORY	EXAMPLE
Hashtags	— (Self)	#FirstQuestionsAsked
Sports-Related	Athlete (DBp)	Andrew Sheridan
	Sport Events (Self)	Chelsea - Liverpool
	Other Sport Issues (DBp)	Conmebol
	College Coach (DBp)	Mike McQueary
Feelings & Emotions	Positive (Self)	Happy Birthday Britney
	Negative (Self)	RIP Perez Hilton
Places &Buildings	Architectural Structure (DBp)	US Capitol
	Administrative Region (DBp)	Northern Iowa
	Feature (DBp)	La Cartuja
	City (DBp)	Amsterdam
	Populated place (DBp)	Cannes
Celebrities	Agent (DBp)	Hugh Grant
	Office holder (DBp)	Bill Clinton
	Politician (DBp)	John McCain
	Artist (DBp)	Freddie Mercury
	Famous person (IMDb)	Christine Reyes
	Writer (DBp)	Edgar Allen Poe
Entertainment	Album (DBp)	Love After War
	Movies (IMDb)	Finding NEMO
	Book (DBp)	Geek Love
	Work (DBp)	Reservoir Dogs
	Character (IMDb)	Batman & Robin
	Video games (DBp)	Death Race
	Film (DBp)	Celda 211
	Single (DBp)	Bad Romance
	TV show (DBp)	American Idol
Companies	Organization (DBp)	RTVE, Spanair
	Privately held company (DBp)	Twitter
	Public company (DBp)	Jackson Hewitt
Others	Diseases (DBp)	HIV
	First name (DBp)	Danielle
	Wide-body aircraft (DBp)	Boeing 767
Unclassified	—	Take Facebook Down

Table 2: List of Semantic classes and categories. For each category we also indicate the source as follows: *DBp* for DBpedia, *IMDb* for IMDb and *Self* for Self-defined categories.

total-active-time in the rest of the section. Note that *net-visibility* is a normalized version of this metric and Twitter penetration, used to compute the *potential-visibility*, is the same for all TTs within a country. Then, results derived with the *total active time* and these other metrics are equivalent.

5.1 Visibility of different semantic classes of TTs

In this subsection we first define an efficient methodology to group TTs by their semantic meaning into different semantic classes. Then we apply this methodology to the Local TTs of the selected countries. Finally, we compute the distribution of the *total active time* for the TTs within each semantic class so that we can report what types of TTs offer higher visibility in each country.

5.1.1 Methodology

Our tool uses the following sources in order to assign a specific TT to a semantic category:

- *DBpedia* is a sub-project of Wikipedia that aims to create an ontology to classify different names, terms, words and expressions available in Wikipedia pages [2]. In particular, it provides a hierarchical ontology that currently covers 359 semantic categories that are described by one or more properties from a pool of 1775.

- *IMDb* is a popular database including information related to a large number of entertainment resources such as movies, TV shows, actors/actresses, etc [8]. Contrary to DBpedia, IMDb does not provide a structured classification for the stored resources.

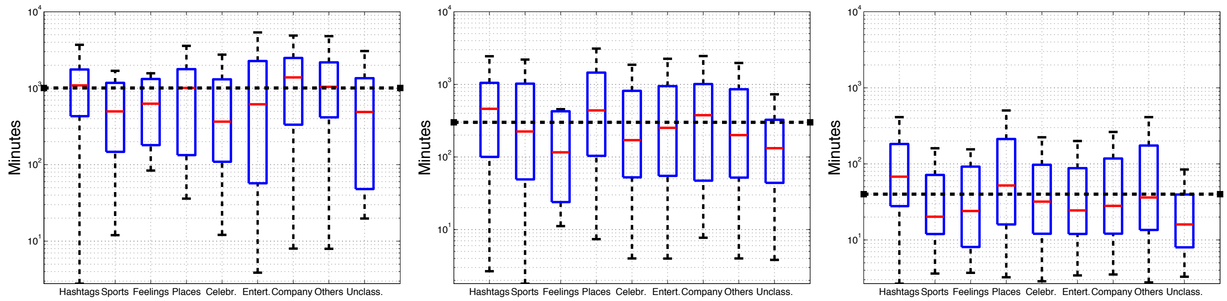
- *Self-defined categories*: Manual inspection of TTs reveals some common semantic categories that although easily identifiable for a human being are not recognized by either DBpedia or IMDb. In particular, we identify two of these categories: (i) *Sport Events*, our manual inspection reveals that TTs are commonly used to reflect events related to different sport games, such as the score of football games. Examples of this are TTs such as 'Arsenal 1-2 Manchester United' or 'Gol de Benzema'. (ii) *Feelings/Emotions*, our manual inspection also suggests that TTs are used to express emotions, feelings, preferences, greetings, etc. Therefore it is common to find TTs including words such as 'Happy', 'Love' or 'Hate'. Examples of these TTs are 'Happy Birthday Andy Carroll' or 'We Love Hunger Games'. Therefore, our tool classifies those TTs that include one (or more) emotion-related word(s) and neither DBpedia nor IMDb are able to classify in the *Feelings/Emotions* class.
Moreover, we consider *Hashtags* as a separate category. As indicated in the Introduction hashtags are a special functionality of Twitter that is widely used and thus understanding whether they offer a higher/lower visibility than "non-hashtaged" topics is of high interest for commercial and advertisement purposes.

The large number of potential output categories provided by DBpedia and the lack of structure of IMBb would make infeasible to conduct a meaningful analysis of the semantic context of TTs using their provided results. To address this issue, we have performed a careful merging process in which we group semantic categories obtained from DBpedia, IMDb and our self-defined categories into a handful set of semantic *classes* that permits us to present a meaningful discussion. Note that for this process we have used as reference the 18 classes defined in [36]. Indeed, the 18 classes defined in [36] can be easily merged into the 9 classes resulting from our process (with the exception of hashtags). We have decided to define a smaller number of classes because using 18 classes results in few of them being scarcely populated.
Table 2 lists the defined semantic classes and, for each class, presents the most important categories along with its original source (i.e., DBpedia, IMDb or self-defined categories). In particular, we use the following preference order in our semantic classification process for a given TT: we first try to classify it using DBpedia in a semantic category and class. If DBpedia fails we use IMDb and in case it also fails we use our Self-defined categories. Those topics that are not classified after these three steps are added to the *Unclassified* class. Finally, our manual inspection of the TTs within the Unclassified class reveals that most of these topics correspond to complex sentences similar to some hashtags but without the initial '#'. Some examples are: 'Tomorrow is Friday', 'Bieber Fever Is Incurable', 'Ian Is Our Pride', 'M or P' and 'Lin is 6'. It can be noticed that some of them are difficult to be semantically classified even for a human being without the required context knowledge (e.g., 'M or P').

5.1.2 Performance Evaluation

We have used the described methodology to classify the TTs included in our datasets. Table 3 summarizes the percentage of TTs that have been classified as well as those that our tool is unable to classify for each analyzed country (Unclassified). The results suggest that our tool is fairly efficient since it is able to automatically classify more than 90% of the TTs in the worst considered case (UK).

However, the effectiveness of a classification tool is not measured by the percentage of resources that it is able to classify but the percentage that it is able to classify correctly.



(a) Active Time NZ (b) Active Time IE (c) Active Time UK

Figure 8: Distribution of the *active time* across TTs within each semantic classes for NZ, IE and UK (the horizontal dashed line shows the median active time of the all the Local-TTs of the correspondent country).

In particular, we define two types of errors for our classification tool: (*i*) *false positives* are those TTs that our tool assigns to a wrong class and (*ii*) *false negatives* are those TTs that our tool was unable to classify but a human being would be able to classify in any of the defined semantic classes.

The detection of false positives and negatives needs to be done manually. Note that this is a common practice used in previous works [36, 42]. Conducting such an experiment for all the TTs from our dataset is a very tedious and time consuming task. Therefore, we have selected a random set of 1000 TTs and three different persons[12] have manually detected the false positives and negatives for this subset of TTs. Note that the differences between the classification done by these three persons over the same random set varies less than 1%. This suggests that the error introduced by human beings is negligible and thus the result of the manual classification can be considered a good approximation to the ground truth. In addition, sampling introduces an error in the proportion of Trending Topics per category used during the validation with respect to the actual proportions. This error can be computed using a hypothesis test for a proportion [41]. This is a well-known tool widely used to compute confidence intervals for the results of surveys. In particular, in our case in which we use a sample of 1K TTs, the error introduced by sampling in the proportion of Trending Topics in any class is $\leq 3.1\%$ (with 95% confidence) for any size (i.e., number of Trending Topics) of the dataset. This suggests that: first, the obtained results are reasonably accurate and, second, the used methodology scales well since manually inspecting a sample of 1000 TTs (that as we have demonstrated is doable for a human being) suffices to not incur in high errors in the considered proportions for different classes.

Our detection experiment reveals that, one the one hand, 41% of the unclassified TTs are false negatives. Since the *Unclassified* class represents less than 10% of our TTs, we conclude that overall only around 4% of the TTs corresponds to false negatives. On the other hand, false positives are also infrequent and represent only 5% of the inspected TTs. In a nutshell, these results indicate that our semantic classification tool is quite accurate and its automatic process is able to classify more than 91% of the TTs correctly.

[12]These three persons were not connected to our research project to guarantee the objectivity.

	NZ	IE	UK
Hashtags	53,72%	47,31%	39,13%
Sports-Related	3,46%	5,96%	10,80%
Feeling & Emotions	0,80%	0,84%	0,76%
Places & Buildings	3,86%	7,39%	4,58%
Celebrities	7,45%	9,79%	14,19%
Entertainment	8,64%	7,58%	8,85%
Companies	3,59%	3,95%	3,06%
Others	13,70%	11,67%	8,99%
Unclassified	4,79%	5,51%	9,65%

Table 3: Distribution of Local TTs from UK, IE and NZ across the defined semantic classes.

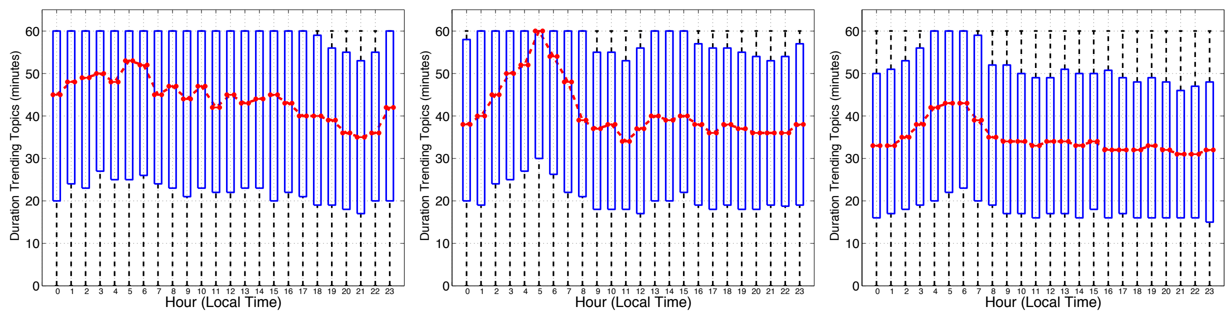
5.1.3 Visibility of TTs across semantic classes

Figure 8 depicts the distribution of the *total-active-time* for every semantic class of the three analyzed countries in the form of boxplot. In addition, we plot a horizontal dashed line that indicates the median *total-active-time* for all TTs in the country for reference.

First of all we observe a high variability among the visibility offered by different TTs within each class. Despite this variability, we still can derive useful observations. For instance, "Hashtags" and "Places" are the only two cateogries whose median *total active time* is above the median value of the country, for all three countries. Interestingly, this result along with results in Table 3 suggest that adding a # in front of the term to be advertised seems to increase the chances to become TT and to enjoy a longer active time. Surprisingly, categories such as "Sport" and "Celebrities" that attract a fair amount of attention from media do not appear among those offering higher visibility. This may indicate that Twitter users do not get excited about these topics for long time. Finally, we observe differences across countries that indicate that each *national market* shows preferences for different types of topics. For instance, TTs related to companies present the highest visibility in NZ whereas TTs in this category show a rather low active time in UK. Furthermore, TTs related to "Sports" present a quite low visibility in NZ and UK but not in IE.

5.2 Daily Pattern of Trending Topics Visibility

Internet traffic as well as most on-line services present a daily usage pattern bound to the daily schedule of their users [40]. In this subsection we focus on understanding whether the visibility offered by Local TTs presents an identifiable



(a) Active Time NZ (b) Active Time IE (c) Active Time UK

Figure 9: Distribution of the *total active time* of TTs instances within each one of the 24-hour slots of a day.

daily pattern. For this purpose, we divide a day in its 24 one-hour slots[13] and for each slot we calculate the distribution of the *active time* for the TT instances present in that slot. Note that the maximum *active time* that a TT instance can have in a slot is 60 minutes.

Figure 9 shows the obtained results for UK, IE and NZ. The x-axis shows the 24 time slots described in the previous paragraph and the y-axis shows the distribution of the *active time* of the TTs present in each time slot in the form of boxplot. Note that the time slots represent local time for each country. We observe that there is a marked daily pattern in the distribution of the *active time* for the different hour-slots. In fact, for every country we can see the presence of few slots where TTs tend to have a higher *active time*. Specifically, these slots correspond to the night (sleeping) hours in which a lower activity of Twitter users helps TTs to remain visible longer time. However, the higher *net-visibility* enjoyed in those hours does not really lead to a higher *potential-visibility* since the number of users connected to Twitter at those hours is likely to be significantly smaller than in the morning, afternoon or evening. We have repeated this experiment for the 35 countries in our dataset. The results can be found in our TR [28]. In summary, most of the countries show the previously reported daily pattern, with few exceptions such as Japan, US and some Latin-American countries (e.g., Colombia or Venezuela), in which we observe a flatter shape. Thus, the difficulty of getting a TT in these countries is independent of the time of the day. Finally, we have separately studied the daily-pattern for week days and weekends for every country without noticing major differences.

6. RELATED WORK

Measurement and Analysis of Trending Topics: Kwak et al. [35] performed the most exhaustive characterization of Twitter so far. As part of this study the authors briefly analyze Trending Topics using coarse temporal metrics and quantitative metrics to classify Trending Topics in few externally defined (i.e., artificial) categories. Furthermore, Asur et al. [26] use quantitative metrics to analyze the formation, persistence and decay phases of Trending Topics. Both works rely on quantitative metrics that, as shown by Morstatter et al. [38], may lead to unreliable results due to the best effort nature of Twitter APIs. Finally, Huang et al. [34] studied the differences between the tagging pattern in Twitter and other OSN systems. The authors present the phenomenon of the Twitter *micro-meme*: emergent topics for which a tag is created, used widely for a few days and then disappears. Although these papers provide initial valuable results, they focus on specific aspects of Trending Topics different to the one addressed in our paper, i.e., the characterization of the visibility offered by TTs in different countries.

Semantic classification of Trending Topics: Lee et al. [36] use a dataset formed by around 800 Trending Topics and classify them into 18 different categories using a text- and a network-based methodologies that achieve an accuracy of 65% and 70%, respectively. In our study we consider a set of Trending Topics 3 order of magnitude larger. Furthermore Zubiaga et al. [42] assign 15 different properties to Trending Topics (including some unreliable quantitative properties) to classify them into 4 classes using a similar text-based methodology as the one used in [36]. They validate their technique using a training and a test sets with 600 and 436 Trending Topics, respectively. In this case they report an accuracy of 78.4%.

7. CONCLUSION AND FUTURE WORK

Despite Trending Topics are a well-know feature regularly exploited in the context of marketing and advertisement, we still stand on preliminary ground in terms of understanding this tool. In this paper we characterize the visibility of Trending Topics across 35 countries. In particular, we present a measurement methodology along with a methodology to thoroughly analyze the visibility of Trending Topics that we believe can be of high value for experts of different disciplines in marketing and advertisement contexts. The results obtained applying these methodologies indicate that, in general, Trending Topics present a comparable visibility to other traditional advertisement channels and thus they can be considered a useful tool in marketing and advertisement contexts. However, the high variability on the visibility offered by Trending Topics across (and within) countries suggests that we should apply the described methodology to obtain accurate results for each specific case.

As future work we plan to apply our methodology to online advertisement in order to compare the visibility offered by TTs with that offered by other online media such as ban-

[13] Slot 0 includes information for the 60 minutes between 12AM and 1AM, slot 1 includes information for the 60 minutes between 1AM and 2AM and so on.

ners in popular websites. Furthermore, we will explore different strategies that companies may use to create Trending Topics in Twitter as well as their associated costs in comparison with traditional advertisement channels.

8. ACKNOWLEDGEMENTS

We would like to thank our shepherd Balachander Krishnamurthy and anonymous reviewers for their valuable feedback. The research leading to these results has been partially funded by the European Union's FP7 Program under the project eCOUSIN (318398), the Spanish Ministry of Economy and Competitiveness under the eeCONTENT project (TEC2011-29688-C02-02), and the Regional Government of Madrid under the MEDIANET project (S2009/TIC-1468).

9. REFERENCES

[1] Are gross rating points really the answer for digital? http://www.imediaconnection.com/content/32278.asp.

[2] DBpedia. http://dbpedia.org/About.

[3] EM - Weka 3. http://weka.sourceforge.net/doc/weka/clusterers/EM.htmls.

[4] Facebook friends Twitter-like Trending Topics. https://www.cnbc.com/id/100942887.

[5] Gross Rating Point Metrics Will Be Good for Online Advertising. http://www.mediapost.com/publications/article/155743/.

[6] HBR Blog Network. http://blogs.hbr.org/cs/2013/05/surprise_is_still_the_most_powerful.html.

[7] How Long is a Radio Ad. http://suite101.com/article/how-long-is-a-radio-commercial-a152539.

[8] Internet Movie Database (IMDb). http://www.imdb.com/.

[9] Marketing with Twitter. https://business.twitter.com/marketing-twitter.

[10] Optimum Scheduling for Radio Advertising. http://avenueright.com/entries/71/optimum-scheduling-for-radio-advertising-frequency-is-key.

[11] The perfect length for a radio commercial. http://danoday.com/blog/2010/11/radio-commercial-length/.

[12] Promoted Trends | Twitter for Business. https://business.twitter.com/products/promoted-trends.

[13] Radio Ad Frequency. http://www.marketingprofs.com/ea/qst_question.asp?qstid=6388.

[14] Spanish revolution. https://wiki.digitalmethods.net/Dmi/DmiSummer2011SpanishRevolution.

[15] Statistic Brain. http://www.statisticbrain.com/twitter-statistics/.

[16] The World Bank. http://www.worldbank.org/.

[17] Twitter API Documentation. https://dev.twitter.com/.

[18] Twitter Blog: Let's Fly. http://blog.twitter.com/2011/12/lets-fly.html.

[19] Twitter blog: To Trend or Not to Trend... http://blog.twitter.com/2010/12/to-trend-or-not-to-trend.html.

[20] Twitter marketing agency. http://twittermarketingagency.com/.

[21] Twitter marketing company. http://www.outsourcegeek.com/marketing-services/twitter-marketing-service/twitter-marketing-company.

[22] Twitter promoted trends tracked for one month. http://allthingsd.com/20130409/big-media-loves-promoted-trends-twitters-big-dollar-digital-billboards/.

[23] Twitter Rate Limiting. https://dev.twitter.com/docs/rate-limiting.

[24] Twitterholic Politicians. http://litteramedia.wordpress.com/2011/12/13/twitterholic-politicians/.

[25] UMD Web Site. http://www.umassd.edu/cmr/studiesandresearch/2011fortune500/.

[26] Sitaram Asur, Bernardo A Huberman, Gabor Szabo, and Chunyan Wang. Trends in social media: Persistence and decay. In *ICWSM*, 2011.

[27] Jan Panero Benway. Banner blindness: The irony of attention grabbing on the World Wide Web. In *HFES Annual Meeting*. SAGE Publications, 1998.

[28] J. Carrascosa, R. Gonzalez, R. Cuevas, and A. Azcorra. Are Trending Topics Useful for Marketing? Visibility of Trending Topics vs Traditional Advertisement. Technical report available at: http://www.it.uc3m.es/~rcuevas/techreports/TT_TR2013.pdf, Universidad Carlos III de Madrid, 2013.

[29] John A Davis. *Measuring Marketing: 110+ Key Metrics Every Marketer Needs*. Wiley. com, 2012.

[30] Paul W Farris, Neil T Bendle, Phillip E Pfeifer, and David J Reibstein. *Marketing metrics: The definitive guide to measuring marketing performance*. Pearson Education, 2010.

[31] Roberto Gonzalez, Ruben Cuevas, Angel Cuevas, and Carmen Guerrero. Understanding the locality effect in Twitter: measurement and analysis. *Personal and Ubiquitous Computing*, 2013.

[32] Silvia Heinz, Markus Hug, Carina Nugaeva, and Klaus Opwis. Online ad banners: the effects of goal orientation and content congruence on memory. In *CHI*, 2013.

[33] Guillaume Hervet, Katherine Guérard, Sébastien Tremblay, and Mohamed Saber Chtourou. Is banner blindness genuine? Eye tracking internet text advertising. *Applied Cognitive Psychology*, 2011.

[34] Jeff Huang, Katherine M Thornton, and Efthimis N Efthimiadis. Conversational tagging in twitter. In *21st ACM conference on Hypertext and hypermedia*, 2010.

[35] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is Twitter, a social network or a news media? In *WWW*, 2010.

[36] Kathy Lee, Diana Palsetia, Ramanathan Narayanan, Md Mostofa Ali Patwary, Ankit Agrawal, and Alok Choudhary. Twitter trending topic classification. In *ICDMW*, 2011.

[37] Hubert W Lilliefors. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 1967.

[38] Fred Morstatter, Jurgen Pfeffer, Huan Liu, and Kathleen M Carley. Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. 2013.

[39] Rik Pieters and Michel Wedel. Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing*, pages 36–50, 2004.

[40] Kevin Thompson, Gregory J Miller, and Rick Wilder. Wide-area Internet traffic patterns and characteristics. *Network, IEEE*, 1997.

[41] Mario F Triola, William Martin Goodman, Gerry LaBute, Richard Law, and Lisa MacKay. *Elementary statistics*. Pearson/Addison-Wesley, 2006.

[42] Arkaitz Zubiaga, Damiano Spina, Víctor Fresno, and Raquel Martínez. Classifying trending topics: a typology of conversation triggers on twitter. In *CIKM*, 2011.